Context-Dependent Trading of Binaural Spatial Cues in Virtual Reality

By

Travis M. Moore

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
in
Hearing and Speech Sciences
December 15, 2018
Nashville, Tennessee

Approved:
G. Christopher Stecker, Ph.D.
Frederick J. Gallun, Ph.D.
Benjamin W. Y. Hornsby, Ph.D.
Erin M. Picou, Au.D., Ph.D.

ACKNOWLEDGMENTS

Despite the hurdles along the way, looking back over the past four-and-a-half years brings about an unexpected calmness. Part of this feeling is a sense of accomplishment, but another facet stems from the enduring support I have received. My sincerest gratitude goes out to everyone who has contributed to this study and to the advancement of my graduate training.

Bethany: The past few years have both dragged and flown by, but you have been a constant throughout all of it. Thank you for your hard work supporting the two of us financially while I synthesized weird sounds in MATLAB for all these years. I have leaned heavily on your limitless patience and unwavering support and you did not falter. Your data are always the best data.

Chris: Thank you for welcoming a second-year graduate student into your lab, largely sight-unseen. Without your generosity I would not be where I am today. Your calmness and adaptability have greatly impacted my approach to science, teaching and mentoring. Trying to emulate the breadth and depth of your expertise has been both humbling and inspiring. Thank you for your guidance, and your example.

Erin: You have run the gamut of roles with me during my time at Vanderbilt: student colleague, TA, lab mentor, landlord, employer, instructor, collaborator, coauthor, and co-instructor. There is not enough room to thank you for all of your cheerful guidance, and your most important role as friend. So much of so many things would have fallen apart without your constant support.

Ben: It has been ten years since my first day in your psychoacoustics course, and you have not stopped teaching me since. Thank you for sharing your wisdom with me for so long. I think it’s fair to say you had the most challenging task in turning a German major into a scientist. I know it wasn’t easy, but lucky for me, easy is not your style.

Erick: Thank you for agreeing to serve on my dissertation committee with little advance warning and only a cursory explanation of the project. I am grateful for your insightful comments, which you always had despite the unfavorable conditions created by video conferencing. Your expertise improved this project, and I hope we collaborate again.

TABLE OF CONTENTS

Page

Chapter

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

HMD	head-mounted device
ILD	interaural level difference
$\mathrm{ILD}_{\mathrm{adj}}$	adjusted value of ILD to offset a fixed ITD
ISI	interstimulus interval
ITD	interaural time difference
$\mathrm{ITD}_{\mathrm{adj}}$	adjusted value of ITD to offset a fixed ILD
ITI	intertrial interval
JND	just noticeable difference
MOA	method of adjustment
MOCS	method of constant stimuli
TR	trading relation
VR	virtual reality

Chapter 1

INTRODUCTION

1.1 Background

Initial reports suggested that human perception of sound source laterality depended solely on the differences in level of a signal at the two ears (e.g., Rayleigh, 1875); however, this notion was amended in favor of the duplex theory of sound localization (Rayleigh, 1907), which has persisted for more than a century. The duplex theory states that sound source localization in the horizontal plane is determined primarily by two frequency-dependent binaural cues: the difference in level at the two ears (interaural level difference; ILD), and the difference in the arrival time, or phase, of sound at the two ears (interaural time difference and interaural phase difference; ITD and IPD). Specifically, Rayleigh suggested the ILD is more effective for high frequencies, above about 1.5 kHz, where the head is able to attenuate short wavelengths and create a level imbalance between the two ears, with the more intense side indicating the origin of the sound source. Conversely, frequencies below about 1.5 kHz are able to bend around the head, and therefore provide much weaker ILDs, if any. Instead, the comparatively longer wavelengths below about 1.5 kHz provide ITDs that indicate laterality based on arrival time, with perceived azimuth drawn toward the earlier side. As will be discussed below, subsequent work has provided a more comprehensive, but ultimately incomplete understanding of the relationship between these two binaural cues.

1.1.1 Binaural cues for azimuthal sound source localization

In an elegant experiment, Mills (1960) put the duplex theory to the test by comparing the just noticeable differences (JNDs) in time and level of a sound-field stimulus to JNDs in level obtained using a dichotic stimulus under headphones, across a range of frequencies (octaves and inter-octaves from 0.25 – 8 kHz, and 10 kHz). Results revealed that the dichotic intensity JND function matched that of the sound-field JND function between 1.5 – 6 kHz, while the functions diverged significantly at frequencies below 1.5 kHz, as predicted by the duplex theory. Complimentary findings were reported by Zwislocki and Feldman (1956), who measured JNDs in dichotic phase as a function of frequency. They showed interaural phase JNDs increased with increasing frequency, becoming too large to measure at approximately 1.3 kHz. Taken together, these studies provide evidence in favor of the frequency-specific nature of ITDs and ILDs for pure tone stimuli.

However, as pure tones rarely exist in natural settings, it is important to examine the nature of binaural cues using more complex stimuli. Wightman and Kistler (1992) asked listeners to judge the perceived locations of wideband and 5 kHz highpass filtered signals. They filtered all stimuli using sound-field-to-eardrum transfer functions measured for each participant, which simulated sound-field cues under headphones (see Wightman & Kistler, 1989 for details of the stimulus generation). Specifically, they manipulated the interaural phase to indicate a static location (0°, -45°, or 90°) and recorded listener responses of apparent location to ILDs (and spectral cues) indicating various other directions. Results showed that listeners' perception corresponded to the ITD for wideband stimuli (which contained both high and low frequencies), regardless of the ILD value. In contrast, the perceived azimuth of the 5 kHz highpass stimuli relied on the ILD. These results support the general duplex theory, and also suggest that ITD is the dominant cue for localization when both high and low frequencies are present in the same signal.

McFadden and Pasanen (1976) measured lateralization accuracy using noise signals of various bandwidths (centered around either 500 or 4000 Hz), as well as two-tone complexes that varied in rate and depth of modulation. They found that as stimulus bandwidth increased, smaller ITDs were able to achieve the same lateralization accuracy as larger ITDs. Interestingly, with a bandwidth of 800 Hz, lateralization performance was the same for stimuli centered around 500 Hz and 4000 Hz. For two-tone complexes, ITD cues were more effective for larger frequency separations (up to a point) and deeper modulation depths. McFadden and Pasanen interpreted their findings as suggesting three types of interaural timing cues: (1) onset time differences (arrival time is earlier at one ear); (2) ongoing time differences (fine structure phase relationship); and (3) envelope time differences (slow envelope fluctuations present in a signal at least 1 ms in duration). The existence of envelope differences in complex signals invalidates the frequency specificity of the duplex theory for signals other than pure tones; that is, envelope cues provide timing differences at frequencies above the physiological limits of phase locking.

Hafter and Dye Jr (1983) further studied the effects of rate of change in the stimulus envelope by manipulating the interclick interval (ICI) present in trains of clicks (ICIs tested were 1, 2, 5, or 10 ms). The listeners' task was to indicate the perceived directional shift from one interval to another in a two-interval, forced choice experiment. Results showed that with longer ICIs, most listeners were able to use the ITD present in each click with equal efficiency. However, at shorter ICIs, performance decreased, suggesting ITD information was not fully integrated at high click rates. Findings from Freyman et al. (1997) aid in understanding these results by demonstrating a relationship between onset delay and the degree of ambiguity of ongoing cues. Specifically, Freyman et al. (1997) showed that lateralization is determined largely by onset cues when ongoing cues are ambiguous (as in the short ICI trains of Hafter and Dye, 1983). Conversely, if the ongoing cues are salient then onset dominance does not play a major role in lateralization. Macpherson and Middlebrooks (2002) confirmed several of the findings already discussed, as well as adding that monaural spectral cues have little or no influence on perceived azimuth (see also Hofman et al., 1998).

Thus, the duplex theory provided early insight into the relationship between ITD and ILD cues for pure tone stimuli, but fails to account for the observed data involving complex stimuli. While subsequent work has accounted for several behavioral phenomena in violation of the duplex theory, the relationship between interaural differences in time and level have yet to be fully understood. For instance, the seemingly straightforward question of determining the ITD required to offset an opposing ILD has yielded a variety of results that depend on a number of factors (e.g., David Jr et al., 1959; Hafter & Carrier, 1972; Ignaz et al., 2014; McFadden et al., 1972; Young Jr & Levine, 1977). The discrepancies in perceived azimuth when directly setting one binaural cue against the other represents a fundamental unexplained phenomenon in the processing of binaural cues for spatial hearing, and is the focus of the current study. The sections that follow discuss the literature pertaining to binaural interaction and propose a novel research study to address an existing gap in the current knowledge base.

1.1.2 Trading of interaural time and level differences

A direct method used to investigate the relationship between ITDs and ILDs is to set the cues into opposition; that is, to create a time delay favoring one ear, and a level difference favoring the opposite ear. According to Blauert (1997, p. 165), the first reports of cue trading were Klemm (1920) and (Wittmann, 1925). Klemm (1920) positioned a listener between two telephones that had been modified to produce identical intensity and timber. Among various other binaural hearing experiments, the following scenario and experimental question emerged:

> …ließ sich ein Intensitätsunterschied herstellen, unter dessen Wirkung das subjektive Hörfeld sicher auf die Seite des stärkeren Schalls hinüberrückte. Läßt sich nun dieser Einfluß des Intensitätsverhältnisses durch einen entgegengesetzten Zeitunterschied so ausgleichen, daß das Hörfeld wieder in die Mitte rückt? (pg. 130)

> …a level difference was introduced, with the effect that the subjective auditory sensation was clearly pushed over to the side of the stronger sound. Does this intensity relationship balance out with the introduction of a time difference in the opposite direction, in such a way that the subjective sensation is pushed back to the center? (Translation: Travis Moore)

Classically, the unit of measure of the time or level difference required to offset the complimentary cue has been in ratio form: µs/dB. Shaxby and Gage (1932) coined the term *trading ratio*, which is commonly used in the literature, and reported a value of 1.7 µs/dB. That is, listeners required 1.7 µs of right-leading ITD per decibel of left-favoring ILD to center an intracranial image. Because this terminology assumes a linear relationship between ITD and ILD effectiveness, and that assumption can be violated, this document uses the term trading relation, after Lang and Buchner (2008). The section below discusses several methods used to measure trading relations, followed by a review of the major findings.

1.2 Psychophysical paradigms in the cue-trading literature

1.2.1 Centering

While results of trading studies are often described in µs/dB, there are a variety of methods to obtain the data. For instance, Shaxby and Gage (1932) introduced a centering method of measuring the equivalence function by asking listeners to adjust the amount of right-leading ITD in a stimulus with a fixed, left-biased ILD until the intracranial image was centered at the midline (see also David Jr et al., 1959; Deatherage & Hirsh, 1959; Harris, 1960). The data were plotted as values of ITD (in µs) along the ordinate as a function of several fixed ILD levels tested along the abscissa. A constant was derived from linear fits of the data that explained the linear relationship between the two cues (a trading *ratio* per se).

1.2.2 Pointing

Moushegian and Jeffress (1959) introduced a type of matching procedure using a target and pointer. Experimenters presented a pure tone with a fixed ITD and ILD (the target), while listeners adjusted the ITD of a noise “pointer” until it matched the perceived azimuth of the target (see also Feddersen et al., 1957). They reported a trading relation of 2.5 µs/dB using a 500 Hz pure tone. Whitworth and Jeffress (1961) employed a similar technique, using 500 Hz pure tones for both target and pointer. The target and pointer were presented in alternation, with the listener adjusting the ITD in real time. Trading ratios ranged from 0.3 µs/dB to 20 µs/dB (discussed in depth below). In yet another variation, Young Jr and Levine (1977) asked listeners to adjust a pure tone pointer to match the location of a noise target. They reported trading relations ranging from 40 – 80 µs/dB at 500 Hz. Hafter and Jeffress (1968) tested pure tone and noise stimuli in the same experiment, both with and without a standard diotic reference. They report a range of trading relations, ranging from 20 – 50 µs/dB for tonal stimuli, and 85 to 150 µs/dB for high-pass clicks.

1.2.3 Method of limits

Young (1976) used the method of limits to obtain a trading relation by asking listeners to report the position of an intracranial image using the terms “left,” “right,” or “midline” as the experimenters adjusted the ILD in the presence of a fixed ITD. Listeners made these reports as the auditory image moved from a random starting position and crossed the midline. The intensity needed to center the intracranial image when moving it back across the midline was considered the ILD value required to offset the ITD. At 400 Hz, the trading relation was approximately 80 µs/dB.

1.2.4 Detection

Hafter and Carrier (1972) measured psychometric functions using a same-different method. Each trial consisted of two, 500 Hz tone bursts in a 2-interval, forced choice (2IFC) task (see also Domnitz & Colburn, 1977). For the “same” condition, both signals were diotic, while the “different” condition contained a diotic standard followed by a dichotic test signal. Listeners responded “different” if there were any perceived differences between standard and test, otherwise they responded “same.” Measures of d' were plotted along the ordinate as a function of fixed ILD values tested along the abscissa. The parameter tested was the ITD value. The mean trading relation was approximately 19 μs/dB.

1.3 Response indication techniques

While there are a variety of techniques used to collect participants responses of perceived azimuth, this document divides them into two main types: allocentric and egocentric. Both types are discussed below.

1.3.1 Allocentric responding

This document refers to “allocentric” response techniques as those which require the observer to indicate the perceived location of a target using an external reference. Examples of this technique include asking participants to indicate an apparent source by positioning a dot on a diagram of a head wearing earphones as seen from behind (Lang & Buchner, 2008), or pointing to a location on a sphere positioned in front of the participant (Gilkey et al., 1995).

1.3.2 Egocentric responding

This document refers to “egocentric” response strategies as those which do not require a shift in first-person perspective. Examples of reporting that maintain a participant-centered reference include shining a spotlight on a semicircular strip placed in front of the participant (Butler & Naunton, 1962) and verbally calling out response coordinates (Wightman & Kistler, 1992). A particularly intuitive example was implemented by Stecker (2010). In a centering task, participants were asked to adjust the ILD present in a stimulus with a fixed ITD by rotating their heads. The ILD was calculated to be equal in magnitude and opposite in sign to the head azimuth. Turning the head toward the ITD caused the image to move toward the interaural midline, and turning in the opposite direction moved the image away from center (i.e., when the ILD and ITD favored the same direction). In a separate task, Stecker (2010) also asked participants to indicate the perceived azimuth of single stimulus presentations (using the method of constant stimuli; MOCS) by head-turn. The current study made use of the second type of task.

1.3.3 Head-pointing, lateralization and virtual reality

Gilkey et al. (1995) introduced the “God’s eye localization pointing” (GELP) method, and compared it to several other response techniques used for recording perceived azimuth. The GELP method, mentioned briefly above, makes use of a sphere positioned in front of the listener, who uses a stylus to point to the corresponding location of an acoustic signal. This technique was compared to localization data from studies that recorded perceived azimuth by asking listeners to call out coordinates (Wightman & Kistler, 1989) and point their heads in the direction of the perceived source (Makous & Middlebrooks, 1990). They found the head-pointing technique produced results that most closely matched the actual sound-field locations of the stimuli. It seems reasonable that an intuitive action, such as orienting toward a sound source, yielded more accurate localization judgments than the GELP method, which requires the listeners to alter the frame of reference to an externalized object. Head-pointing also proved more accurate than verbally calling out estimated coordinates, despite the egocentric nature of both tasks. It appears that the instinctiveness of orienting the head in the direction of a sound might offer an advantage over other egocentric techniques.

Considering that head-pointing yields the most accurate localization results, a pertinent question is whether this technique can be used to indicate the perception of an intracranial image presented under headphones. Results from Stecker (2010) suggest head-pointing is a valid method even without the use of sound-field stimuli. Localization data from his study were systematic and sensitive to the study parameters, and participants reported turning their heads in the direction of a perceived intracranial image was quick and intuitive. The ease of translating a “lateralization” task outside the head is in line with a report from Jeffress and Taylor (1961), who compared an externalized lateralization task (assigning an azimuthal position in space to stimuli presented under headphones) to similar data obtained in the sound-field (Stevens & Newman, 1936). They showed that judgments were very similar between headphone and sound-field stimuli, without the need for additional practice to externalize the headphone stimuli. Participants reported the task was easy, despite the fact they perceived the sound intracranially and indicated position using lamps positioned approximately 6 ft away. In a binaural interaction study, Lang and Buchner (2008) trained listeners using head-related transfer functions (to simulate the sound-field and achieve percepts outside the head), but used unfiltered stimuli during the testing session. They also reported systematic and sensitive results, without reported difficulty from participants.

In light of the results described above, the current study used a head-pointing technique to record participant responses. In an effort to increase the intuitive nature of the technique and create improved realism, the head-pointing procedure was performed in a virtual reality (VR) environment. Van Veen et al. (1998) advocate that virtual reality offers several benefits to laboratory tests. For instance, they mention the precise control of stimuli, easy manipulation of parameters, interactivity between subject and environment, improved multisensory realism, and multiple methods of recording responses. The current study utilized VR to simulate an outdoor, free-field environment. This step adds realism to previous head-pointing procedures, which required listeners to orient to sound sources in the presence of a variety of potential visual anchors present in the laboratory (e.g., speakers, wall and floor patterns). VR also offers the potential for consistent visual input when testing across studies and physical laboratory locations. It is important to note that this dissertation does not concern the influence of visual cues or VR on TRs. Rather, these experiments were a first step toward using VR in future studies for increased face validity and more complex manipulation of audiovisual interaction. The VR environment is described in detail in the General Methods.

1.4 Complications in quantifying trading relations

Consistent with the variety of factors at play in determining the frequency selectivity of binaural cues discussed above, multiple parameters affect ITD/ILD equivalence relations: the cue being adjusted (Young Jr & Levine, 1977); task (Lang & Buchner, 2008); adaptation (Thurlow & Jack, 1973); cue magnitude (David Jr et al., 1959); feedback (Carlile et al., 2001); the distance of the cues from the listener (Shinn-Cunningham et al., 2000); interclick interval (Stecker, 2010); relative laterality between cues (Moushegian & Jeffress, 1959); naturalness (Gaik, 1993); masking (Teas, 1962); and whether a reference tone is present (Ignaz et al., 2014).

1.4.1 Incomplete trading

The complex nature of binaural cue interaction is complicated by the finding that the trade between time and intensity is incomplete (e.g., Hafter & Carrier, 1972). In other words, there is no value of one cue that perceptually offsets the other cue completely. Hafter and Carrier (1972) demonstrated this by measuring psychometric functions for listeners’ ability to detect a difference between a diotic and dichotic signal in a 2IFC task (described earlier). Following a standard, diotic signal, a fixed ITD was presented (0, 10, 20, 30 and 40 μs; the parameter) over a range of ILDs (abscissa). Listeners' sensitivity was calculated as *d'*, which was plotted along the ordinate. The range of ILDs tested included values that were both higher and lower than the ILD that yielded poorest detection, which produced "V" shaped functions with minima indicating the least sensitive combination of ITD and ILD values. Notably, all of the function minima revealed sensitivity above *d'* = 0, implying incompleteness of the trade between time and intensity. From these results, Hafter and Carrier determined (1) the trade between time and intensity is incomplete; and (2) a partial trade does exist (function shapes depended on the ITD and ILD values). A third observation was that the weight of each cue differed widely across participants, despite well-practiced listeners (no less than 32,000 observations).

1.4.2 Perception of dual images

A related complicating factor when measuring trading relations is the perception of two auditory images reported by some studies examining binaural interaction under headphones (e.g., Banister, 1926; Hafter & Jeffress, 1968; Whitworth & Jeffress, 1961). Whitworth and Jeffress (1961) investigated a phenomenon described by Banister (1926), wherein opposing interaural cues led to the perception of two separate auditory images. As described above, Whitworth and Jeffress (1961) asked listeners to adjust the ITD of a 500 Hz pointer until it coincided with the perceived azimuth of a fixed 500 Hz target. The target ILD was always 0 dB, with the ITD fixed at one of seven values (0, ±90, ±180, and ±270 μs). The pointer ILD was also fixed at one of seven values (0, ±3, ±6, and ±9 dB) while listeners adjusted the ITD. The results were plotted as the ILD of the signal along the abscissa, the ITD of the signal as the parameter, and the ITD adjustment, in μs, along the ordinate (see Figure 1). The data revealed listeners were able to use the acoustic pointer to indicate the perceived azimuth of two intracranial images: one image that was determined almost entirely by the ITD (the “time” image; lower plot of Figure 1), and one image that was determined by a combination of time and intensity (termed the “intensity” image; upper plot of Figure 1). Whitworth and Jeffress reported TRs with values of 0.3 µs/dB and 20 µs/dB for the time and intensity images, respectively.

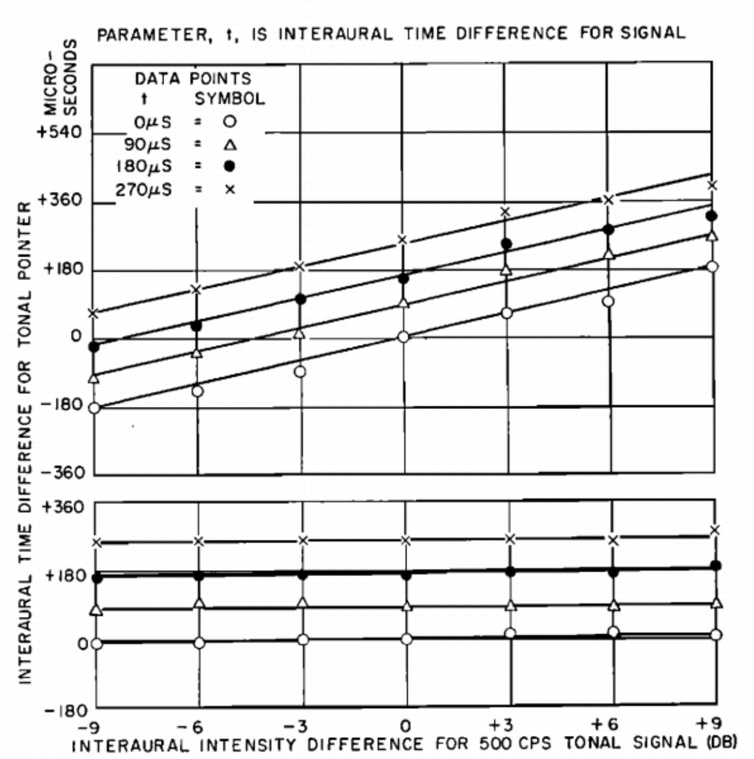


FIG. 1. “Intensity” results (upper) and “time” results (lower) for RHW.

Figure 1. Perception of dual auditory images taken from Whitworth and Jeffress (1961). The upper part of the plot shows responses for the time/intensity image when binaural cues were set into opposition. The slope depicts the change in perceived azimuth depending on the fixed ILD value (abscissa). The lower portion of the plot shows responses when listeners focused on the time image. The flat slope indicates the time image was not affected by the fixed ILD. From “Time vs Intensity in the Localization of Tones,” by R. H. Whitworth and L. A. Jeffress, 1961, *The Journal of the Acoustical Society of America, 33,* pg. 1441–1445. Copyright 1961 by AIP Publishing. Reprinted with permission.

While the absolute values of the time and intensity relations reported by Whitworth and Jeffress (1961) do not always agree precisely with other reported values, a consistent finding across studies of TRs shows the time image produces a comparatively smaller ratio than the intensity image. Hafter and Jeffress (1968) suggested that the wide range of trading relations across studies may be due to the presence of dual images. They argue that because it takes extensive training for a listener to perceive and interact with double images, participants may unwittingly respond to one image on some trials and the complimentary image on others, within the same experimental session. For example, a listener who responded to the intensity image on one trial would require a larger offsetting ITD than when the same listener responded to the time image.

The existence of a poorly understood, confounding factor that leads to variations in trading relations is a fundamental deficit in obtaining reliable information pertaining to binaural cue interaction. As mentioned above, two main hypotheses have been proposed to account for the different trading relations that seem to depend on either time or a combination of time and intensity cues (discussed below).

1.4.3 Existing hypotheses for cue-dependent trading

Regression

In contrast to the theory proposed by Hafter and Jeffress (1968), that variations in trading relations can be explained by the perception of dual images, Trahiotis and Kappauf (1978) proposed a judgmental bias can account for the cue-specific trading data. They cite similar differential results in the psychophysical literature at large when using the MOA to measure a common function obtained by matching variables of different dimensions. They discuss vibrotactile data from Sheldon (1973), as discussed by Kappauf (1975). In brief, participants matched the abruptness of a vibrotactile standard by adjusting one of two parameters of a similar vibrotactile target: rise time, and final amplitude. A different “equal-surge contour” was produced depending on whether the participants adjusted rise time or amplitude. Specifically, when time was adjusted, the final value was closer to the rise time of the standard. Conversely, when amplitude was adjusted, the final value was closer to the amplitude value of the standard.

In other words, “the observer’s matching settings regress toward the level of the standard on the dimension being adjusted” (Kappauf, 1975). Thus, adjusting the ILD of a pointer to match a diotic standard would result in a smaller ILD denominator because the adjusted ILD value would regress toward 0 dB, resulting in an artificially large trading value. For example, a TR of 33 µs/dB (ITD = 400 µs, ILD = 12 dB) increases with an ILD biased closer to 0 dB: 80 µs/dB (ITD = 400 µs, ILD = 4 dB). The opposite effect on the ratio occurs when the ITD is adjusted. The original 33 µs/dB would shrink to just 8 µs/dB (ITD = 100 µs, ILD = 12 dB). These trading relations are biased in the same direction as the reported time- and intensity-based equivalence relations reported in existing studies.

Attentional upweighting

Lang and Buchner (2008, 2009) propose a different account for the difference in TRs depending on the cue being adjusted. In a first experiment, TRs were measured using the MOA, where listeners used a slider presented on a computer screen to adjust the ILD (or ITD) of a target with a fixed, opposing value of the complimentary cue. The listeners were instructed to center the auditory image. The stimulus was played in a loop (ISI = 500 ms) until adjustments were completed, and the final values of ITD and ILD that produced a centered percept were recorded. In a second experiment, participants judged the location of stimuli presented a single time that contained the same values of ITD and ILD required to center the target from the previous experiment. Listeners indicated perceived azimuth by positioning a red dot in relation to a representative drawing of a head. The results revealed that previously centered percepts obtained using the MOA no longer appeared at midline when presented in isolation. Instead, Lang and Buchner reported a “shift-back” effect, wherein the previously centered stimulus was perceived closer to the location of the static cue presented during the adjustment experiment (Figure 2).

Lang and Buchner (2008, 2009) argue that increased attention to the cue being adjusted during the MOA task results in a perceptual upweighting of the adjusted cue. For example, the artificially inflated weight of the ILD cue during adjustment would lead to an ILD insufficient to offset the opposing ITD when both cues were presented as a single stimulus in a localization task (i.e., when neither cue benefited from increased attention). The insufficient ILD creates an imbalance favoring the ITD, resulting in a percept shifted more toward the location indicated by the now-dominant ITD. The implication for measuring TRs is that adjusting the ILD leads to smaller required level differences, and thus larger trading ratios. Conversely, adjusting the ITD leads to smaller required time differences and smaller trading ratios. This pattern of cue-specific trading relations is consistent with the regression hypothesis, as well as the values reported in the binaural interaction literature.

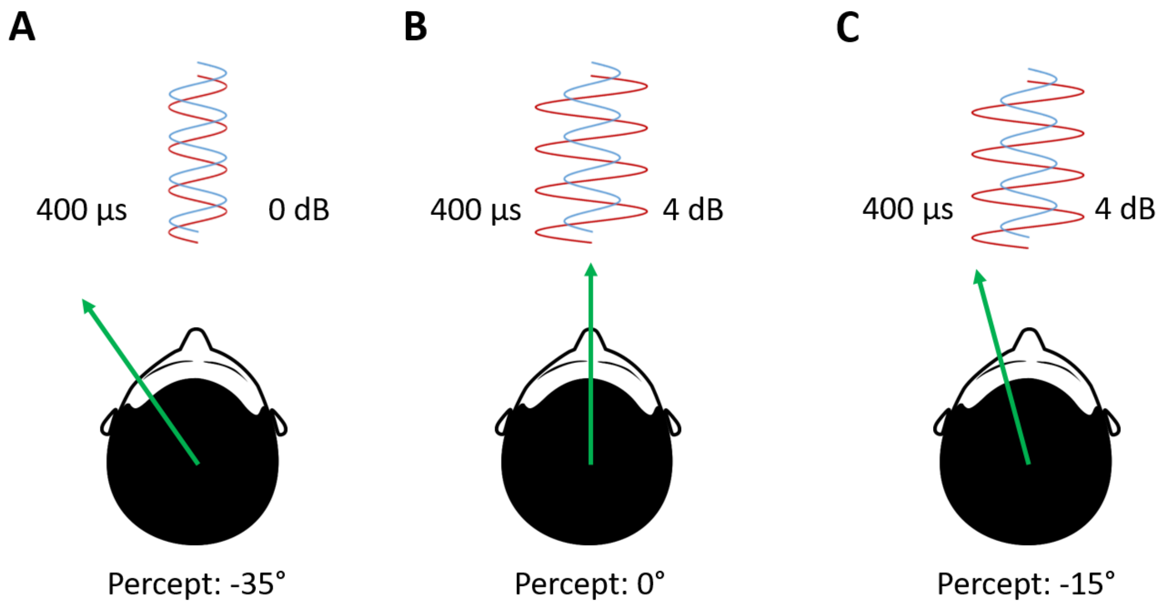


Figure 2. Graphical representation of the shift-back effect described by Lang and Buchner (2008). A. The auditory image is perceived 35° to the left due to a 400 µs ITD favoring the left ear. B. Method of adjustment experiment, requiring the participant to offset the fixed ITD by adjusting the ILD. The green arrow shows the perceived location of the intracranial image at midline after the participant introduced an opposing 4 dB ILD. The effectiveness of the ILD is increased due to attending to that cue during adjustment. C. Localization experiment, presenting the same values obtained in A, in isolation (i.e., without the attentional benefit). The green arrow shows the intracranial image is no longer sufficient to center the auditory image, and the percept has “shifted back” toward the fixed ITD.

Regression and attentional upweighting

In an effort to determine whether regression or upweighting acted alone or in concert to influence trading relations, Ignaz et al. (2014) measured equivalence relations both with and without the presence of a reference tone in the same participants. They found that while cue-specific trading relations occurred in the absence of a reference tone, confirming the experiments of Lang and Buchner (2008, 2009), the shift-back effect was greater when a reference was presented in alternation with the target, in support of the regression hypothesis. Taken together, the data reveal the existence of a perceptual phenomenon that differentially affects trading relations depending on the cue being adjusted in an MOA task. The mechanism may involve topdown control from attentional processes, but is also modulated by stimulus parameters (i.e., the presence or absence of an acoustic reference).

Adaptation

The current study suggests a third possibility to account for the cue-dependent nature of TRs: auditory spatial adaptation. The section below introduces the concept of spatial adaptation in the auditory system and illustrates how adaptive processes can account for the existing binaural cue trading relationship findings.

1.5 Adaptation in the spatial auditory system

1.5.1 Adaptive localization aftereffects

Flügel (1920) first investigated the effect of prolonged, monaural exposure to sound on the azimuthal localization ability of the human auditory system. He showed that while binaural presentation of a tone resulted in a centered percept in the head, following monaural exposure to an adapting tone (from 0.25 – 12 minutes), the same binaural presentation resulted in a perceived shift in the auditory image away from the adapted ear. Because the auditory image shifted in apparent location away from the adapted ear, Flügel reasoned the adaptor induced fatigue in the exposed ear, creating a preponderance of perceptual sensitivity favoring the unadapted ear. However, Bartlett and Mark (1922) found similar results using a binaural adaptor, suggesting the mechanism is more nuanced than simple neuronal fatigue (see also Jones & Bunting, 1949). Thurlow and Jack (1973) systematically tested the lateral placement of adaptors and probes for both ITD and ILD cues. Consistent with the early literature, they always noted a shift in the probe away from the adaptor. Specifically, eccentric adaptors of either cue type caused eccentric probes of the same cue type to shift toward the midline, while midline adaptors caused probes to shift away from the midline (Figure 3). The effects of using adaptor/probe pairs of mixed cue types revealed results that were inconclusive.

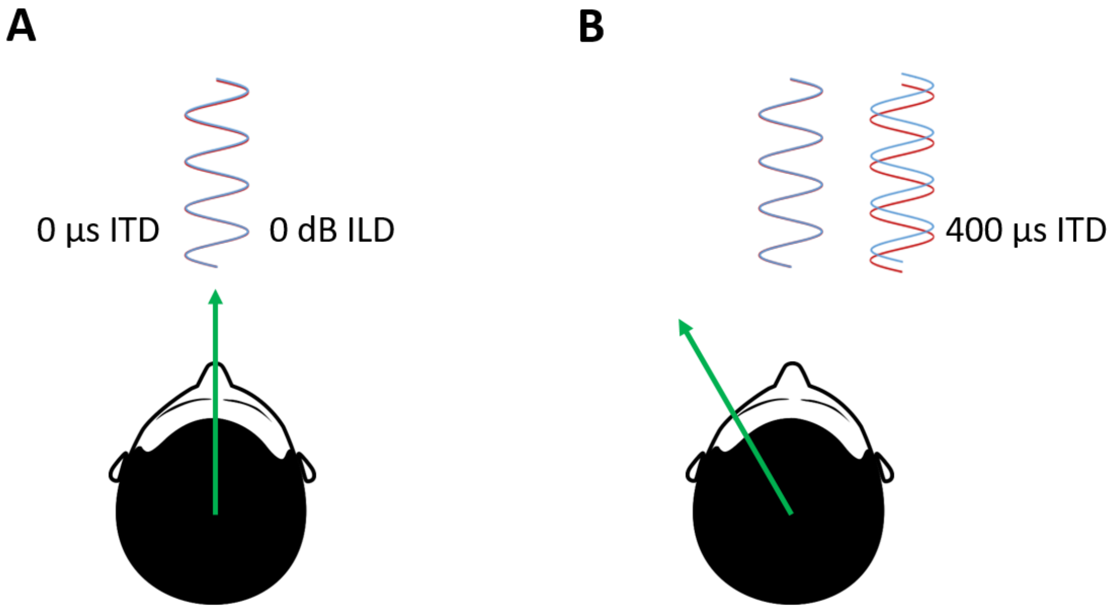


Figure 3. An illustration of the auditory localization aftereffect from Thurlow and Jack (1973). A. The green arrow illustrates midline perception of the intracranial image resulting from a diotic stimulus. B. The green arrow shows that following a preceding adaptor (the signal carrying a 400 µs ITD), the same diotic stimulus is perceived displaced away from the adaptor (i.e., away from the midline).

Canévet and Meunier (1994) measured the shift of a 15-degree probe following a midline adaptor, and found increasing drift of the probe with increasing duration of the adaptor. Canévet and Meunier (1996) repeated their findings in the sound-field as well as under headphones. Meunier et al. (1996) tested the adaptive aftereffect over a range of stimulus frequencies and bandwidths. They found the shift of the probe was larger using a narrow-band-noise (NBN) centered at 4 kHz compared with an NBN centered at 1 kHz. They also found that the effect was greatest when a broadband adaptor (2 – 6 kHz) overlapped in frequency with the probe (i.e., greater effect for 4 kHz NBN than 1 kHz NBN). The direction of the shift was always away from the position of the adaptor.

Kashino and Nishida (1998) systematically examined the frequency and ITD selectivity of the localization aftereffect. They found the effect was greatest when the frequencies of the adaptor and probe were similar, with the perceived shift disappearing for frequency separations greater than one-half octave. The results also suggested the aftereffect is selective to ITD value. The magnitude of the shift in the probe's apparent location was greatest when ITDs differed by 250 µs, and decreased for ITDs greater and lesser than this value for a tone at 4000 Hz. Consistent with previous work, the perceptual shift in probe location was always away from the position of the adaptor. Braasch and Hartung (2002) also confirmed the seemingly repulsive effect of an adaptor on a probe, and showed that the effect was greater in reverberation, compared to anechoic conditions.

1.5.2 Similarities across psychophysical tasks

An interesting observation is the similarity in methodologies between studies of trading relations using the MOA and studies of auditory spatial adaptation. Many of the trading relation studies using the MOA described earlier use a paradigm that presents a standard and target in alternation. The result is repeated exposure to a static cue, followed by a changing (adjusted) complimentary cue. Kopčo et al. (2007) showed that displacement of a probe in the presence of a preceding stimulus can occur with a single adaptor presentation, with an adaptor duration of only 2 ms. One interpretation of these results is that a lengthy period of adaptation is not a prerequisite to elicit the localization aftereffect.

Another similarity across MOA and adaptation tasks is the interstimulus interval (ISI). The ISIs used in MOA studies of binaural interaction typically range from roughly 200 to 500 ms (e.g., Hafter & Carrier, 1972; Lang & Buchner, 2008). This range overlaps with ISIs that produce the localization aftereffect (e.g., Kashino & Nishida, 1998; Kopčo et al., 2007; Phillips et al., 2006). Furthermore, the standard and target in MOA tasks can be of the same frequency (e.g., Lang & Buchner, 2009; Whitworth & Jeffress, 1961), which has been shown to produce spatial adaptive aftereffects of the greatest magnitude (e.g., Kashino & Nishida, 1998).

It seems reasonable that cue-specific trading relations reported by studies of binaural interaction using the MOA could be contaminated by introducing the auditory localization aftereffect, due to the similar methodologies across studies.

1.5.3 Adaptation and binaural interaction

While it seems reasonable that similar methods could lead to similar perceptual effects, an important consideration is whether one binaural cue can adapt the complimentary cue at all. Phillips et al. (2006) investigated the relationship between ITD and ILD by measuring psychometric functions for each cue type alone, and then following an adaptor of the complimentary cue type (e.g., an ITD adaptor followed by an ILD probe). The psychometric functions were consistently displaced from the adaptor, indicating that complimentary cue types can indeed serve as adaptors. Consistent with the same-cue adaptation literature, the probe was always displaced away from the adaptor, suggesting a common mechanism.

1.5.4 Adaptation and cue-specific trades

It seems possible that auditory spatial adaptation could account for the observed findings in trading relations obtained using the MOA. As in the discussion concerning regression and the attention-shift model, a 33 μs/dB ratio (ITD = 400 μs, ILD = 12 dB) decreases when adjusting the ITD in an MOA centering task. The repeated presentation of a fixed standard ILD (ITD = 0 µs) favoring the left could serve as an adaptor to the adjusted ITD target. The result is that a smaller ITD is required to offset the ILD to center the percept, because as the target approaches the ILD "adaptor" it is displaced toward midline. Thus the full ITD required to offset the ILD and center the image is not necessary using the MOA. This effect could conceivably occur with or without a reference tone. In the absence of a reference, the static ILD could serve as an adaptor and the changing ITD cue as the probe. The physical presence of a midline reference tone would lead to an even greater displacement of the percept from midline, because the repeated reference tone would become the adaptor. This scenario could account for the same pattern of results demonstrated by Ignaz et al., (2014) (i.e., greater shift-back effect in the presence of a reference tone).

Another possible role for adaptation is to increase neural thresholds for the static cue over time, creating an artificial imbalance favoring the adjusted cue. This is, in essence, a scenario opposite that proposed by Lang and Buchner (2008), who argued that attention to the adjusted cue led to greater weighting. There is in fact precedence for the weakening of a cue leading to changes in TR, rather than an increase in weighting. Stecker (2010) showed that decreasing the interclick interval between Gaussian-filtered impulses below 5 ms abolished the envelope cues necessary to extract ITD. This led to a shift in the equivalence function that favored the ILD. Subsequent analysis confirmed the shift was due to weakened ITD cues, rather than an increase in ILD effectiveness. Consistent with the findings of Stecker (2010), adaptation of the repeated cue would create a preponderance of activation favoring the adjusted cue due to a reduction in neural response to the static cue – not an increase in firing to the adjusted cue.

1.6 Purpose of the current study

It has been shown that despite great advances in our understanding of the relationship between the azimuthal cues for sound source localization, current knowledge cannot explain the cue-specific nature of their interaction. The motivation behind this study was to provide novel insight into the fundamental nature of binaural spatial cues in order to advance current understanding of basic auditory spatial perception. To that end, this study investigated the potential influence of the auditory localization aftereffect on binaural cue TRs, using a head-pointing technique in a virtual reality environment.

Three experiments were carried out. Experiment 1 measured TRs obtained using the MOA. Listeners adjusted the amount of ITD required to center a stimulus containing one of several fixed ILDs, and vice versa. Experiment 2 measured TRs obtained using a head-pointing technique similar to Stecker (2010). Combinations of ITD and ILD were presented in isolation, and the oriented head angle indicated perceived azimuth. Experiment 3 was identical to Experiment 2, with the addition of an adapting train preceding each probe.

It was hypothesized that the results from Experiment 1 and Experiment 3 (the MOA task and the adaptation paradigm) would produce similar TRs. That is, adaptation present in the MOCS adaptor conditions would reproduce the cue-dependent effects obtained from the MOA task. Accordingly, the results obtained from Experiment 2 (the no-adaptor head-pointing task) should differ from Experiments 1 and 3, because the no-adaptor MOCS task does not allow for adaptation. Specifically, the TR from the no-adaptor MOCS task should lie between those obtained from the other experiments. If these hypotheses are validated, similar TRs between the MOA and adaptation paradigm will provide evidence suggesting auditory spatial adaptation is involved in trading ITDs and ILDs. Such a finding would have implications for the interpretation of past work and for the design of future studies investigating binaural cue interaction.

Chapter 2

GENERAL METHODS

2.1 Participants

Ten adult listeners were recruited from Vanderbilt University for this study. One participant was excluded due to inability to complete the task. The remaining nine participants (8 females; aged 24 – 33 years; M = 28 years) completed all tasks. All participants had normal, symmetrical hearing at octave frequencies from 250 – 8000 Hz (< 25 dB HL), verified using standard audiometric procedures for air conduction thresholds. There was no history of neurogenic or otologic disease, as evidenced by self-report. All participants reported normal, or corrected normal visual acuity and color vision. Participants were compensated for their time. This study was approved by the Vanderbilt Institutional Review Board.

2.2 Testing environment and apparatus

All sessions were conducted in a sound-treated room. Participants wore an Oculus Rift virtual reality headset (https://oculus.com), while seated in a swivel chair approximately 1 m from dual motion sensors. The custom virtual environment was coded using the Unity3D game engine (https://unity3d.com; version 2018.2.1f1) on a custom-built PC running Steam VR (version 2017-01-30, Valve Corporation, Bellevue WA USA). The virtual environment placed the participant in the center of a circular platform, with red helium balloons “tied” around the outer platform perimeter in 1-degree steps. The only orienting cue was that the balloon at midline was green. The larger area was an outdoor setting consisting of uniform grass and clear sky to avoid visual reference points, while also creating the visual equivalent of a free field (Figure 4).

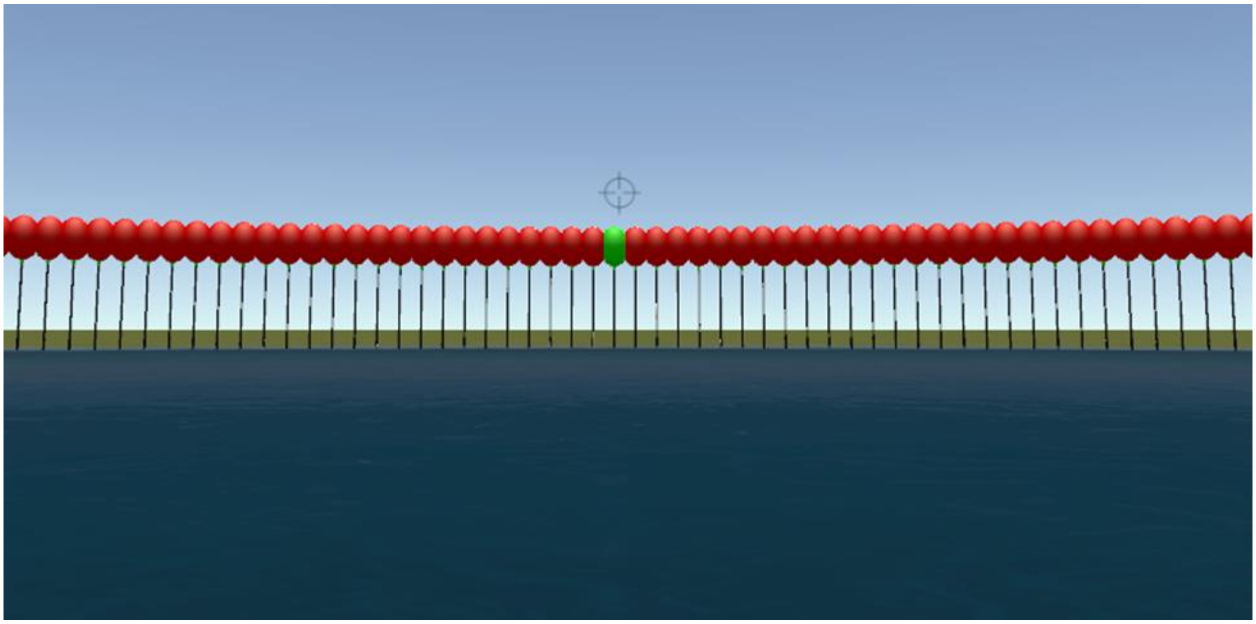


Figure 4. The virtual reality environment seen during the localization experiments. The green balloon visually marks the midline. The reticle (above green balloon) moves with the head and is used to indicate perceived azimuth.

Participants interacted with the environment to make responses via standard Oculus handheld controllers. Each controller had two push buttons, a thumbstick button, a trigger, and a grip button. Various input methods were used for each experiment (discussed in the experiment-specific methods). The spatial position of the head-mounted device (HMD) was tracked using the Rift’s onboard gyroscope.

A reticle in the center of the visual field followed participant head movements, allowing them to aim at individual balloons simply by orienting the head. This paradigm was also used to maintain proper head position at the onset of each trial. Participants were instructed to keep the reticle centered on the green balloon (midline) either throughout the experiment (Experiment 1), or to begin a new trial after head pointing (Experiments 2 and 3). If the reticle moved away from the green balloon, the experiment stopped and a green box appeared at midline. The experiment continued only after the reticle was returned to the green box for 2 seconds.

A second PC (Dell, Inc.) running MATLAB (Mathworks, Natick, MA) communicated with the presentation computer running the Unity3D game engine via transmission control protocol/Internet protocol (TCP/IP). Behavioral tasks for all experiments were coded in MATLAB. These scripts also controlled the virtual reality environment via triggers to call custom Unity3D functions (e.g., balloon pop, reset environment), and to store responses and HMD position data. A diagram of the setup is provided in Figure 5.

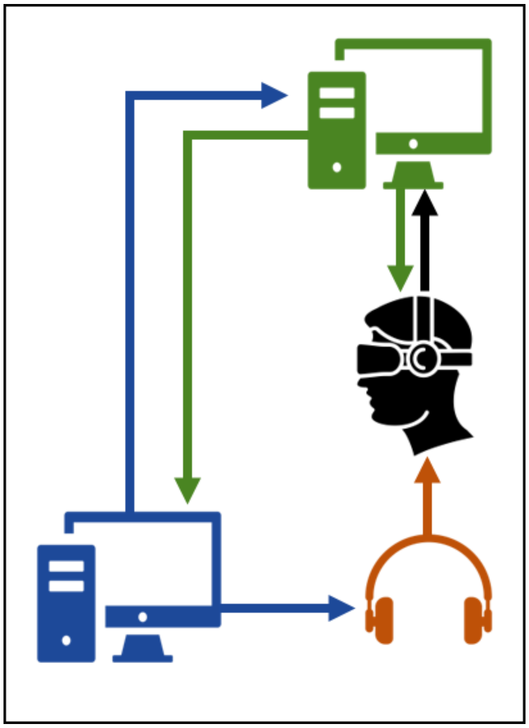


Figure 5. The experimental setup. The blue (bottom) computer runs MATLAB, which controls the virtual scene, rendered by the green (top) computer. The blue computer also delivers audio directly to the insert earphones. The green computer records and sends responses and head position to the blue computer for storage.

2.3 Stimuli

All sounds were synthesized using MATLAB. Because synchronization between auditory and visual stimuli were not of importance to the study, no timing calibration measurements were made between the audio onset and Unity3D function execution. Auditory stimuli were synthesized at 48.828 kHz (Tucker-Davis Technologies RP2.1, Alachua, FL) and presented via ER-2 insert earphones (Etymotic, Elk Grove Village, IL). Stimuli were presented from the MATLAB PC, bypassing the Unity3D audio device completely. All stimuli consisted of 500 Hz pure tones with a duration of 500 ms. Unless modified by introducing an ILD, all stimuli were presented at a level of 65 dBA. Tones were gated using raised cosine ramps of 20 ms duration to avoid spectral transients. Differences in arrival time at the two ears were computed by shifting the whole waveform of one channel relative to the other in time. Level differences were achieved by halving the desired ILD and applying offsets as a reduction to one channel, and as an increase to the other channel.

Pure tones at 500 Hz were chosen for several reasons. First, 500 Hz provides a robust, lateralizing cue for manipulation of ITDs (Zwislocki & Feldman, 1956), while also being sensitive to ILDs under headphones. Second, 500 Hz tones allow the results of the current study to be compared with existing studies of binaural cue interaction, which have commonly used 500 Hz pure tone stimuli (e.g., Harris, 1960; Lang & Buchner, 2009; Whitworth & Jeffress, 1961). Third, the use of 500 Hz pure tones allowed for precise control over the frequency content of the stimuli. Due to the exploratory nature of this study, we felt it reasonable to accept the drawbacks of using pure tones, which include being contrived, laboratory stimuli, and the relative ineffectiveness of ILD cues at low frequencies in sound-field listening (Mills, 1960).

2.4 Procedure

A total of three experiments were conducted. Experiment 1 sought to replicate the MOA literature, by using a centering task to obtain TRs when adjusting the ITD and ILD, respectively. Experiment 1 consisted of two visits, each consisting of approximately two hours. Experiment 2 obtained TRs using a single-presentation localization task, in order to remove the possibility of adaptive effects. This experiment consisted of one visit of approximately 2 hours. Experiment 3 was the same as Experiment 2, but introduced adapting trains before the single presentations of the probe stimulus. This experiment resulted in TRs affected by either ITD or ILD adaptors that matched the corresponding cue in the probe. Experiment 3 consisted of two visits of approximately 2 hours each. The total time required to for each participant was then approximately ten hours over five visits.

Where possible, each experiment presented stimuli using one of two presentation patterns: cue types were either intermixed within a single experimental session (e.g., trials contained ITD and ILD adaptors within the same block), or only a single cue type was presented in any one session (e.g., session 1 contained ITD adaptors only, session 2 contained ILD adaptors only). This was done to examine whether repeated exposure to a single cue type over time rendered listeners more sensitive to adaptive effects. To this end, one group of participants was presented intermixed cues (the Mixed group; S1, S2, S3, S5 and S6) and one group of participants was presented a single cue type during any one session (the Fixed group; S4, S7, S8, S9). Conditions were counterbalanced for Experiment for both groups, and further counterbalanced by cue type for the Fixed group.

Chapters 3 through 5 present each experiment as a self-contained data set. Comparisons across experiments are first made in Chapter 6. Chapter 7 serves as a general discussion of the study as a whole, and Chapter 8 is a concise summary of the major conclusions.

Chapter 3

EXPERIMENT 1: MOA (CENTERING TASK)

3.1 Experimental methods

3.1.1 Stimuli

The stimuli in Experiment 1 were synthesized using the parameters described in the General Methods. Trials consisted of looped, alternating presentations of standard and target tones. The standard tone always carried a 0 dB ILD and a 0 µs ITD. The target tone consisted of either a fixed ILD (0, ±3, ±6, or ±9 dB), or a fixed ITD (0, ±100, ±200, or ±300 µs) and a variable, complimentary cue used to center the test tone to midline. This yielded 14 different conditions. The cue values were chosen after Whitworth and Jeffress (1961), who had successfully demonstrated incomplete trading using these values with 500 Hz tones. The variable complimentary cue was adjusted by the participant (described below in the Procedure), and started at a random value ranging from ±3 to ±9 dB, or ±100 to ±300 µs for ILD and ITD cues, respectively. The standard and target tones were separated by a 400 ms interstimulus interval. Each standard-target pair was separated by a silent interval of 600 ms. The increased duration of the intertrial interval was introduced to render standard-target pairs more easily recognizable, due to reported difficulty segregating the pairs during pilot testing (see also Domnitz & Colburn, 1977).

3.1.2 Procedure

Participants completed a centering task with insert earphones. Stimuli were presented using the method of adjustment (MOA). Participants initiated each trial by pulling the trigger on the right Oculus Rift controller. A brief animation (three balloons bobbing) indicated the trigger pull had been read and the trial had begun. In this experiment, the virtual environment served to ensure participants kept their heads centered during the task (described in the General Methods) and to provide visual consistency across experiments, but otherwise there was no interaction with the VR surroundings for the MOA task. The variable cue of the target tone was adjusted by the participant using the handheld Oculus Rift controllers until the target tone was perceived as coming from the midline (i.e., overlapping the standard tone in perceived azimuth). The right controller increased the time or level advantage to the right ear (arrival time lead or higher level). The left controller increased the time or level advantage to the left ear. Adjustments were made by pressing one of the push buttons or the grip button of each controller. When the adjustable cue was the ITD, pressing the push button increased the time lead in steps of 10 µs, and pressing the grip button increased the time lead in steps of 100 µs (up to a ±900 µs maximum). When the adjustable cue was the ILD, pressing the push button increased the level difference in steps of 0.1 dB, and pressing the grip button increased the level difference in steps of 1 dB (up to a 15-dB maximum). After participants were satisfied that the target tone had been centered, they pushed the thumbstick on the right controller to end the trial and record the cue value. Another animation (color changes) signaled the thumbstick press had been read and the trial had ended.

Each session began with at least 8 practice trials. During this time, participants could ask questions and were given as much time as necessary to familiarize themselves with the controls. After 8 practice trials, additional practice was provided until a participant reported comfort with the task. Practice data were inspected to ensure performance was broadly consistent with expectations: e.g., a fixed, right-ear level advantage was perceptually centered by the participant introducing a left-ear time advantage. A total of 8 judgements were made during data collection for each of the 14 conditions (112 recorded responses).

Five participants were presented trials randomly from any of the seven possible ITD (0, ±100, ±200, or ±300 μs) and ILD (0, ±3, ±6, or ±9 dB) fixed cue values. That is, those 5 participants adjusted both cue types intermixed within the same session or day (the Mixed group). The remaining 4 participants were only presented trials of one cue type per session (the Fixed group).

3.1.3 Data and Analyses

The final cue value chosen to center the static, complementary cue was recorded at the end of each trial. The values of the 8 judgments per condition were averaged into a single data point, after removing outliers by determining their the absolute deviation from the median (Leys et al., 2013). A total of 28 outliers were removed across all participants and conditions (approximately 6.5% of data points). All data were plotted with ITD (μs) along the ordinate, and ILD (dB) along the abscissa. Therefore, ITD judgments are fixed along the abscissa according to the fixed ILD value against which the adjustment was made. ITD values indicate timing judgments along the ordinate. Conversely, the ILD judgments are fixed along the ordinate, according to the fixed ITD value against which the level judgments were made. Level judgements are indicated along the abscissa. Data from both conditions are shown within a single plot (see Figure 7).

The data points for ITD and ILD fixed cue values were fit using linear regression. The resulting slope was taken as the trading relation in that condition. In other words, each participant produced two TRs: one based on the slope of the data points when adjusting the ITD (henceforth ITD_{adj}), and one based on the slope of the data points when adjusting the ILD (henceforth ILD_{adj}). The group mean $\mathrm{ITD}_{\mathrm{adj}}$ and $\mathrm{ILD}_{\mathrm{adj}}$ TRs were compared using a bootstrapped paired-samples *t*-test. If cue trading required differing TRs based on the cue being adjusted to center the auditory percept (as hypothesized and consistent with the literature), the *t*-test will reveal the mean TRs are statistically different from each other.

The reliability of the data over time was measured using the split-half method. This approach groups the first 4 judgments and the second 4 judgments for each cue condition. The correlation between the early and later judgments is an indication of the extent to which first-half and second-half responses contributed to the mean response. If the scores are well correlated to each other, the data are considered reliable.

3.2 Results

Data from all nine participants contributed to the results. No participants were reliably able to offset ILD values of ±6 or ± 9 dB with any amount of ITD. Participants were only consistently able to offset ILD cues at values of 0 dB and ±3 dB. Consequently, the TRs for the MOA task are derived from the slope of three data points per condition instead of the intended seven. It should be noted that the last four listeners were not tested in the ±9 dB ILD condition at all. Potential explanations for the truncated range of testable ILD values are considered in the Discussion.

3.2.1 Descriptive statistics for the MOA task

For all but one listener ($r = 0.29$), the split-half reliability of the data revealed significant correlations between the first 4 and second 4 judgments (range: $r = 0.56$ to $r = 0.94$). Individual plots of the correlations are provided in Figure 6. Overall, the data indicate participant responses were stable throughout the task, excluding learning effects, fatigue or changes in response strategy accounting for the results. Similar statistical results were achieved whether the listener displaying low split-half reliability was retained or not, therefore that participant has been included in all subsequent analyses.

A further quantification of the data is provided in Figure 7, where error bars denote the standard error around the mean for each individual mean judgment, for each condition. In addition to the reliability over time revealed by the split-half test, the standard error bars show little deviation of individual judgments around the mean for both $\mathrm{ITD}_{\mathrm{adj}}$ and $\mathrm{ILD}_{\mathrm{adj}}$ TRs (SEM = 3.45 µs/dB and 5.79 µs/dB, respectively).

The results of the Shapiro-Wilk normality test revealed the MOA task data were normally distributed. However, Bartlett’s test for homogeneity of variance between ITD and ILD conditions failed to reject the null hypothesis (K-squared = 22.93, $p < 0.05$), indicating the variance across conditions was unequal. Appropriate statistical tests were chosen to account for the violation.

3.2.2 Mixed vs. Fixed groups

Unequal variance *t*-tests (Welch two-sample test) comparing the Mixed and Fixed group TRs were conducted to determine whether the manner in which the cues were presented (i.e., mixing cue types, or presenting only a single cue type per session) influenced the results. The *t*-

tests revealed no significant differences between the Mixed and Fixed groups for either the ITD_{adj} TR (M_{mixed} = 25.8 µs/dB; M_{fixed} = 30.5 µs/dB; $t(4.5) = 0.6$, $p > 0.5$) or the ILD_{adj} TR (M_{mixed} = 37 µs/dB; M_{fixed} = 41.5 µs/dB; $t(4.3) = 0.33$, $p > 0.5$), suggesting the MOA is not sensitive to intermixing cue types within a session. Because there were no statistical differences between groups, subsequent analyses of the MOA took place on the pooled data. It is interesting to note, that despite the lack of statistical difference between the Mixed and Fixed groups, there is a visual trend for more consistency in responses over time for the Fixed group (see Figure 6).

3.2.3 TR: Adjusting ITD (fixed ILD)

Individual TRs obtained when participants adjusted the value of the ITD in the presence of various fixed ILDs are shown in Figure 7 (blue points). The mean TR while adjusting the ITD was 27.91 µs/dB (range = 15 to 44.9 µs/dB, SEM = 3.45 µs/dB).

3.2.4 TR: Adjusting ILD (fixed ITD)

Individual TRs when participants adjusted the value of the ILD in the presence of various fixed ITDs are shown in Figure 7 (green points). The mean TR while adjusting the ILD was 39.01 µs/dB (range = 19.9 to 69.1 µs/dB, SEM = 5.79 µs/dB).

3.2.5 Comparison between ITD_{adj} and ILD_{adj} trading relations

A bootstrapped paired-samples *t*-test (10,000 replications) comparing TRs between conditions revealed a significant difference between the ITD_{adj} (M = 27.91 µs/dB) and ILD_{adj} (M = 39.01 µs/dB) for the MOA task ($t(8) = 3.87$, 95% CI(-1.88, 1.81), $p < 0.01$, $d = 1.29$). Individual (thin lines) and mean (thick lines) slopes are superimposed in Figure 8.

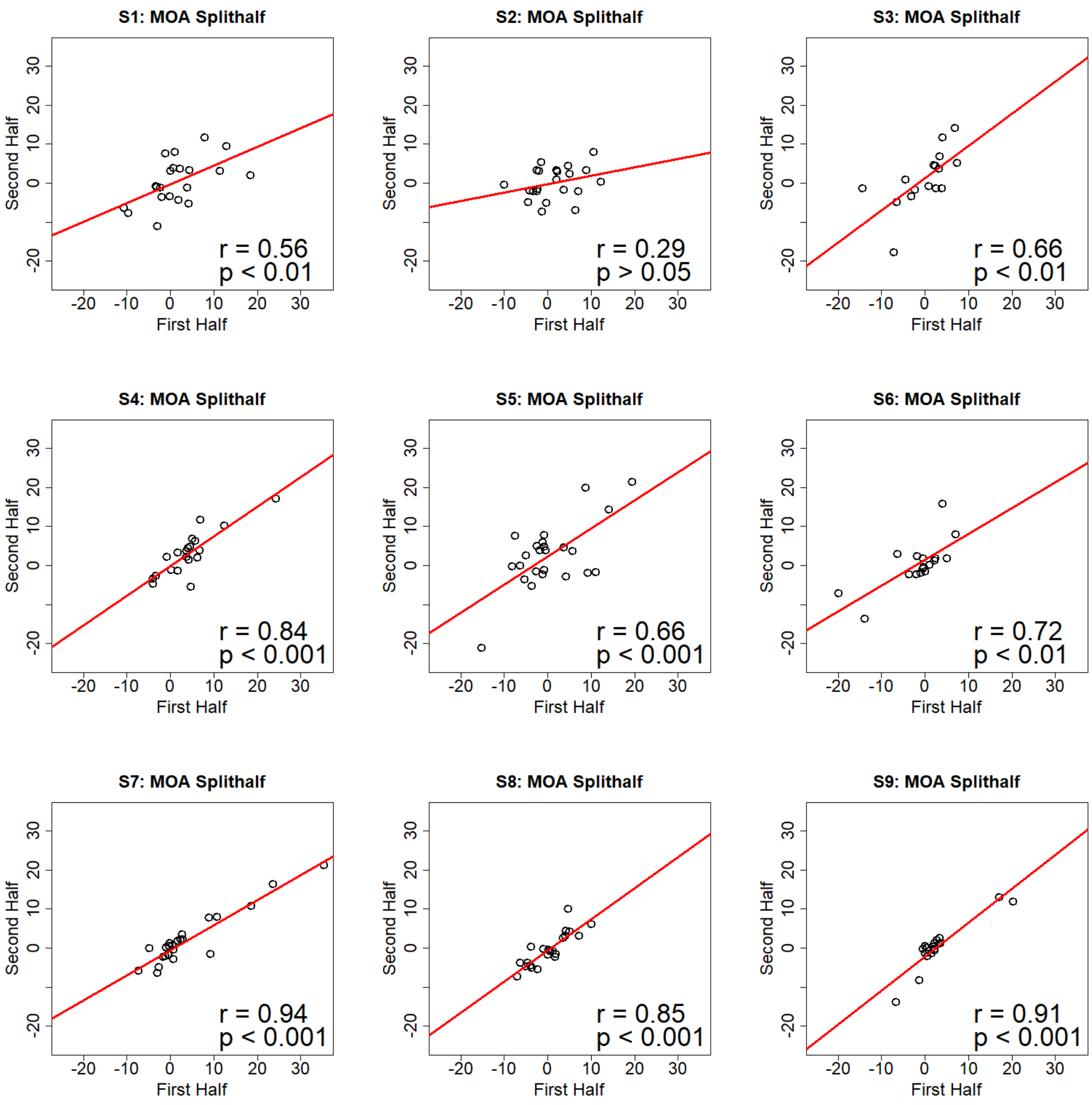
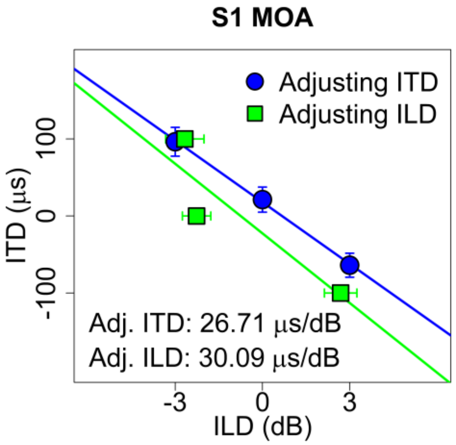
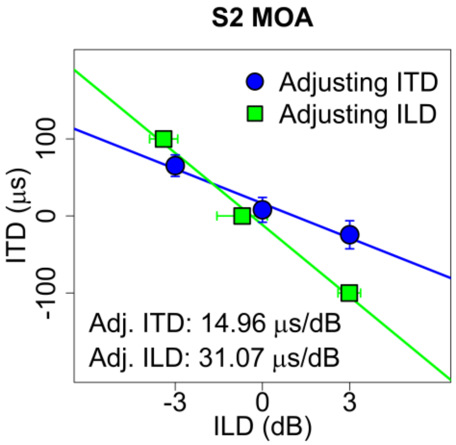
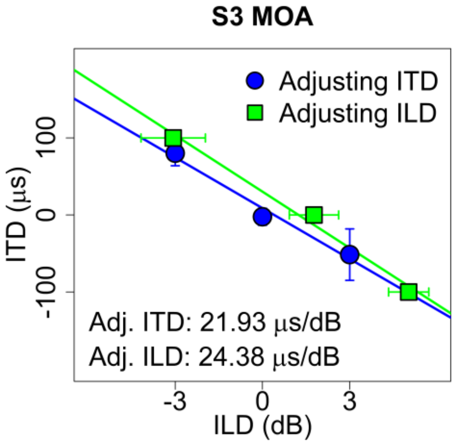
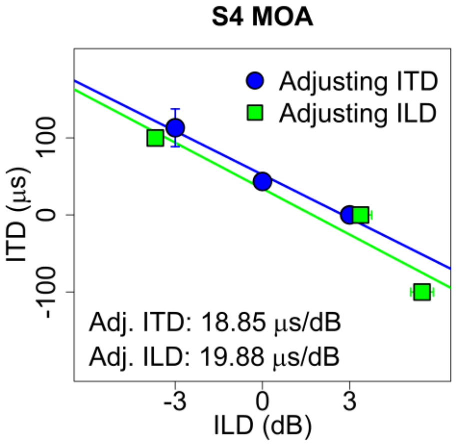


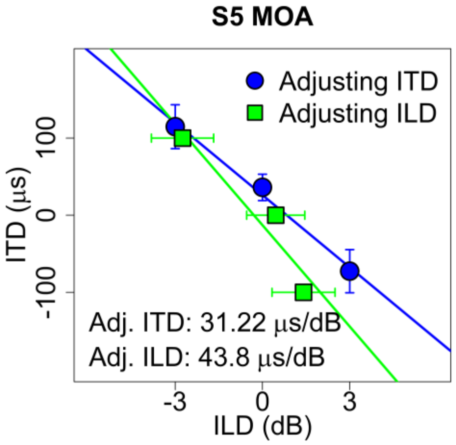
Figure 6. Split-half reliability for Experiment 1, for each participant. The last 4 judgments are plotted as a function of the first 4 judgments. Each panel includes responses collapsed across ITD_{adj} and ILD_{adj} conditions. Each circle represents a single response. The red line shows the best linear fit of the data.

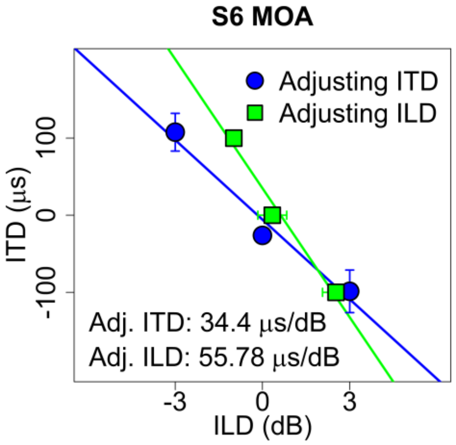

S1 MOA
Adjusting ITD
Adjusting ILD
ITD (µs)
100
0
-100
Adj. ITD: 26.71 µs/dB
Adj. ILD: 30.09 µs/dB
-3
0
3
ILD (dB)


S2 MOA
Adjusting ITD
Adjusting ILD
ITD (µs)
100
0
-100
Adj. ITD: 14.96 µs/dB
Adj. ILD: 31.07 µs/dB
-3
0
3
ILD (dB)


S3 MOA
Adjusting ITD
Adjusting ILD
ITD (µs)
100
0
-100
Adj. ITD: 21.93 µs/dB
Adj. ILD: 24.38 µs/dB
-3
0
3
ILD (dB)


S4 MOA
Adjusting ITD
Adjusting ILD
ITD (µs)
100
0
-100
Adj. ITD: 18.85 µs/dB
Adj. ILD: 19.88 µs/dB
-3
0
3
ILD (dB)


S5 MOA
Adjusting ITD
Adjusting ILD
ITD (µs)
100
0
-100
Adj. ITD: 31.22 µs/dB
Adj. ILD: 43.8 µs/dB
-3
0
3
ILD (dB)


S6 MOA
Adjusting ITD
Adjusting ILD
ITD (µs)
100
0
-100
Adj. ITD: 34.4 µs/dB
Adj. ILD: 55.78 µs/dB
-3
0
3
ILD (dB)

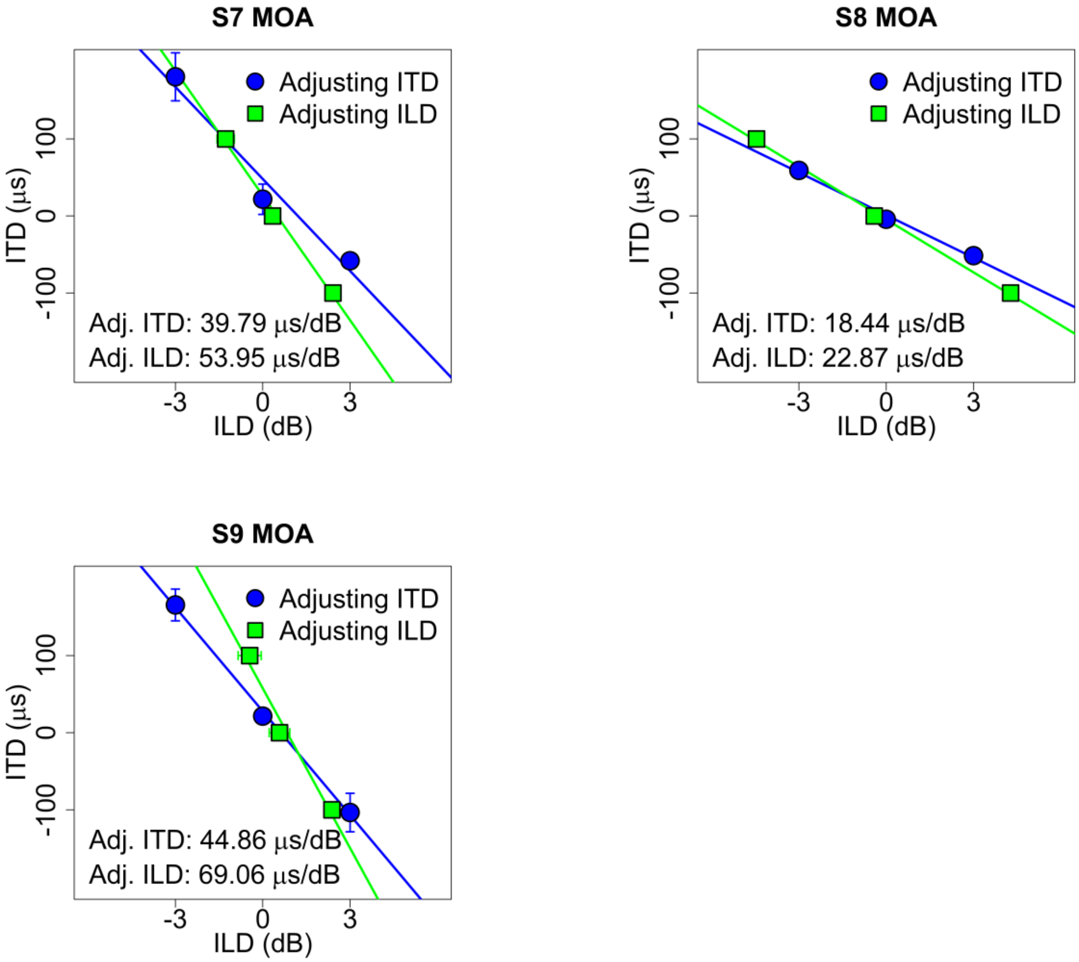


Figure 7. Individual TRs from Experiment 1. Blue circles indicate the required ITD (μs) to offset a variety of fixed ILDs (fixed values labeled along the abscissa in dB). Green squares indicate the required ILD (dB) to offset a variety of fixed ITDs (fixed values labeled along the ordinate in μs). Error bars denote standard error of the mean. Each panel represents data from one participant. The slopes of the respective data points were taken as trading relations and are given in the lower left of the panels.

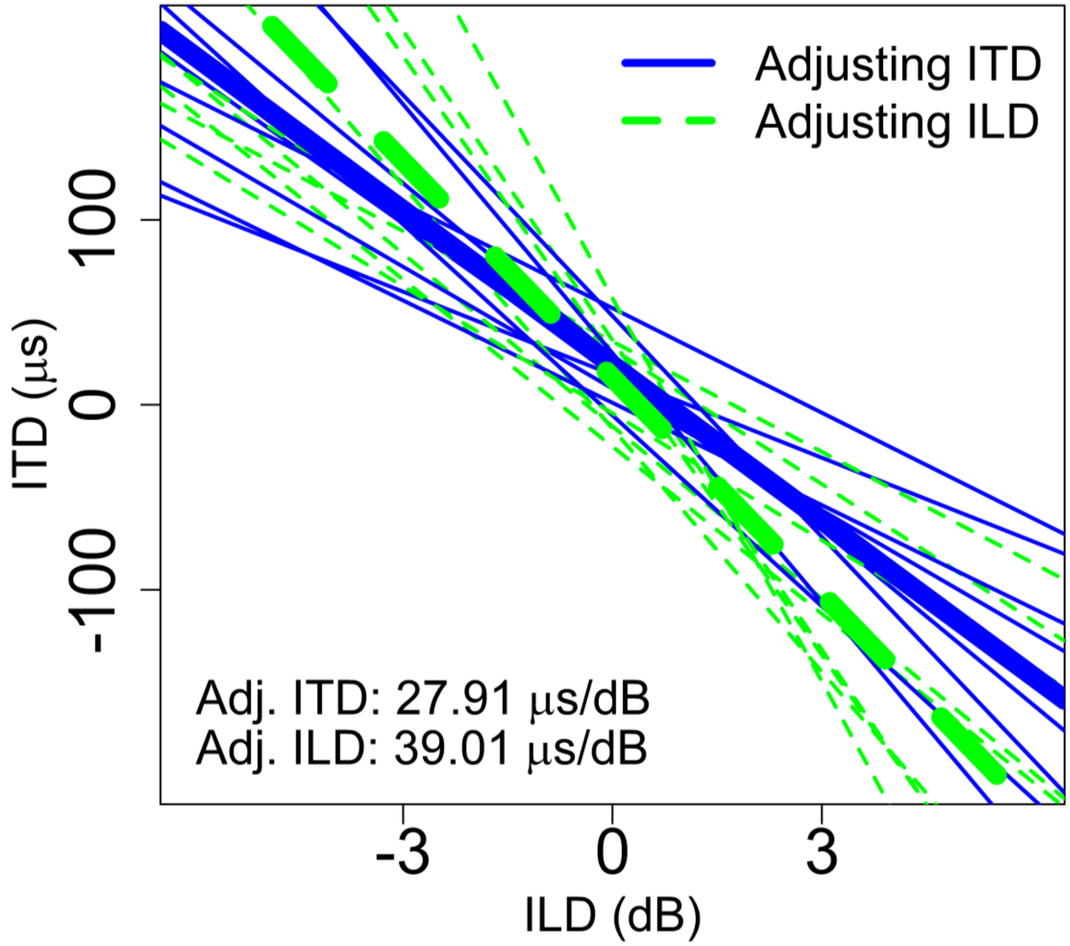


Figure 8. Mean TRs from Experiment 1. Thin, solid blue lines represent the slopes from all participants while adjusting the ITD. Thin, dashed green lines represent the slopes from all participants while adjusting the ILD. Thick lines (solid and dashed) show the group mean slopes (i.e., TRs) when participants adjusted the ITD and ILD, respectively.

3.3 Discussion

Consistent with the literature, the results of this study show that TRs obtained using the MOA differ based on the cue being adjusted. Furthermore, the relationship between the TRs in this study replicate existing findings showing the $\mathrm{ITD}_{\mathrm{adj}}$ produces a smaller TR (i.e., shallower slope) compared with the $\mathrm{ILD}_{\mathrm{adj}}$ TR. These results will serve as a basis for comparison with TRs obtained using the MOCS within the same individuals (Experiments 2 and 3). The current study findings are considered within the context of the existing literature below.

3.3.1 TRs from MOA tasks in the literature

Centering tasks

Röser (1965) reported on a number of interaural trading values from various studies, which ranged from 1.7 to 230 µs/dB. In this most general test, the mean TRs of the current experiment fit within the extremes of previously observed limits. The actual TRs from a number of studies using the MOA are provided in Table 1. The table includes centering tasks, similar to the method used here, as well as related pointing type tasks (described in the Introduction). When comparing TRs, it is important to keep in mind that early centering studies measured TRs for the $\mathrm{ITD}_{\mathrm{adj}}$ condition only.

The $\mathrm{ITD}_{\mathrm{adj}}$ TR measured here (27.9 µs/dB) is most similar to the ~28.3 µs/dB reported by Harris (1960). Other early $\mathrm{ITD}_{\mathrm{adj}}$ TRs range from 25 µs/dB (David Jr et al., 1959) to 63.6 µs/dB (Deatherage & Hirsh, 1959), with larger TRs resulting from stimuli other than pure tones (e.g., clicks and impulses). The current findings are most dissimilar to the earliest reported TRs: Klemm (1920) observed a TR of 130 µs/dB, while Shaxby and Gage (1932) measured a TR of 1.7 µs/dB.

The TR from Klemm (1920) likely differs due to the disparity in stimulus type. In Klemm's study, participants adjusted a telephone ring stimulus, thus it seems more appropriate to compare the TR from Klemm to the clicks used by David Jr et al. (1959). Depending on sensation level, David Jr. and colleagues reported click TRs over 200 µs/dB at 10 dB SL and approximately 100 µs/dB presented over a range of 30 dB SL to 50 dB SL, which are more comparable with Klemm (1920). Conversely, the extremely small TR from Shaxby and Gage (1932), may be due to "dual images," which have been postulated to skew results if participants unwittingly respond to a mixture of "time" and "intensity" images during the same session/experiment (e.g., Hafter & Jeffress, 1968).

More recently, Lang and Buchner (2009) reported an $\mathrm{ITD}_{\mathrm{adj}}$ TR roughly half that of the present study (i.e., 12.9 µs/dB). They also measured an $\mathrm{ILD}_{\mathrm{adj}}$ TR, whose value was 19.9 µs/dB, which is also approximately half that of the $\mathrm{ILD}_{\mathrm{adj}}$ seen here. Thus the absolute magnitude of the TRs differ from this experiment, but the relationship between TRs is roughly consistent (within about 3 µs/dB). Stecker (2010) obtained an $\mathrm{ILD}_{\mathrm{adj}}$ TR higher than both values just mentioned (i.e., 72.4 µs/dB), but used an unorthodox response technique (i.e., adjusting cue via head-turn) and click train stimuli, which have been shown to result in higher TRs (e.g., Deatherage & Hirsh, 1959). Considering the wide variability in TRs between subjects and across tasks, this survey of the literature suggests existing TRs are generally in agreement with the results reported here. It seems reasonable to assert the current MOA task successfully replicated the existing literature.

Pointing tasks

Studies employing pointing tasks mostly focus on cue-dependent trading through the use of training listeners to become sensitive to two "images" when complementary cue types are put into opposition. The main difference from centering tasks is that instead of directly adjusting one cue to offset the complementary cue to midline, participants adjust the ITD (most typically) of a pointer with a 0 dB ILD (most typically) to match the position of an eccentric marker consisting of opposing cue values. While positioning the pointer, listeners match either an image driven almost entirely by the ITD, or another image composed of a combination of ITD and ILD. In this type of paradigm, the time image TR is typically quite small, as evidenced in Table 1. In essence, it seems participants can always identify the ITD, but do not perceive an image driven solely by the ILD. While this phenomenon is quite interesting, it may not add much to the topic of cue trading per se. First, the ITD has been shown to dominate the ILD in wideband stimuli (e.g., Macpherson & Middlebrooks, 2002), suggesting ITD is generally the more potent cue. Second, low frequency ILD cues do not occur naturally (i.e., only under headphones) and may be more easily ignored by the auditory system. Third, the auditory image has been shown to become more diffuse and harder to localize with increasing ILD, rendering the intensity cue itself less discernible under certain conditions. It may be due to these factors that well-practiced listeners are sometimes able to parse the binaural stimulus to find the ITD cue. This author feels the time image contributes little to the percept arising from the mixture of antagonistic ITD and ILD cues. TRs derived from the level image have always included the effect of both cues; in fact, "*level-image*" is shorthand for an image that contains both time and intensity (for review see Hafter & Jeffress, 1968; Whitworth & Jeffress, 1961). This point is discussed further in the next subsection. For clarity, the level-image is henceforth referred to as the *time/intensity* image. Participants in the current study were instructed to center the leftmost image should two images be perceived, in order to account for the possibility of randomly centering one or the other image within a session and thereby artificially reducing the observed TR.

Collapsing across tasks

When considering TRs derived from the time/intensity image, the values are similar to $\mathrm{ITD_{adj}}$ TRs from centering experiments. For example, the time/intensity TRs from Table 1 range from approximately 18 µs/dB (Moushegian & Jeffress, 1959) to 30 µs/dB (Domnitz & Colburn, 1977; Hafter & Jeffress, 1968). TRs from both centering and pointing tasks, excluding those obtained from the time image, all cluster around 20 µs/dB to 50 µs/dB when measured using stimuli similar to the current study. Thus, the image of interest, the time/intensity image, seems relatively stable across task type. It is also interesting to note in centering tasks that directly offsetting the cues leads to TRs well above ~1 µs/dB for both $\mathrm{ITD_{adj}}$ and $\mathrm{ILD_{adj}}$. One of the few studies to measure both $\mathrm{ITD_{adj}}$ and $\mathrm{ILD_{adj}}$ TRs using a centering task, Lang and Buchner (2009) also show TRs well above the traditional time image, reporting TRs of 12.9 µs/dB and 19.9 µs/dB for $\mathrm{ITD_{adj}}$ and $\mathrm{ILD_{adj}}$, respectively.

Taken together, one interpretation of the literature as a whole is that the cue-dependent nature of TRs may be of a smaller magnitude than indicated by “dual image” studies, where time/intensity TRs are compared with time image TRs. That the time image can be discerned with adequate practice might be more of an indication the ITD is the dominant cue under certain conditions, and that ITD and ILD cues are ultimately processed in separate perceptual channels. However, the percept arising from a mixture of opposing cues, the time/intensity image, suggests binaural cues are capable of interaction. The nature of this interaction, as investigated via directly offsetting one cue with the complementary cue, suggests there is at least some tendency for the $\mathrm{ITD_{adj}}$ TR to be smaller than the $\mathrm{ILD_{adj}}$ TR; however, this discrepancy might be accounted for by the type of task. If it can be shown that binaural spatial adaptation can be induced using timing and stimulus parameters common to MOA tasks, it would constitute a step toward accounting for the difference between ITD_{adj} and ILD_{adj}.

Table 1. Trading relations from the literature using tasks and stimuli similar to the current study. When multiple intensities or frequencies were tested, the reported trading relations reflect those most similar to the parameters of this study. Under the stimulus column, “pointer” refers to the stimulus adjustable by the listener, while “target” refers to a stimulus off-midline that the pointer must match in perceived laterality.

Study	**Task**	**Stimulus**	**Trading Relation**
Klemm (1920)	Centering ITD_{adj}	Telephone ring	130 µs/dB
Shaxby and Gage (1932)	Centering ITD_{adj}	Tone (0.5 kHz)	1.7 µs/dB
David Jr et al. (1959)	Centering ITD_{adj}	1) Impulse 2) Click	1) ~ 25 µs/dB 2) ~ 50 µs/dB
Deatherage and Hirsh (1959)	Centering ITD_{adj}	Click train	~ 63.6 µs/dB
Harris (1960)	Centering ITD_{adj}	Tone (0.5 kHz)	~ 28.3 µs/dB
Young Jr and Levine (1977)	Centering ITD_{adj}, ILD_{adj}	Tone (0.5 kHz)	ITD_{adj}: ~ 40.4 µs/dB ILD_{adj}: ~ 79.4 µs/dB
Lang and Buchner (2009)	Centering ITD_{adj}, ILD_{adj}	Tone (0.5 kHz)	ITD_{adj}: ~ 12.9 µs/dB ILD_{adj}: ~ 19.9 µs/dB
Stecker (2010)	Centering ILD_{adj}	Click train	72.4 µs/dB
Moushegian and Jeffress (1959)	Pointing ITD_{adj}	Pointer: Noise (0.1 – 3 kHz) Marker: Tone (0.5 kHz)	SS1: 2.5 µs/dB SS2: 18 µs/dB SS3: 27 µs/dB
Whitworth and Jeffress (1961)	Pointing ITD_{adj}	Pointer: Tone (0.5 kHz) Marker: Tone (0.5 kHz)	Time: 0.65 µs/dB Int: 19.9 µs/dB
Hafter and Jeffress (1968)	Pointing ITD_{adj}	Pointer: Tone (0.5 kHz) Marker: Tone (0.5 kHz)	Time: ~ 5 µs/dB Int: ~ 30 µs/dB
Domnitz and Colburn (1977)	Pointing ILD_{adj}	Pointer: Tone (0.5 kHz) Marker: Tone (0.5 kHz)	30 µs/dB

3.3.2 Range of tradable ILD values

While the overall results of this study were consistent with the existing literature, participants were only able to offset a truncated range (i.e., 0 dB and ±3 dB) of the fixed ILD values included here. This finding warrants attention, because the full range of ILD values presented here have been used successfully in other studies. For example, Hafter and Jeffress (1968), Whitworth and Jeffress (1961) and Moushegian and Jeffress (1959) all obtained responses out to an ILD of ±9 dB. All of these studies also used the pointing method, whereas the current study used the centering method. Thus, one potential explanation for the difference in tradable ILD values may be due to specific differences across tasks. For instance, it may be an easier task to manipulate the pointer, which consists of an adjustable cue with the complementary cue held at midline (i.e., either 0 dB or 0 µs), which allows a participant to "point" across the entire useable range of a given azimuthal cue. This is in stark contrast to the centering technique, where cues are directly opposed and must offset the other completely to midline.

This logic is supported by studies using the centering task. For instance, Harris (1960) stated: "As the ILD is increased, the image spreads out and becomes harder to locate. For large intensity differences, the image sometimes splits." For this reason, he restricted the range of fixed ILDs to no greater than 6 dB. Despite this precaution, he still found that centering accuracy decreased with increasing ILD. Harris (1960) also reported pure tone stimuli produced less accurate data compared to clicks, with participants indicating pure tone stimuli sounded very diffuse. Lang and Buchner (2009) measured TRs with fixed ILD values up to ±7.5 dB. They included a "Not enough" option participants could use if they were unable to center the stimulus. The "Not enough" option was checked most often when the fixed ILD value was ±7.5 dB.

It has become clear that the present study should have chosen fixed ILD values based on the centering literature, rather than a study using a pointing task (i.e., Whitworth & Jeffress, 1961). Future studies should take as many study parameters into account as possible when building on existing work relating to TRs.

Chapter 4

EXPERIMENT 2: MOCS (HEAD-POINTING TASK)

4.1 Experimental methods

4.1.1 Stimuli

The stimuli in Experiment 2 were synthesized using the parameters described in the General Methods. Similar to Experiment 1, cue combinations consisted of fixed ILD values (0, ±3, ±6, or ±9 dB), and fixed ITD values (0, ±100, ±200, or ±300 μs). In this experiment, all possible combinations of ITD and ILD were used, resulting in 49 different combinations.

4.1.2 Procedure

Participants completed a “localization” task with insert earphones. That is, unfiltered pure tones resulted in intracranial images that participants mentally extrapolated into space. Stimuli were presented using the method of constant stimuli (MOCS). Participants wore the Oculus HMD and were immersed in the same virtual environment as Experiment 1 (see Figure 4). Participants were seated in a swivel chair and held the right Oculus controller. Pulling the trigger started a brief animation (three balloons bobbing) to indicate the trigger pull had been read and the trial had begun. Participants had to position the head-locked reticle into a green box at midline in order for the stimuli to play. Each trial consisted of the presentation of a single, 500-ms tone containing one of the 49 possible cue combinations. Participants were required to keep their heads centered until the stimulus had completely finished playing. Participants were then instructed to extrapolate the intracranial image outside the head, and to position the reticle over the balloon at that location via head turn. Consistent with existing reports (e.g., Jeffress &

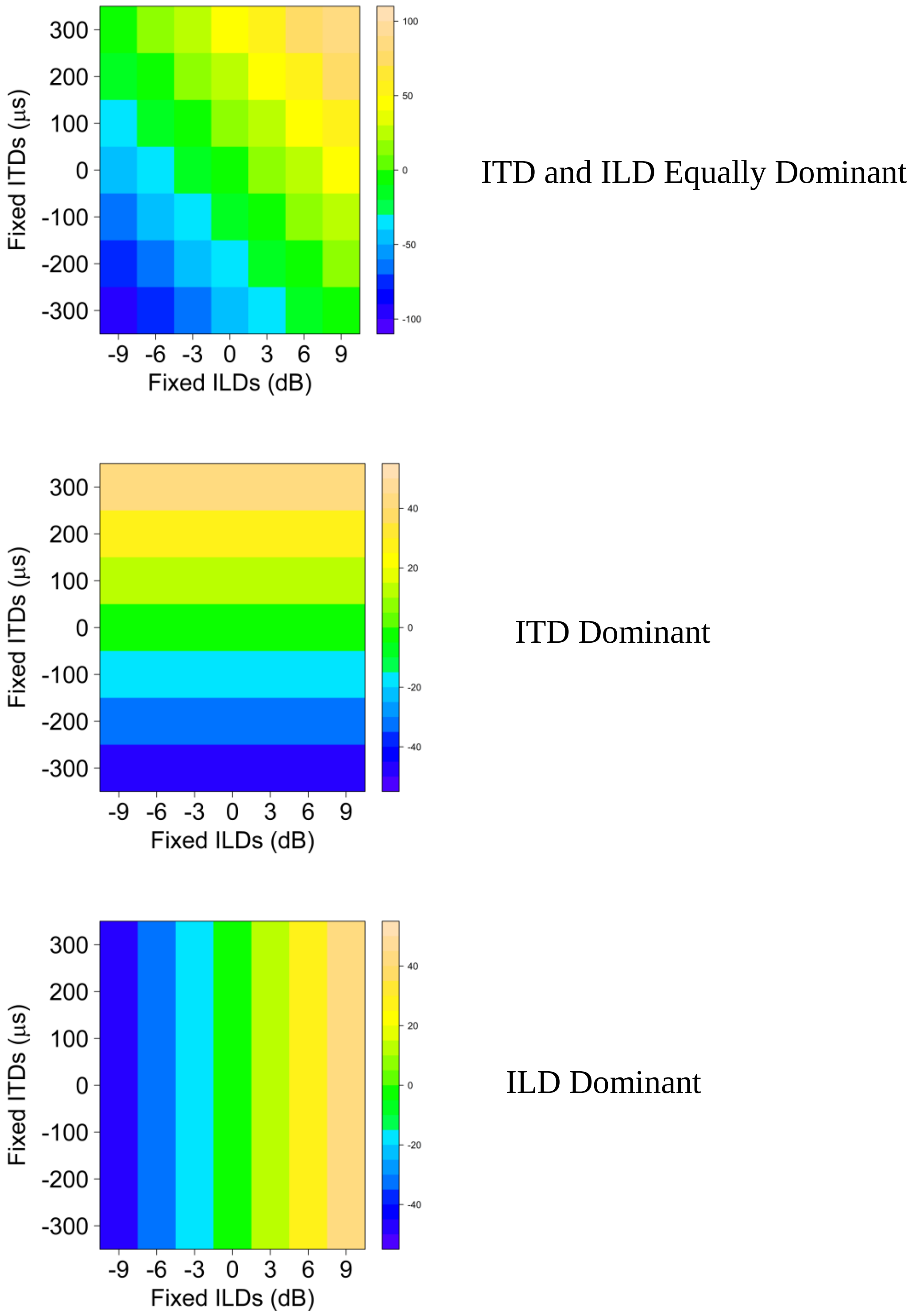
Taylor, 1961; Stecker, 2010), participants had no difficulty extrapolating the lateralized stimulus to an external location. Once the reticle was satisfactorily aligned with the perceived azimuth of the tone, the participant pulled the trigger on the Oculus Rift controller. Immediately after the trigger pull, the balloon “popped,” providing visual confirmation of the selection, and recording the current head position. The next trial began after the participant recentered the reticle inside the green box at midline, with a delay of 2 seconds. Participants made 8 judgments for each cue combination (392 responses).

There were no mixed or fixed presentation patterns for this experiment, due to the nature of the task: participants were all presented single presentations from the same pseudorandomly chosen 49 cue combinations.

4.1.3 Data and Analyses

The azimuth of the head position when a participant responded was recorded for each trial. That is, the data collected represent the participants’ perceived azimuth of the stimuli. To account for differences in potential bias and range of responses across the session, individual data were normalized to z-scores. The 8 judgments per condition were averaged into a single data point for plotting and analyses. Participant responses to all 49 cue combinations were plotted in a 7 X 7 matrix with individual values represented as colors (i.e., a heatmap). Following the convention from Experiment 1, fixed ITD values are plotted along the ordinate, while fixed ILD values are plotted along the abscissa. The color scale ranges from blue (to the extreme left) to tan (to the extreme right) with midline represented by green. For comparison, idealized heatmaps are given in Figure 9, and represent maps resulting from a completely dominant ITD, a completely dominant ILD, and equal dominance between the cues.

TRs were calculated from an individual’s data by fitting contour lines to the heatmaps. Because the MOA data from Experiment 1 were the result of a centering task, wherein participants offset the complementary cue so that the auditory percept appeared at midline, the contour line of interest for the MOCS task was at 0, indicating the cue values at which participants perceived the stimuli at midline. TRs were calculated by fitting the data points of the contour lines at 0 using linear regression, taking the resultant slope as the TR.



ITD and ILD Equally Dominant

ITD Dominant

ILD Dominant

Figure 9. Idealized heatmaps for the MOCS task, showing scenarios where the ITD and ILD cues are equally dominant (top), the ITD is dominant (middle), and where the ILD is dominant (bottom). Fixed cue values are plotted along the axes, the response parameter is perceived azimuth.

4.2 Results

4.2.1 Descriptive statistics for the MOCS task

Reliability of the data over time was measured using the split-half method (described in Experiment 1). Individual results are shown in Figure 10. Judgments were similar across time, with correlations ranging from $r = 0.79$ to $r = 0.94$. This finding indicates the first and second half of judgments made similar contributions to the mean.

Individual heatmaps using color to display the magnitude of the SEM are provided in Figure 11. These plots show the extent of deviation around the mean of the 8 judgments for each cue combination, after conversion to z-scores. Visual inspection revealed the lowest deviations typically occurred for smaller values of ILD, consistent with Harris (1960); however, some participants exhibited very low deviation around the mean even at the largest fixed ILD values.

The contour lines at 0 were fit using linear regression to characterize the slope of the values at which participants perceived the stimuli at midline. Linear functions accurately represented the contour lines at 0, with correlations ranging from $r = 0.88$ to $r = 0.99$.

A Shapiro-Wilk normality test revealed the data for this experiment were normally distributed.

4.2.2 TRs obtained using the MOCS

Data from all 9 participants were included in the analyses. Individual heatmaps using color to display perceived azimuth are shown in Figure 12. Individual and mean slopes of the contour line at 0 (i.e., TRs) are superimposed in Figure 13. The mean TR was 40.8 µs/dB (range = 20.2 to 64 µs/dB; SEM = 5.08 µs/dB). The overall diagonal pattern of the data indicates the ITD and ILD carried similar perceptual weight for most cue combinations.

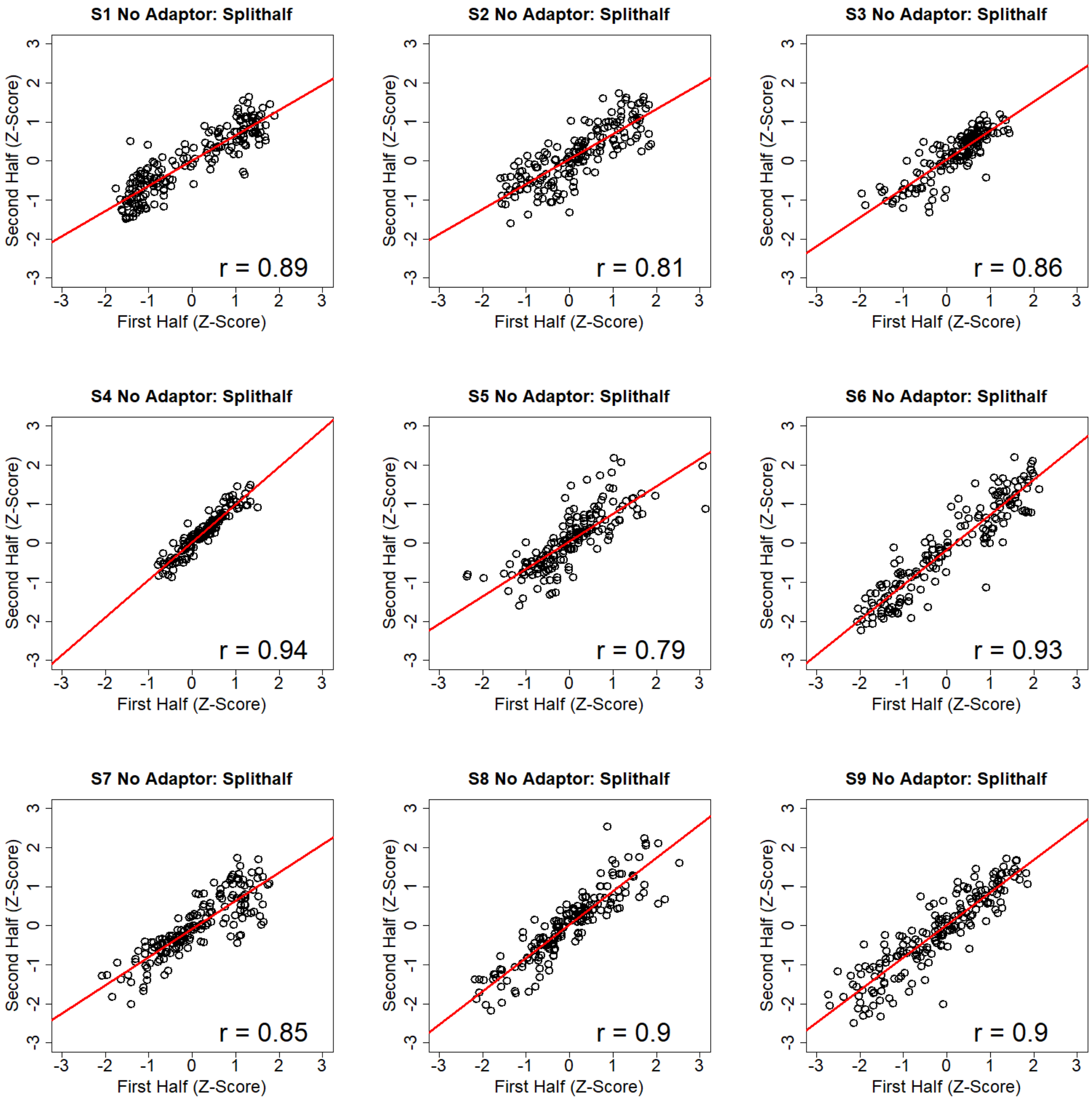
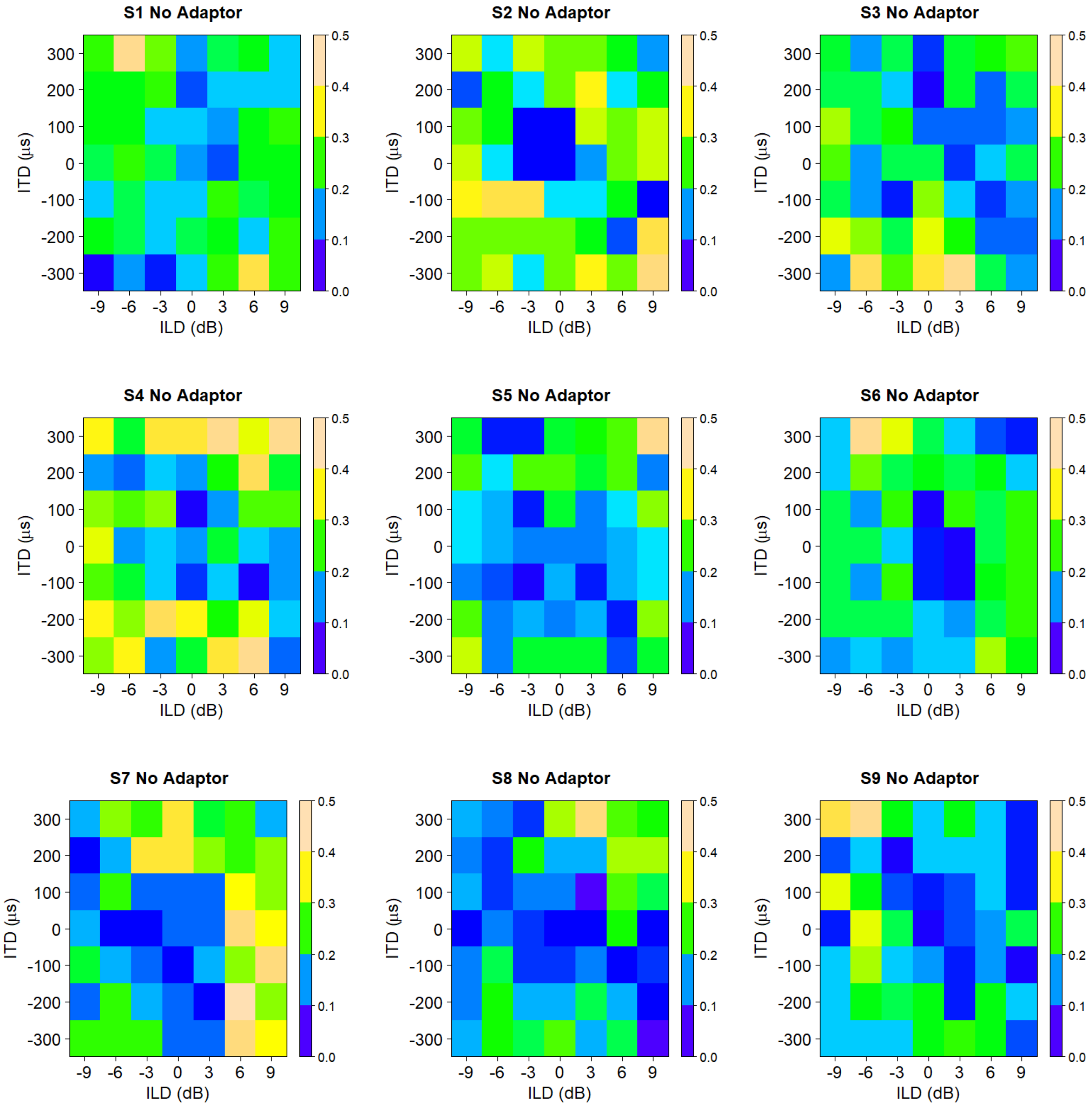


Figure 10. Split-half reliability for Experiment 2, for each participant. The first- and second-half participant responses are plotted on the abscissa and ordinate, respectively. Each data point represents a single judgment. The red line shows the linear regression fit of the data.



S1 No Adaptor

S2 No Adaptor

S3 No Adaptor

S4 No Adaptor

S5 No Adaptor

S6 No Adaptor

S7 No Adaptor

S8 No Adaptor

S9 No Adaptor

Figure 11. Individual SEM heatmaps for Experiment 2. The color in each square displays the amount of deviation around the mean of the 8 judgments in each cue combination (after conversion to z-scores).

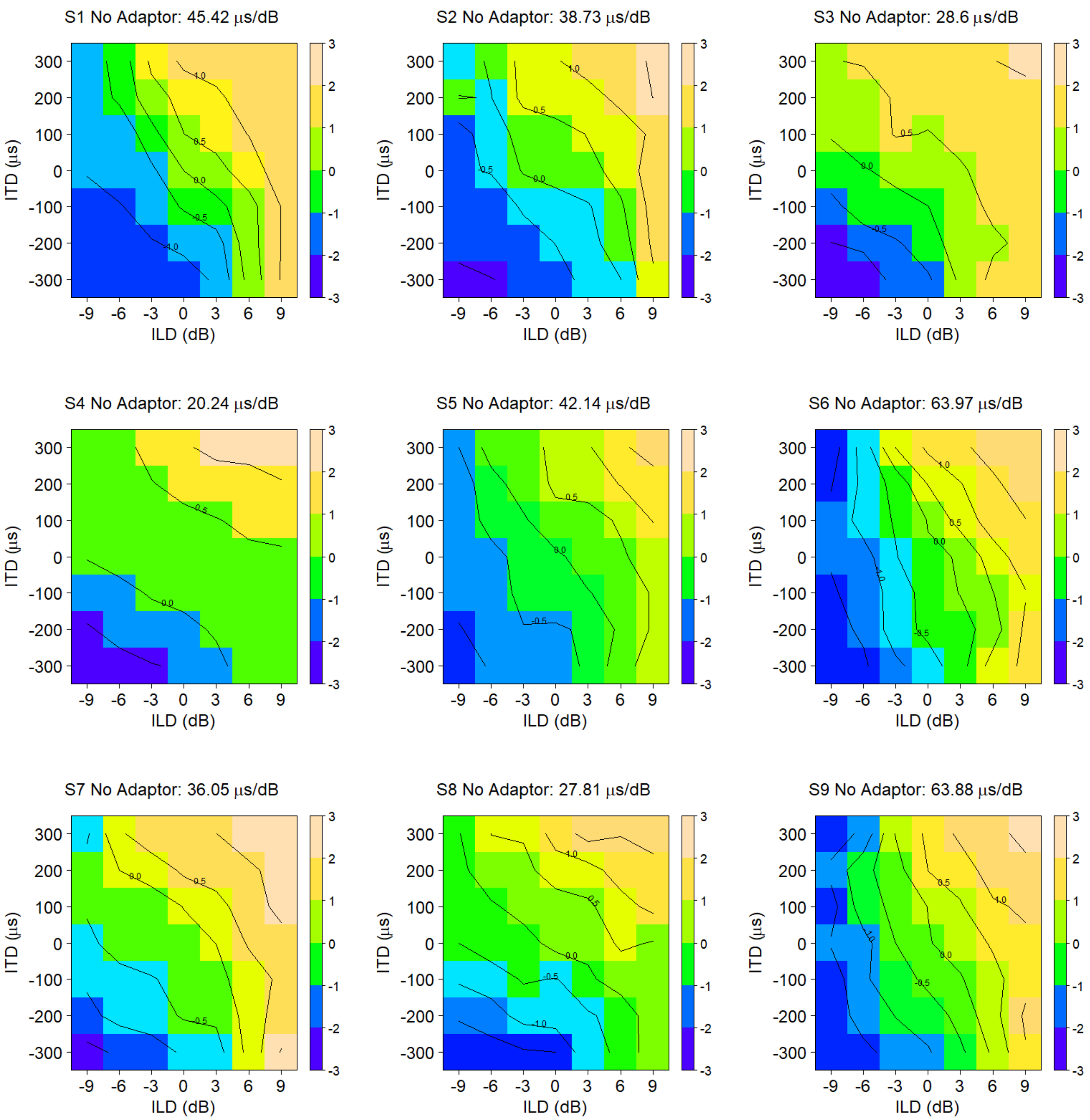


Figure 12. Individual perceived-azimuth heatmaps for Experiment 2. Colors represent perceived azimuth, and are fit with contour lines. TRs were calculated from the slope of the contour line at 0, using linear regression. The TR values are provided in the panel titles in µs/dB. The color scale represents left, midline and right locations using blue, green and tan, respectively.

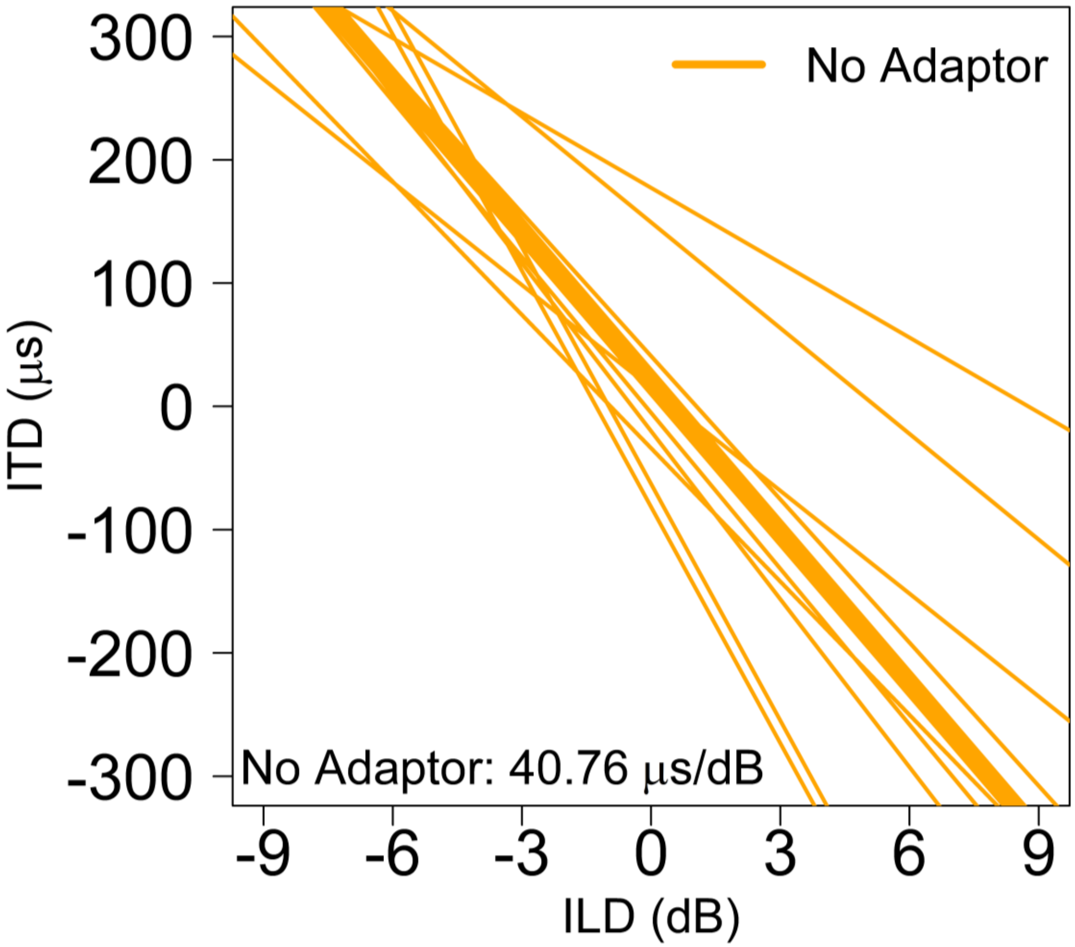


Figure 13. Mean and individual slopes for Experiment 2 (thick and thin lines, respectively). Slopes are plotted on the same scale as the heatmaps, where they were derived from the contour line at 0 (cue combinations at which listeners perceived the stimulus at midline). The value of the mean slope (i.e., mean TR) is given in the lower left-hand corner.

4.3 Discussion

4.3.1 TRs from MOCS tasks in the literature

Stecker (2010) reports TRs for a paradigm similar to the current experiment. Participants indicated the perceived location of single stimuli containing different combinations of ITD and ILD (-225 to 225 μs in 50 μs steps; and -5 to 5 dB in 2 dB steps). The response technique utilized an electromagnetic position sensor that captured participant head movements, allowing for a response technique that was similar to the present study; listeners indicated the perceived location of the stimulus by orienting their heads to the extrapolated azimuth of the intracranial image. Despite the similarities in task and response technique, Stecker (2010) observed a TR of 80.2 μs/dB; double that found in the current study. This is most likely due to Stecker's use of click train stimuli, which have been shown to yield larger TRs than 500 Hz pure tones (e.g., Deatherage & Hirsh, 1959). Thus, the main conclusion to be drawn between the TRs of these two studies is further confirmation that broadband stimuli provide larger TRs than pure tones.

Lang and Buchner (2009), as well as studies in the same vein (Ignaz et al., 2013, 2014; Lang & Buchner, 2008), have also used the MOCS in the context of cue trading; however, the MOCS task stimuli consist solely of the chosen cue values used to center the stimuli during an MOA task. That is, after centering a stimulus using the MOA, these authors then presented listeners with single presentations of the cue values that were judged at midline in a MOCS task. Thus, the traditional TR in μs/dB is exactly the same across task types. The analyses in these types of studies concern the differences in perceived azimuth depending on task type, rather than TRs. For this reason, TRs cannot be compared here, but the findings of these studies are discussed in Chapter 7.

Unfortunately, the cue trading literature is rather sparse regarding lateralization tasks using the MOCS, compared with the number of studies using the MOA (centering and pointing). The majority of work not using the MOA has employed some form of forced choice task (e.g., Babkoff et al., 1973; Hafter & Carrier, 1972). For instance, Hafter and Carrier (1972) employed a 2-interval forced choice (2IFC) discrimination task using 500 Hz tones. The ISI between standard and test signals was 200 ms, which is near the most effective value of the auditory localization aftereffect (250 µs) for same-frequency adaptor and probes at 400 Hz (Kashino & Nishida, 1998). These parameters render TRs susceptible to adaptative effects, similar to the MOA. Indeed, TRs obtained from Hafter and Carrier (1972) were around 25 µs/dB, which is quite similar to the ITD_{adj} reported here from Experiment 1 (i.e., 27.6 µs/dB). Due to the sensitivity of TRs to task and stimuli, further discussion of discrimination tasks is outside the scope of this section.

Chapter 5

EXPERIMENT 3: MOCS (HEADING-POINTING WITH ADAPTORS)

5.1 Experimental methods

5.1.1 Stimuli

The stimuli were the same as in Experiment 2, with the addition of a train of 5 pure tones preceding the probe. All tones in the train were synthesized using the parameters given in the General Methods so that the train and the probe stimuli were identical (e.g., 500 Hz, 500 ms in duration), except for the binaural cues they carried. The pure tone train served as an adaptor and always contained one of the same binaural cues as the probe. The unadapted cue was presented at midline during the adaptor (i.e., 0 µs or 0 dB, depending on the cue), with the probe containing a pseudorandomly-chosen value from the same range as the previous experiments (i.e., ±300, ±200, ±100 and 0 µs, or ±9, ±6, ±3 and 0 dB, for ITD and ILD respectively). These combinations yielded the same 49 cue combinations as in Experiment 2, with the addition of a probe-matched ITD or ILD adaptor in the preceding train. Therefore, the data set from this experiment consisted of two matrices of judgments (perceived azimuth) obtained in the presence of an ITD or ILD adaptor.

In order to approximate the parameters used in the MOA task, each tone in the entire stimulus (5 adaptors and 1 probe) was separated by an ISI of 400 ms; the same ISI used in Experiment 1 to separate the standard and target tones in the MOA task. The goal of this experiment was to evoke adaptive processes during the MOCS task in a controlled, cue-specific manner, while employing the same timing and frequency parameters as the MOA task.

5.1.2 Procedure

The task was identical to that in Experiment 2, except each trial contained a stimulus composed of 6 tones (5 adaptors and 1 probe) instead of a single tone. Participants were instructed to ignore the first 5 tones, and to indicate via head turn the perceived azimuth of the last tone only. Participants were required to keep their heads centered by placing the reticle in a green box at midline until the entire stimulus finished playing. A single adaptor type was presented for all 49 combinations in any given block. Listeners made 8 judgments for each type of adaptor, for a total of 784 responses ($49_{\text{combination}}$ * 8_{judgment} * 2_{adaptor}).

As in Experiment 1, the participants were divided into Mixed and Fixed groups. Participants were assigned to the same group as in Experiment 1. Five participants completed blocks with probes preceded by either type of adaptor within the same session (Mixed), and 4 listeners completed blocks with the same adaptor type in any given session (Fixed). Trials for the Fixed group also began with 1 second of uncorrelated white Gaussian noise, at an RMS level of 65 dB relative to the level of a 500 Hz pure tone at 65 dB. Each Fixed trial thus consisted of Noise → 1.5 s Silence → Stimulus. The Gaussian noise was inserted to obviate any carryover effects from one trial to another (e.g., Ignaz et al., 2014).

5.1.3 Data and Analyses

This experiment resulted in the same data structure as Experiment 2: heatmaps depicting perceived azimuth (color) for each cue combination (ITDs along the ordinate and ILDs across the abscissa). TRs were obtained using the same methods as Experiment 2: fitting each heatmap with contour lines, and calculating the slope of the line at 0. The difference between data from Experiment 2 and this experiment is that two sets of TRs were produced: those obtained in the presence of an ITD and ILD adaptor, respectively.

The results of a Shapiro-Wilk normality test revealed the data were normally distributed. Bartlett's test for homogeneity of variance across adaptor types revealed the assumption of homogeneity was valid. Therefore, the data were analyzed using a two-way mixed model ANOVA, with a single within-subjects factor of adaptor type (ITDadaptor, ILDadaptor), and a between-subjects factor of stimulus presentation group (Mixed, Fixed). Effect sizes for the ANOVA are reported as generalized eta squared ($\hat{\eta}_G^2$), which is comparable across a wide range of research designs (Olejnik & Algina, 2003). A general recommendation for interpretation of $\hat{\eta}_G^2$ is 0.02 as small, 0.13 as medium, and 0.26 as large (Bakeman, 2005).

It is important to keep in mind the ITDadaptor condition implies the ILD is the dominant cue, and the ILDadaptor condition implies the ITD is the dominant cue. That is, with ILD plotted along the abscissa, an ITD adaptor would lead to ILD dominance, and therefore more vertical slopes.

5.2 Results

5.2.1 Descriptive statistics for the MOCS-Adaptor task

Individual plots showing the split-half reliability of the data are given in Figure 14 and Figure 15 for the ITDadaptor and ILDadaptor conditions, respectively. The first 4 and last 4 judgments contributed approximately equally to the mean, with correlations ranging from r = 0.84 to r = 0.94 for the ITDadaptor condition, and r = 0.84 to r = 0.92 for the ILDadaptor condition.

Individual heatmaps using color to display the magnitude of the SEM for each cue combination are provided in Figure 16. The heatmaps are plotted side-by-side for ease of comparison, with the ITDadaptor condition in the left column, and the ILDadaptor condition in the right column. Visual inspection shows the SEM was consistent across conditions for an individual listener. The overall trend reveals the largest SEM was for the most extreme cue combinations (e.g., a 9-dB ILD presented with a -300 µs ITD).

Heatmaps of perceived azimuth at each cue combination were fit with contour lines as in Experiment 2. Individual heatmaps are provided in Figure 17, with the ITDadaptor condition heatmaps in the left column, and the ILD condition heatmaps in the right column. Linear functions accurately described the contour lines at 0, with correlations ranging from r = 0.88 to r = 0.98 for the ITDadaptor condition, and r = 0.89 to r = 0.99 for the ILDadaptor condition.

5.2.2 Mixed vs. Fixed groups

To explore the relationship between stimulus presentation pattern (Mixed, Fixed) and adaptor type (ITDadaptor, ILDadaptor), a two-way mixed model ANOVA was conducted. Mauchly's test for sphericity failed to reject the null hypothesis, indicating variances were not significantly different from equal; therefore, no corrections were applied to the degrees of freedom. The analysis revealed no main effect of Group ($F(1,7) = 0.04$, p = n.s.), but did show a main effect of Adaptor ($F(1,7) = 9.11$, $p < 0.05$, $\hat{\eta}_G^2 = 0.31$) and a significant Group X Adaptor interaction ($F(1,7) = 8.58$, $p < 0.05$, $\hat{\eta}_G^2 = 0.29$). These results indicate that listener responses differed significantly depending on the adaptor, but collapsing responses across adaptor type did not result in a difference between groups. The interaction effect reveals that mixed-cue groups and fixed-cue groups differed significantly in how they responded as a function of adaptor type. This finding reveals the effect of stimulus presentation pattern (i.e., Group) had a significant impact on perceived azimuth during the MOCS tasks. Due to the significant interaction effect, data for this experiment will be divided into Mixed and Fixed groups for all further analyses.

5.2.3 Trading relations obtained using ITD and ILD adaptors

Mixed Group

Mean TRs for the Mixed group were 44.2 µs/dB for the ITDadaptor condition (range = 37.4 to 52.7 µs/dB), and 43.7 µs/dB for the ILD condition (range = 27 to 65.2 µs/dB). Individual (thin lines) and mean (thick lines) slopes for each adaptor type are displayed in the bottom panel of Figure 18. These findings reveal no effect of the adaptors on listeners' responses. Data from the Mixed group therefore do not support the hypothesis that binaural spatial adaptation contributes to the cue-dependent nature of TRs obtained using the MOA.

Fixed Group

Mean TRs for the Fixed group were 63.6 µs/dB for the ITDadaptor condition (range = 26.7 to 89.1 µs/dB), and 23.9 µs/dB for the ILDadaptor condition (range = 17.4 to 33.3 µs/dB). Individual (thin lines) and mean (thick lines) TRs for each adaptor type are given in the top panel of Figure 18. A paired-samples *t*-test compared the TRs between the ITDadaptor and ILDadaptor conditions. The result revealed a significant difference ($t(3) = -3.22$, $p < 0.05$, $d = 1.61$), suggesting adaptor type differentially influenced perceived azimuth for listeners in the Fixed group. These findings support the hypothesis that parameters used in MOA tasks in the cue trading literature render the task susceptible to adaptation. The nature of this relationship is explored in depth in the next chapter.

To rule out order effects as a potential cause for the difference between adaptors for the Fixed group, an ANOVA with factors of Adaptor and Day was attempted, but the model was saturated due to the reduced number of participants after subsetting listeners into the Fixed group. Therefore, two separate post-hoc analyses were performed. A Subject X Adaptor ANOVA revealed a main effect of Adaptor ($F(2,6) = 9.51$, $p < 0.05$, $\hat{\eta}_G^2 = 0.44$), indicating participants responded differently depending on the adaptor condition. Conversely, a Subject X Day ANOVA revealed no effect of Day (*F*(2,6) = 0.17, *p* = n.s.), indicating participants responded similarly regardless of session number. These findings support the validity of the Fixed group data.

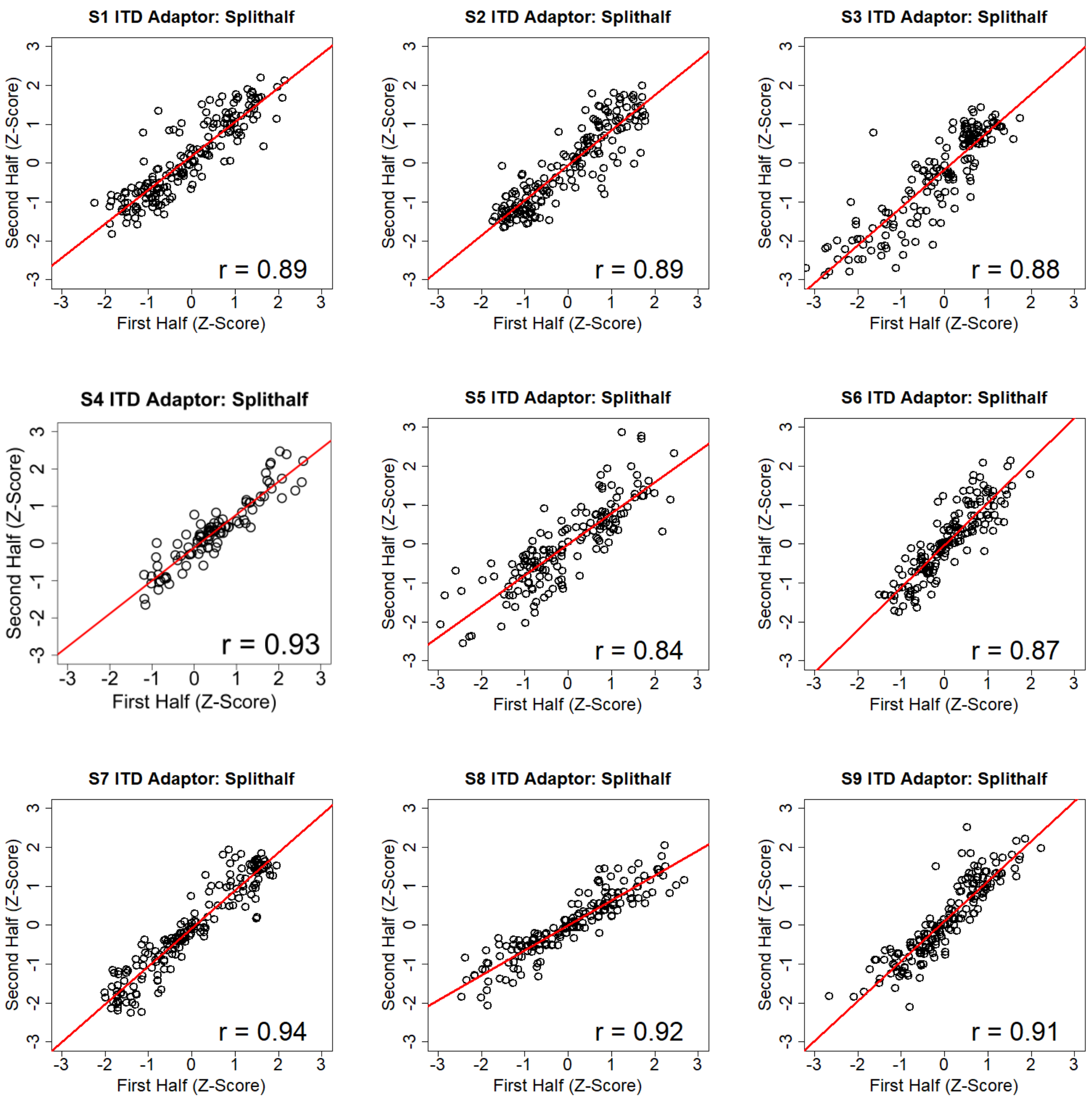


Figure 14. Split-half reliability for Experiment 3 (ITD adaptor), for each participant. The first- and second-half participant responses are plotted on the abscissa and ordinate, respectively. Each data point represents a single judgment. The red line shows the linear regression fit of the data. Participant S4 has half the number of judgments as the other listeners due to technical error.

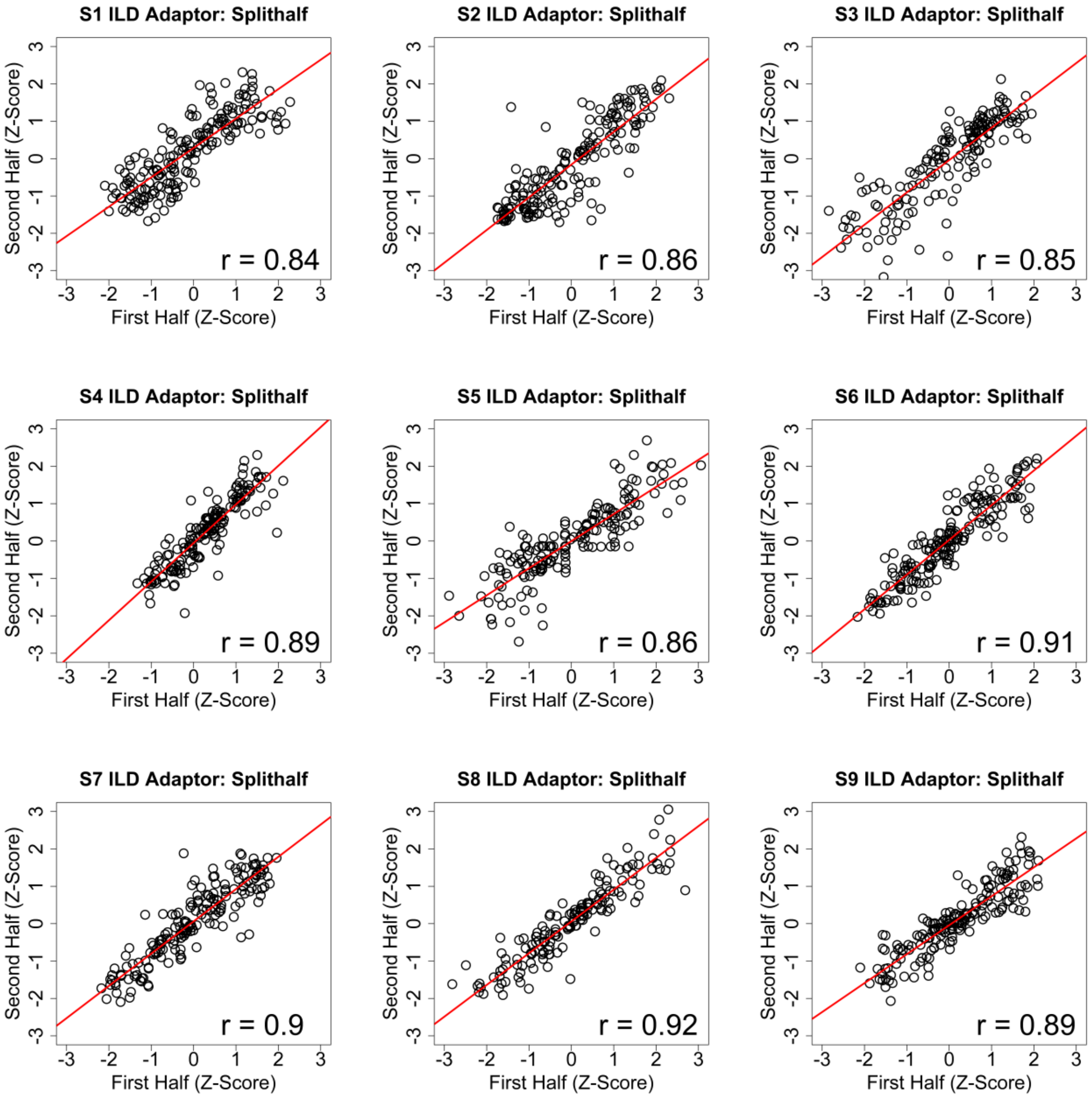
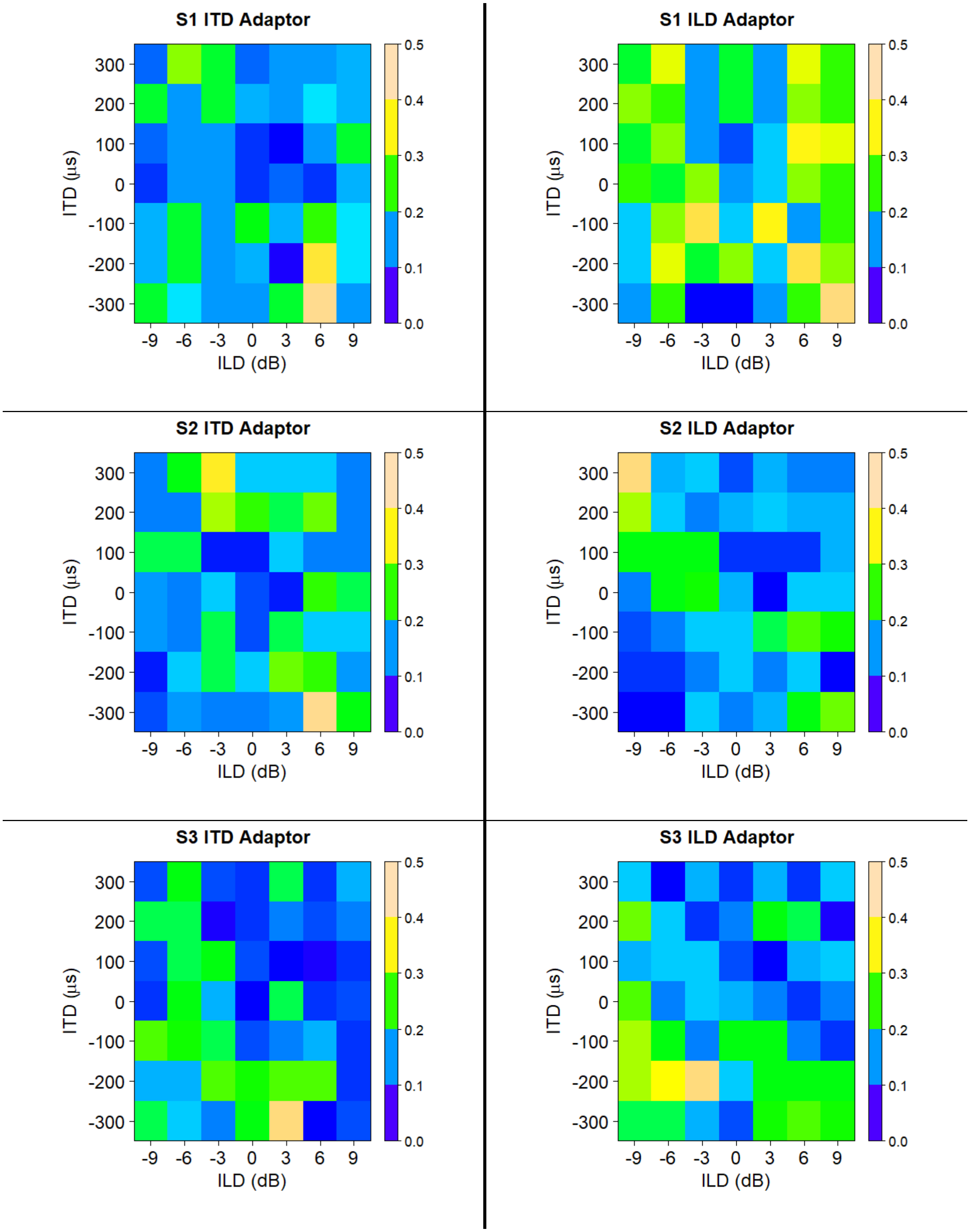
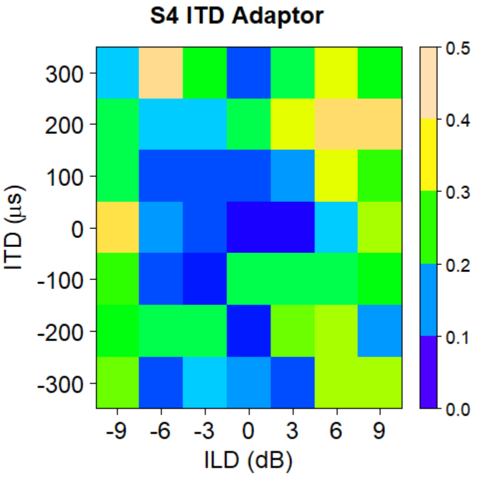
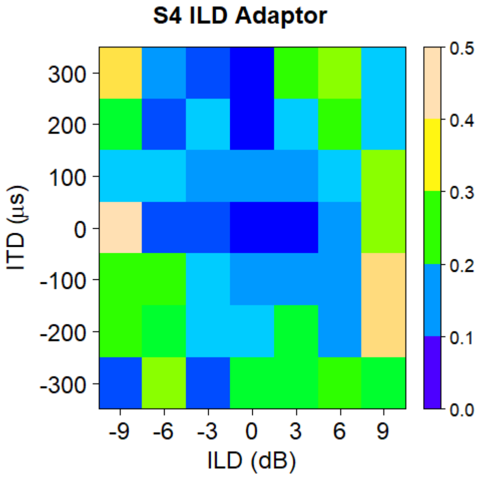
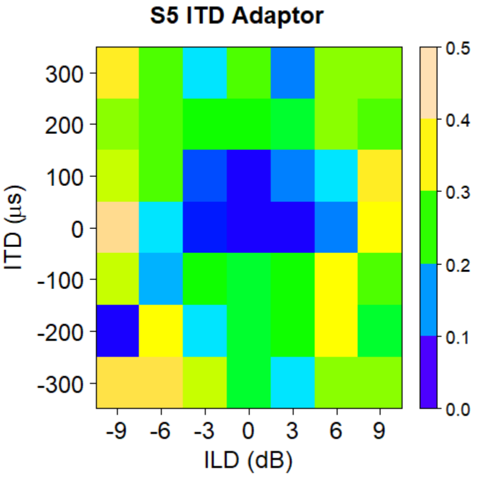


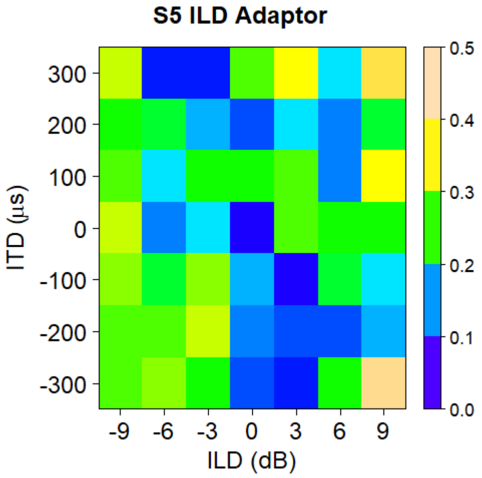
Figure 15. Split-half reliability for Experiment 3 (ILD adaptor), for each participant. The first- and second-half participant responses are plotted on the abscissa and ordinate, respectively. Each data point represents a single judgment. The red line shows the linear regression fit of the data.

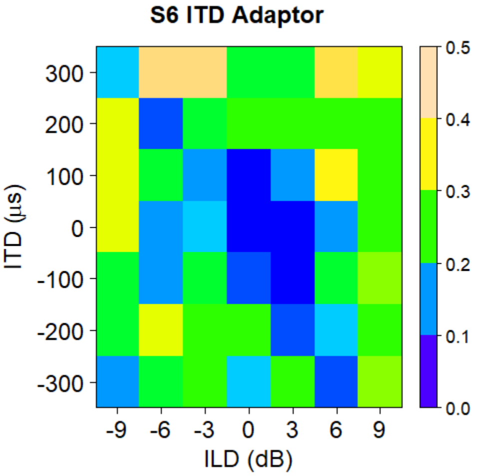

S1 ITD Adaptor
S1 ILD Adaptor
S2 ITD Adaptor
S2 ILD Adaptor
S3 ITD Adaptor
S3 ILD Adaptor
ITD (µs)
300
200
100
0
-100
-200
-300
ILD (dB)
-9
-6
-3
0
3
6
9
0.5
0.4
0.3
0.2
0.1
0.0

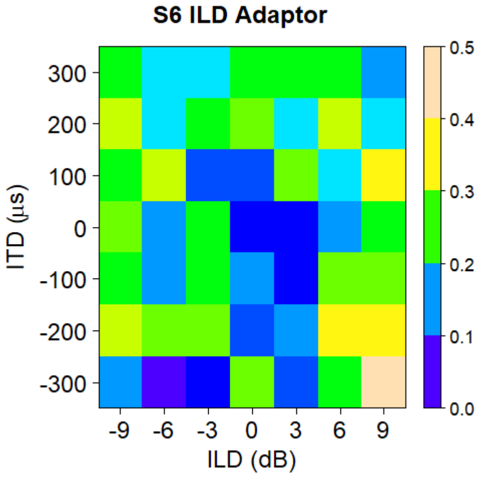

S4 ITD Adaptor


S4 ILD Adaptor


S5 ITD Adaptor


S5 ILD Adaptor


S6 ITD Adaptor


S6 ILD Adaptor

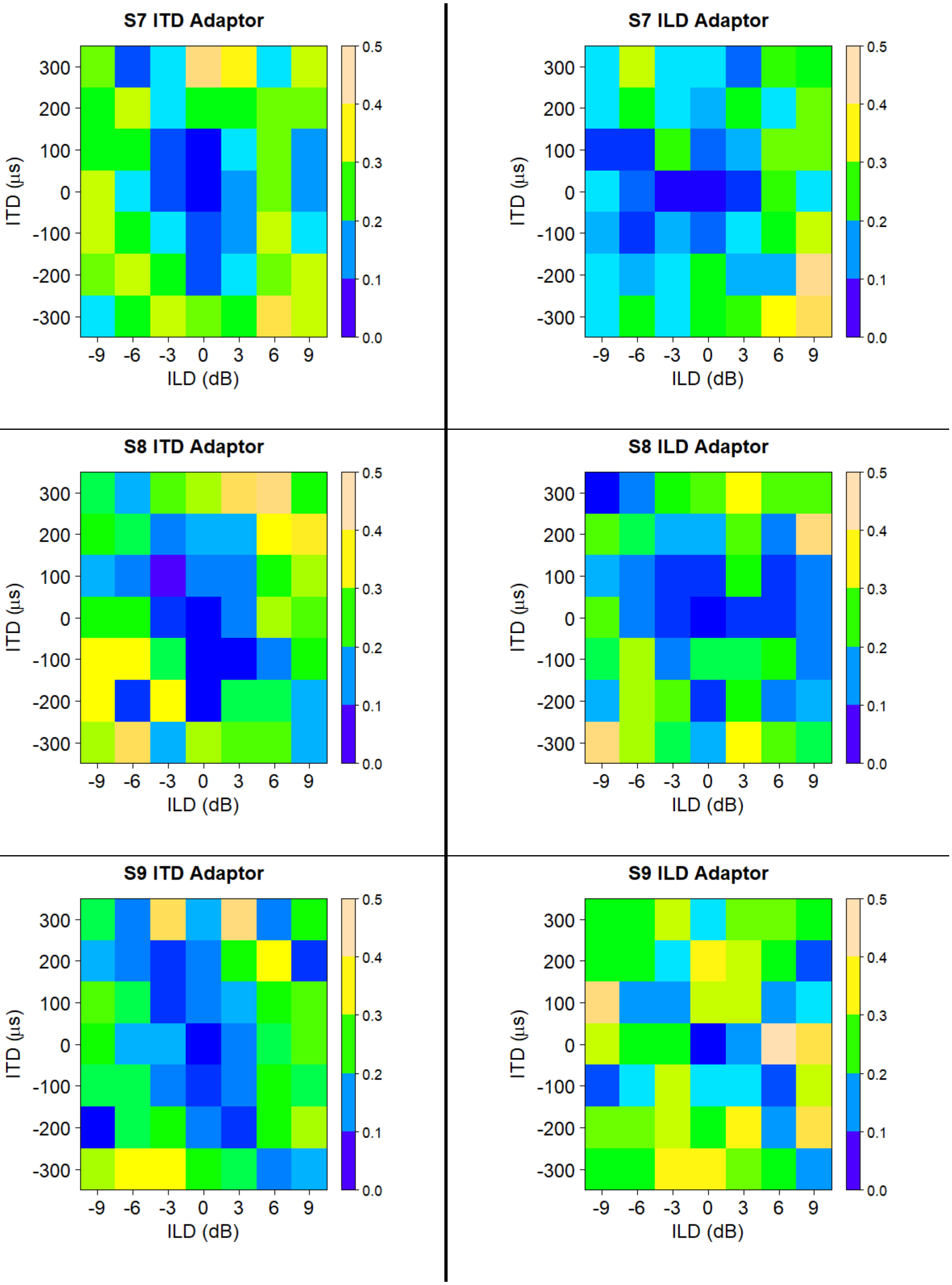
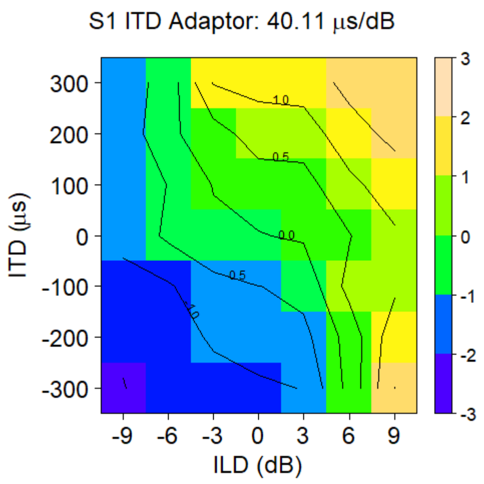
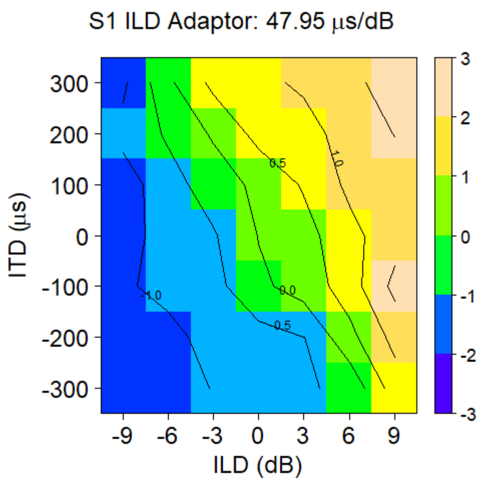
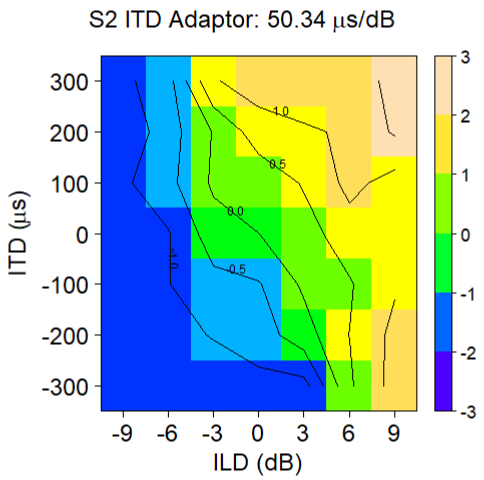
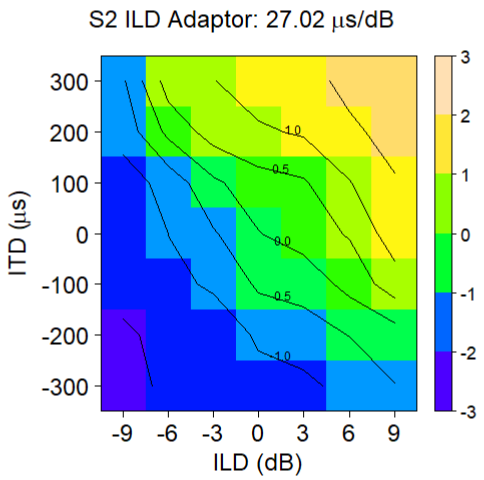


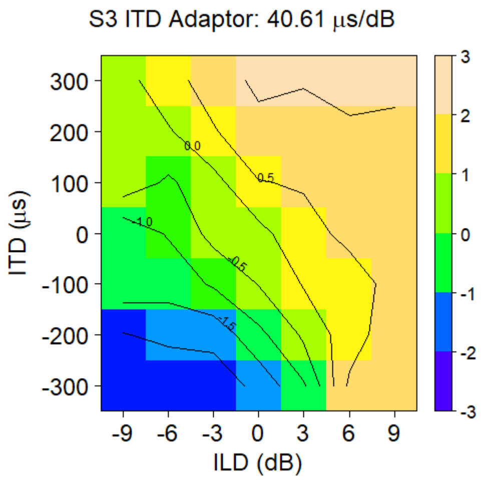
Figure 16. Individual SEM heatmaps for Experiment 3. Data for the ITDadaptor and ILDadaptor conditions are plotted in the left and right columns, respectively. The color in each square displays the amount of deviation around the mean of the 8 judgments in each cue combination (after conversion to z-scores).

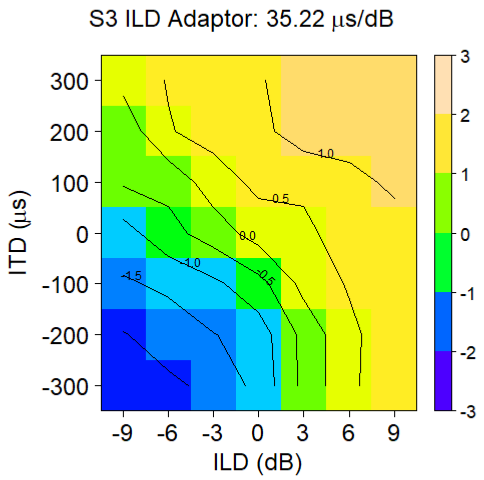

S1 ITD Adaptor: 40.11 μs/dB

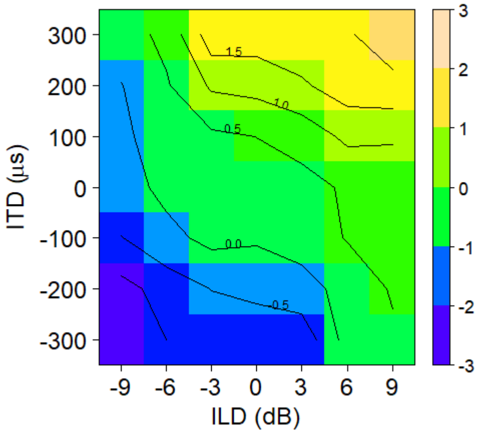

S1 ILD Adaptor: 47.95 μs/dB


S2 ITD Adaptor: 50.34 μs/dB

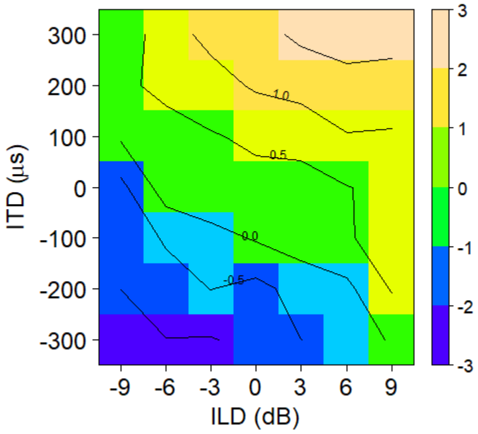

S2 ILD Adaptor: 27.02 μs/dB


S3 ITD Adaptor: 40.61 μs/dB

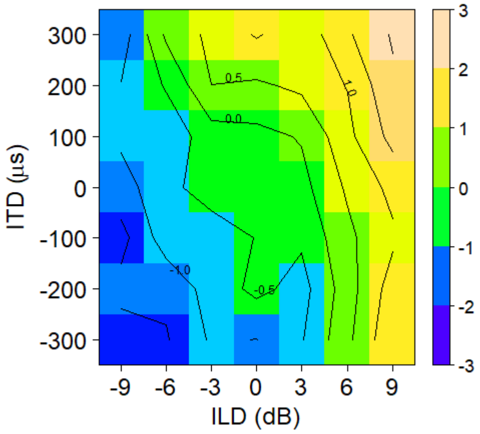

S3 ILD Adaptor: 35.22 μs/dB



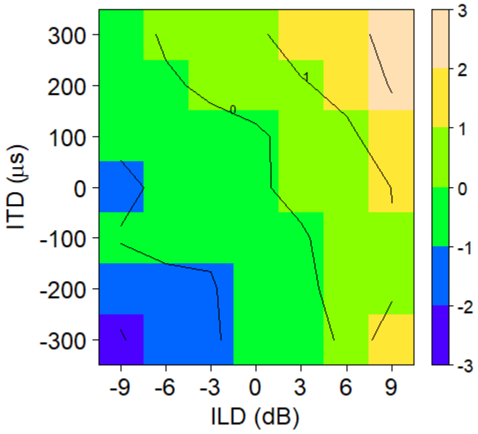
S4 ITD Adaptor: 26.72 µs/dB



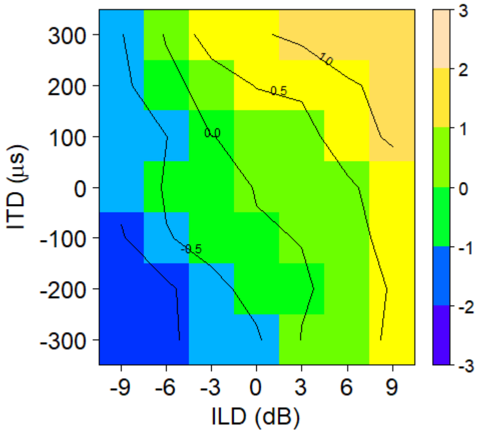
S4 ILD Adaptor: 17.37 µs/dB



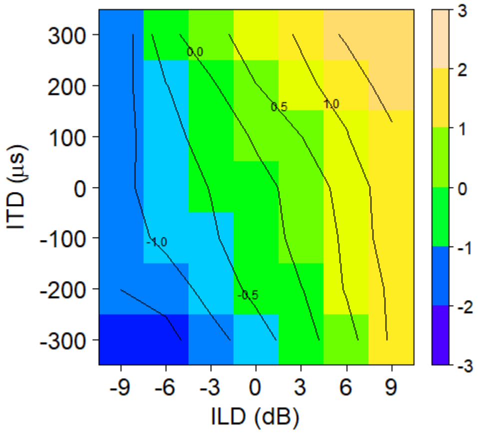
S5 ITD Adaptor: 37.39 µs/dB



S5 ILD Adaptor: 42.98 µs/dB



S6 ITD Adaptor: 52.72 µs/dB



S6 ILD Adaptor: 65.23 µs/dB

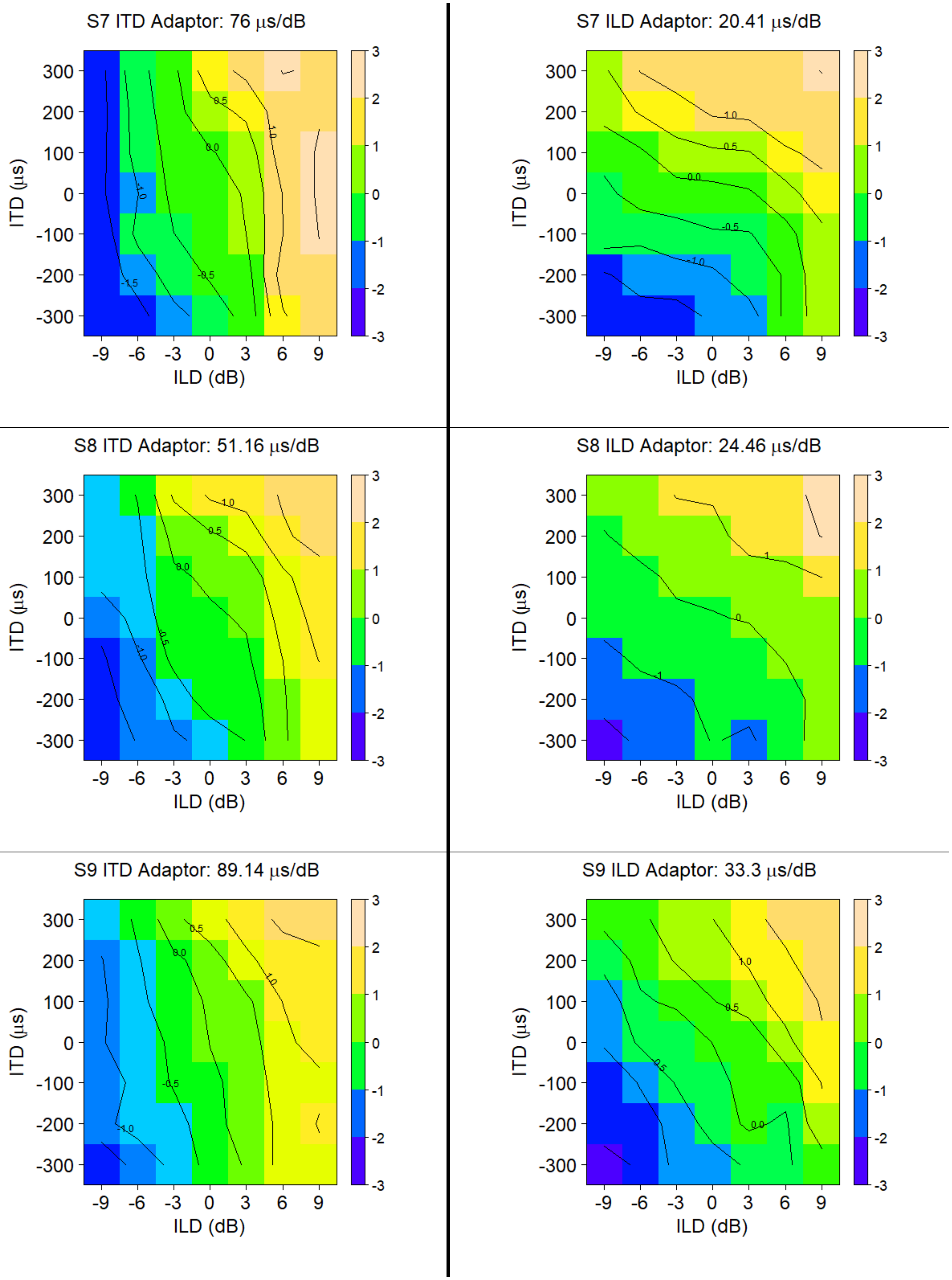


Figure 17. Individual perceived-azimuth heatmaps for Experiment 3, with the ITDadaptor and ILDadaptor conditions plotted in the left and right columns, respectively. Colors represent perceived azimuth. The TR values are provided in the panel titles in µs/dB. The color scale represents left, midline and right locations using blue, green and tan, respectively.

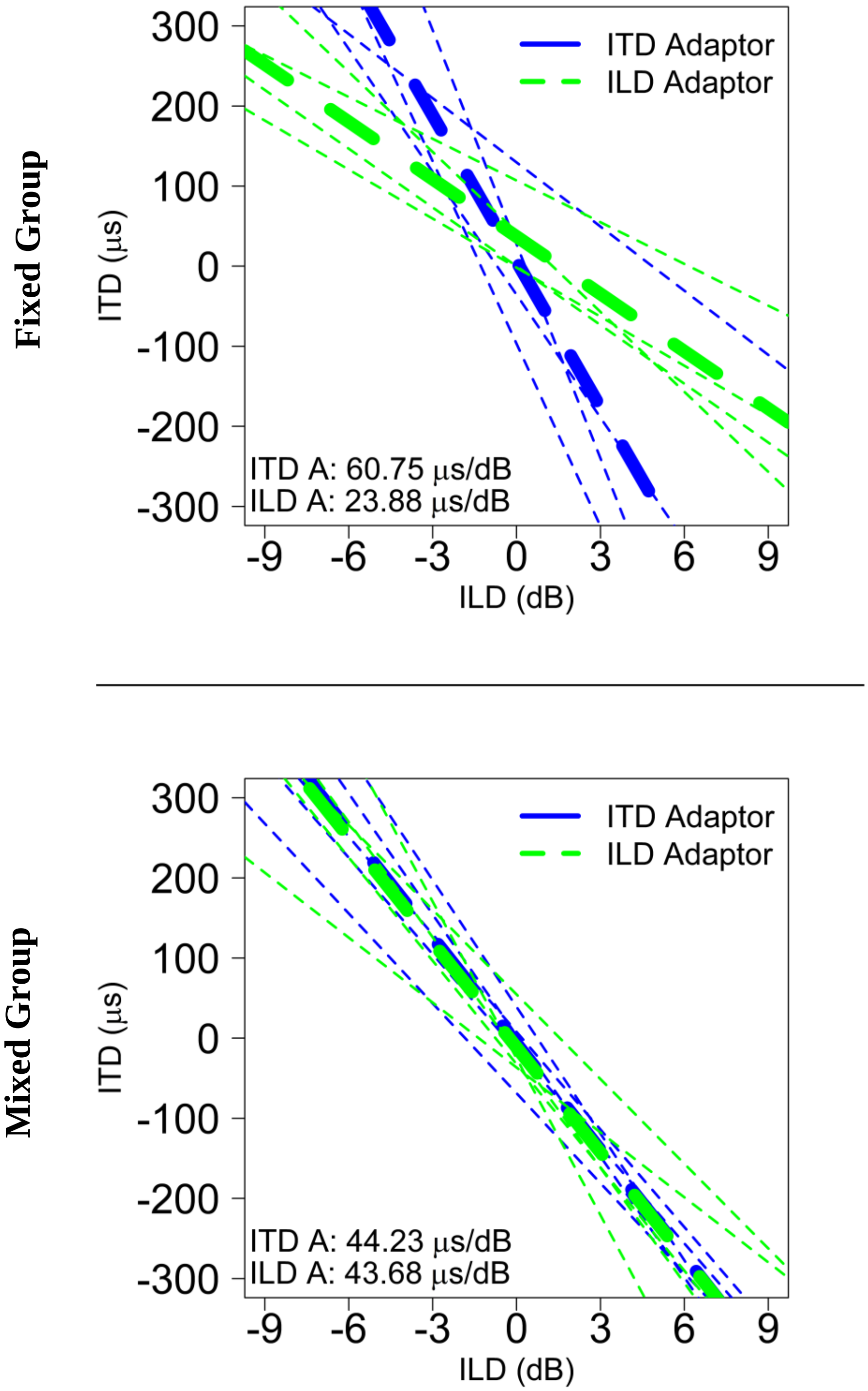


Figure 18. Mean and individual slopes for Experiment 3 (thick and thin lines, respectively). The top two panels show data from the Fixed group, and the bottom two panels show data for the Mixed group. Slopes are plotted on the same scale as the heatmaps, where they were derived from the contour lines at 0 (cue combinations at which listeners perceived the stimulus at midline).

5.3 Discussion

The results of this experiment confirmed the adapting trains were successful in biasing the perceived azimuth of probe tones using a modified MOCS task, at least for the Fixed group. This was an essential step in determining whether adaptation could reasonably occur during the MOA task. At this point it seems reasonable to infer that adaptation can, and likely does, occur during repeated presentations of fixed cues inherent to the MOA. This finding is significant because it can explain the variation in cue effectiveness, and thus differing TRs, while centering an acoustic image using the MOA. More specifically, it can be argued that the adjusted cue changes as the acoustic image is centered and is therefore not adapted; however, the fixed cue remains the same throughout the entire trial and becomes susceptible to adaptive effects. The next chapter examines the nature of adaptation by considering the data from all three experiments.

5.3.1 Mixed and Fixed groups

The difference across stimulus presentation patterns in this experiment is an interesting finding, because previous studies have been of the fixed-cue type by nature of the study design. The existing literature either measured a single type of cue interaction (e.g., adjusting only the ITD; Harris, 1960), or tested cues separately in different experiments (e.g., Lang & Buchner, 2009). The finding that some listeners (i.e., Mixed group) are not sensitive to adaptive effects during an MOCS task with adaptors is a novel finding that has implications for future study design, and reinforces the sensitivity of cue trading to task and stimulus parameters. It is also interesting to note neither the MOA task nor the no-adaptor MOCS task showed differences between Mixed and Fixed groups. More work is needed to investigate this phenomenon.

Chapter 6

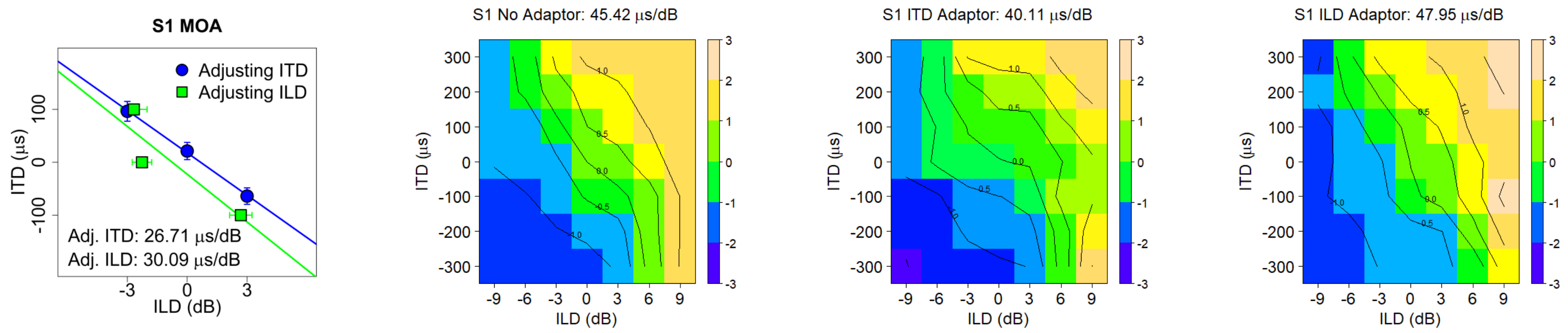
COMPARISON OF TRADING RELATIONS ACROSS EXPERIMENTS

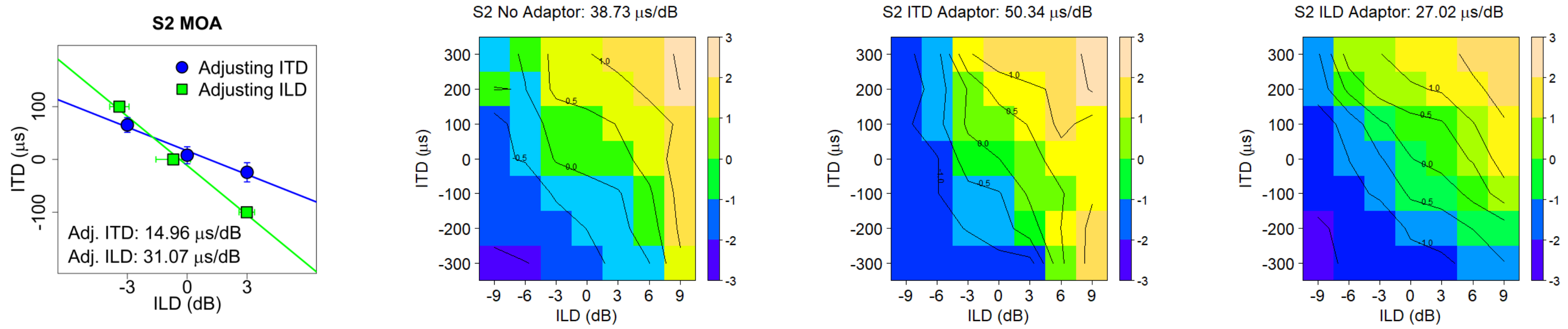
6.1 Results

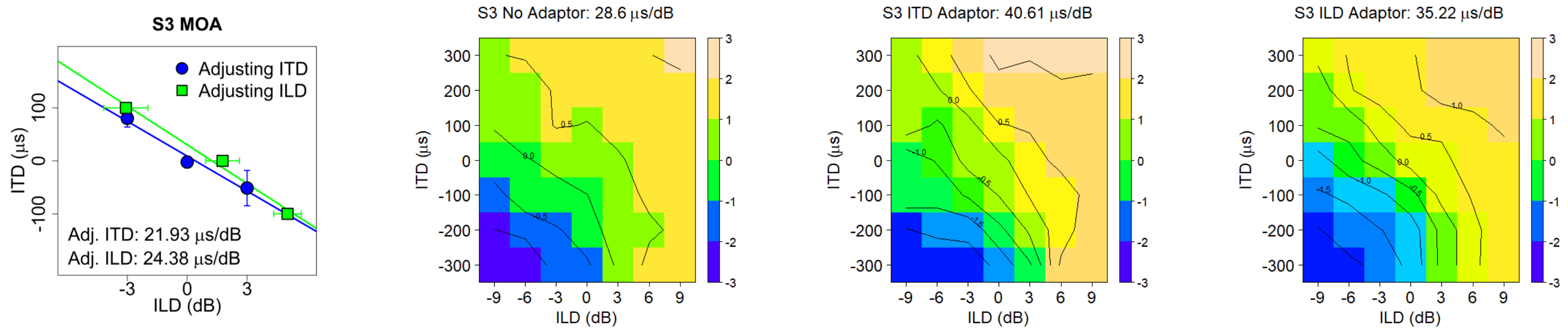
This chapter examines the relationships between TRs obtained across the entire study, with the purpose of discovering whether binaural spatial adaptation influences the cue-dependent TRs obtained using the MOA. The individual data across all experiments and conditions are presented for each listener in Figure 19. This figure is provided for ease of reference. The mean TRs from each experiment and condition are given in Table 2.

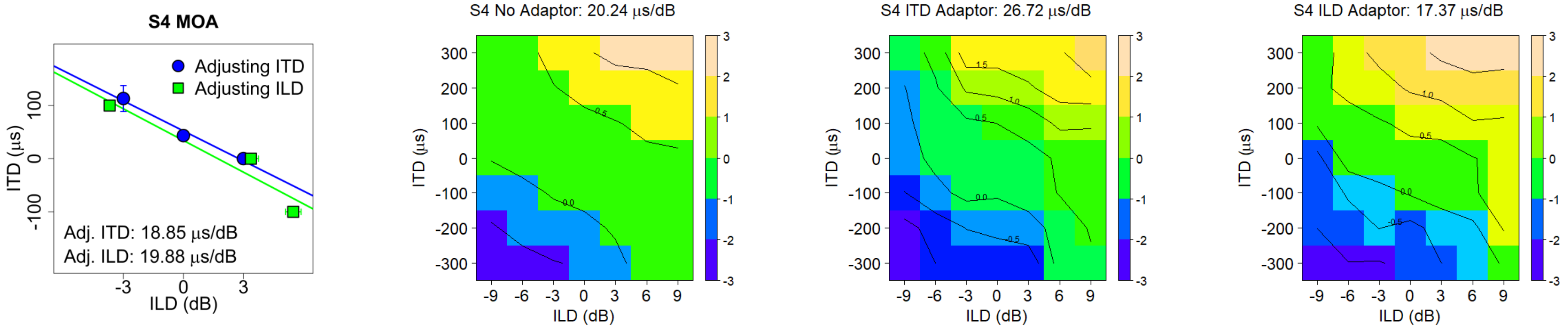
6.1.1 Experiment 2 vs. Experiment 3

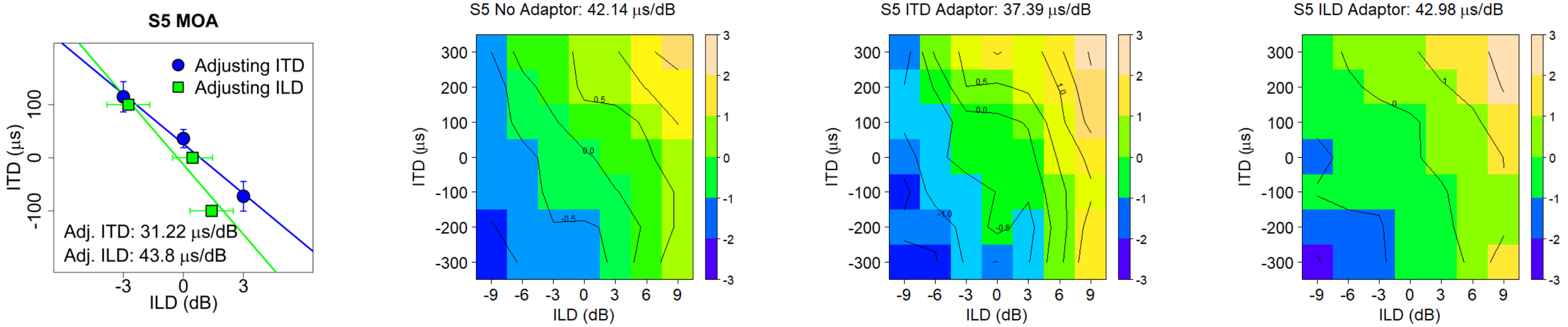
The first comparison of interest in investigating the role of adaptation in cue trading was between the TR from the no-adaptor MOCS task and the TRs from the two adaptor conditions. Because the stimulus parameters were the same as those used during the MOA task, evidence of adaptation between Experiments 2 and 3 would bolster the claim that binaural adaptation occurs during the MOA.

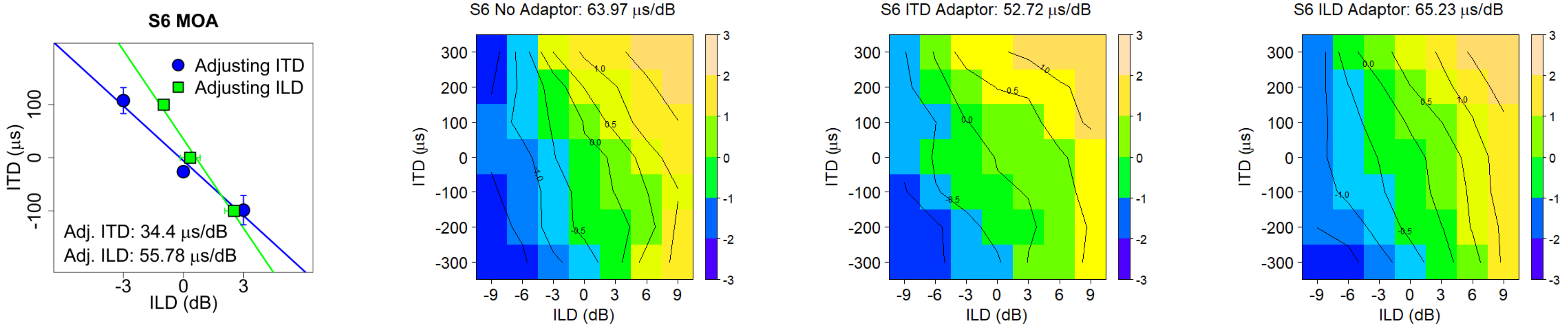

S1 MOA
Adjusting ITD
Adjusting ILD
Adj. ITD: 26.71 μs/dB
Adj. ILD: 30.09 μs/dB
ITD (μs)
ILD (dB)
S1 No Adaptor: 45.42 μs/dB
S1 ITD Adaptor: 40.11 μs/dB
S1 ILD Adaptor: 47.95 μs/dB
-9 -6 -3 0 3 6 9
-300 -200 -100 0 100 200 300
3 2 1 0 -1 -2 -3

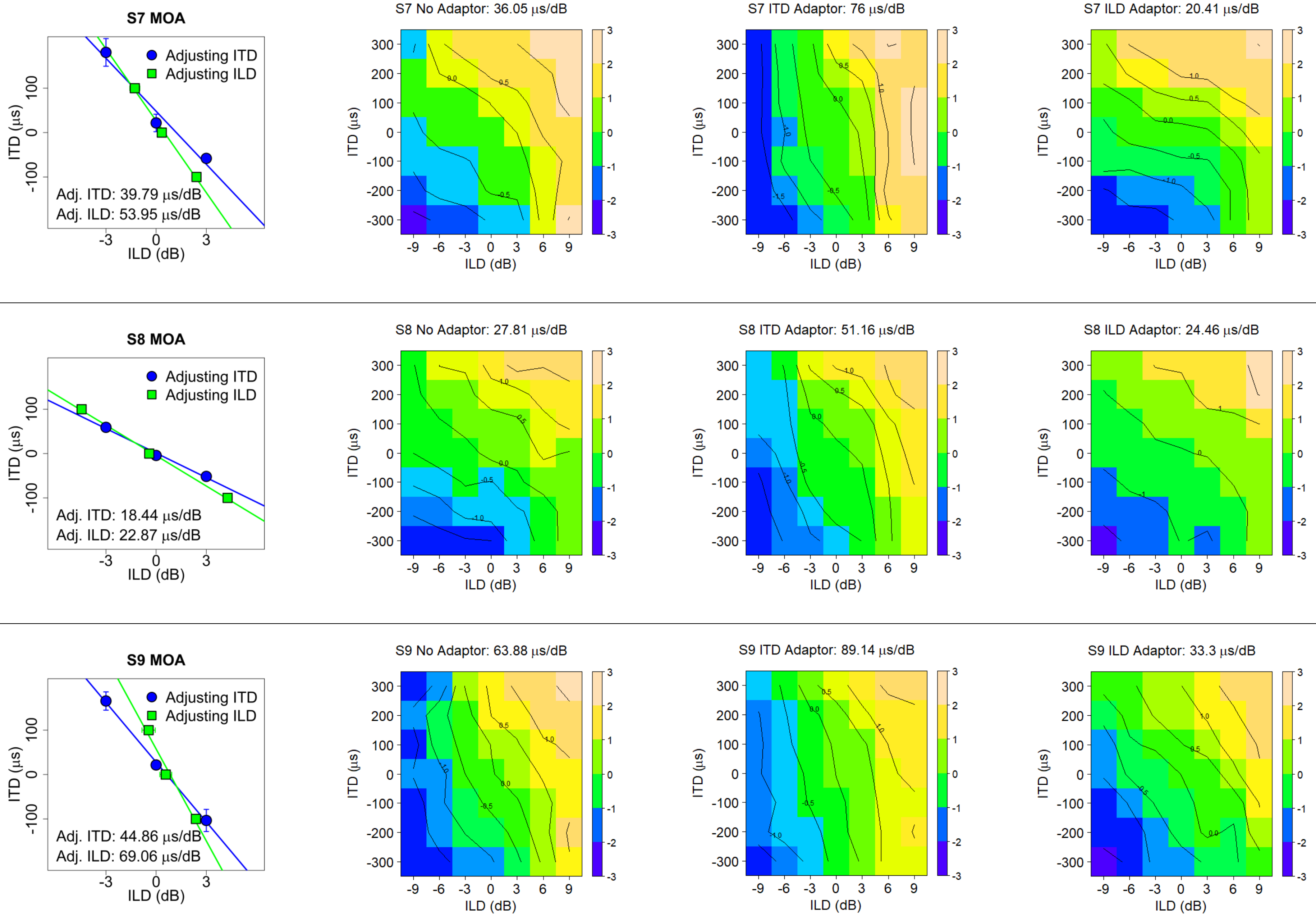

S2 MOA
Adjusting ITD
Adjusting ILD
Adj. ITD: 14.96 μs/dB
Adj. ILD: 31.07 μs/dB
ITD (μs)
ILD (dB)
S2 No Adaptor: 38.73 μs/dB
S2 ITD Adaptor: 50.34 μs/dB
S2 ILD Adaptor: 27.02 μs/dB
-9 -6 -3 0 3 6 9
-300 -200 -100 0 100 200 300
3 2 1 0 -1 -2 -3


S3 MOA
Adjusting ITD
Adjusting ILD
Adj. ITD: 21.93 μs/dB
Adj. ILD: 24.38 μs/dB
ITD (μs)
ILD (dB)
S3 No Adaptor: 28.6 μs/dB
S3 ITD Adaptor: 40.61 μs/dB
S3 ILD Adaptor: 35.22 μs/dB
-9 -6 -3 0 3 6 9
-300 -200 -100 0 100 200 300
3 2 1 0 -1 -2 -3


S4 MOA
Adjusting ITD
Adjusting ILD
Adj. ITD: 18.85 μs/dB
Adj. ILD: 19.88 μs/dB
ITD (μs)
ILD (dB)
S4 No Adaptor: 20.24 μs/dB
S4 ITD Adaptor: 26.72 μs/dB
S4 ILD Adaptor: 17.37 μs/dB
ITD (μs)
ILD (dB)


S5 MOA
Adjusting ITD
Adjusting ILD
Adj. ITD: 31.22 μs/dB
Adj. ILD: 43.8 μs/dB
ITD (μs)
ILD (dB)
S5 No Adaptor: 42.14 μs/dB
S5 ITD Adaptor: 37.39 μs/dB
S5 ILD Adaptor: 42.98 μs/dB
ITD (μs)
ILD (dB)


S6 MOA
Adjusting ITD
Adjusting ILD
Adj. ITD: 34.4 μs/dB
Adj. ILD: 55.78 μs/dB
ITD (μs)
ILD (dB)
S6 No Adaptor: 63.97 μs/dB
S6 ITD Adaptor: 52.72 μs/dB
S6 ILD Adaptor: 65.23 μs/dB
ITD (μs)
ILD (dB)


S7 MOA

S7 No Adaptor: 36.05 µs/dB

S7 ITD Adaptor: 76 µs/dB

S7 ILD Adaptor: 20.41 µs/dB

S8 MOA

S8 No Adaptor: 27.81 µs/dB

S8 ITD Adaptor: 51.16 µs/dB

S8 ILD Adaptor: 24.46 µs/dB

S9 MOA

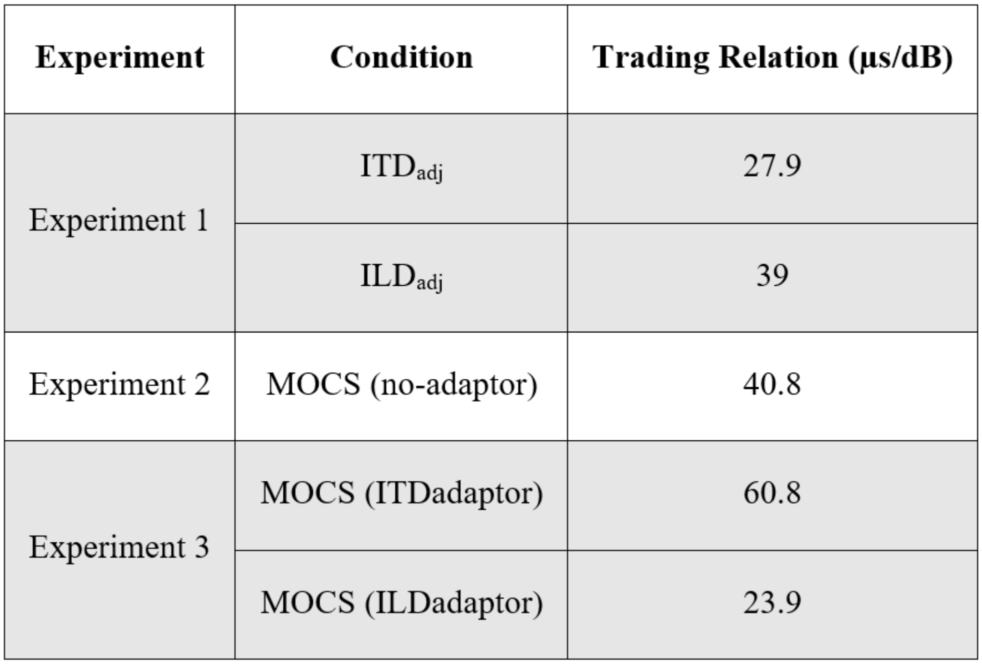
S9 No Adaptor: 63.88 µs/dB

S9 ITD Adaptor: 89.14 µs/dB

S9 ILD Adaptor: 33.3 µs/dB

Figure 19. Individual TRs across all experiments and conditions. Each row displays data for a different listener. From left to right, the panels show TRs from (1) the MOA task; (2) the MOCS task; (3) MOCS ITDadaptor condition, and (4) MOCS ILDadaptor condition.

Table 2. Mean TRs for every experiment and condition. Values given for Experiment 3 are based on the Fixed group only.



Experiment	Condition	Trading Relation (µs/dB)
Experiment 1	ITD_{adj}	27.9
	ILD_{adj}	39
Experiment 2	MOCS (no-adaptor)	40.8
Experiment 3	MOCS (ITDadaptor)	60.8
	MOCS (ILDadaptor)	23.9

A series of planned, paired-sample *t*-tests were conducted between the MOCS TR from Experiment 2, and the two adaptor conditions from Experiment 3. Due to the lack of effect of the adaptors for the Mixed group (see Chapter 5), these analyses were restricted to participants in the Fixed group from Experiment 3 (i.e., S4, S7, S8, S9). Comparison of the no-adaptor MOCS TR and the ITDadaptor TR revealed a significant difference between the conditions ($t(3) = -3.47$, $p < 0.05$, $d = 1.73$). Comparison of the MOCS TR and the ILDadaptor TR showed the values were not statistically different ($t(3) = -2$, $p > 0.05$, $d = 1$); however, the large effect size suggests a larger sample would likely yield a significant finding. Supporting this notion, inspection of Figure 20 shows a clear visual trend for the MOCS and ILDadaptor slopes to differ in the expected direction had adaptation occurred. Reducing the sample size to only those in the Fixed group (i.e., 4 participants) is very likely the reason for lack of statistical significance. These findings confirm that binaural spatial adaptation is indeed possible using the parameters commonly employed in MOA tasks.

It is worth mentioning that the Mixed group produced ITDadaptor and ILDadaptor TRs in Experiment 3 that were similar to the MOCS TR in Experiment 2 (44.2 μs/dB, 43.7 μs/dB, and 40.8 μs/dB, respectively). These results reveal similarity in TRs not only across adaptor types, but also between the adaptor and no-adaptor conditions for the Mixed group. This is strong evidence to suggest adaptation did not occur in the Mixed group.

The stimulus presentation pattern specific to each group provides some insight into the differential effectiveness of adaptation seen here. The Mixed group was presented blocks of either adaptor type during a single experimental session, and there was no Gaussian noise between trials, leading to three possible explanations for the lack of an effect. First, it may be that trial-by-trial adaptation is insufficient to influence perceived azimuth consistently, but that behaviorally relevant adaptation requires extended exposure to a single adaptor type. Second, the Gaussian noise could have effectively segregated stimulus presentations to a greater extent than the no-noise stimuli, potentially rendering trial-by-trial adaptation more potent. Third, a combination of prolonged exposure to a single cue, as well as intertrial noise might be required to see effects. Unfortunately, the current study design cannot distinguish between these possibilities.

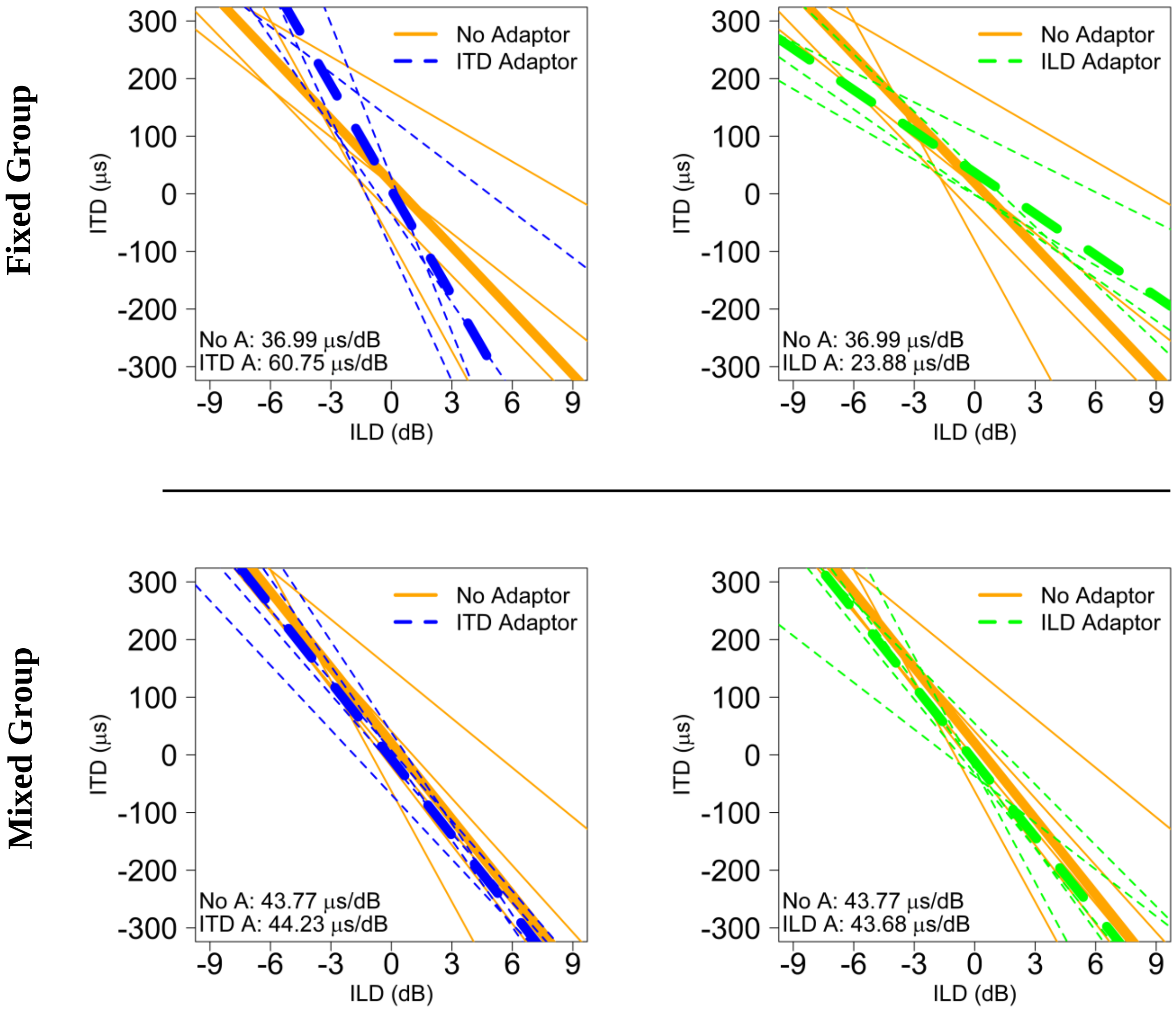


Figure 20. Mean and individual slopes for each adaptor type (thick and thin lines, respectively). The top two panels show data from the Fixed group, and the bottom two panels show data from the Mixed group. Slopes are plotted on the same scale as the heatmaps, where they were derived from the contour lines at 0 (cue combinations at which listeners perceived the stimulus at midline).

6.1.2 Experiment 1 vs. Experiment 2

The next step in determining whether spatial adaptation influences cue trading during an MOA task is to compare the ITD_{adj} and ILD_{adj} TRs to the no-adaptor MOCS TR (i.e., Experiment 2). It was hypothesized that repeated presentations rendered the MOA TRs susceptible to spatial adaption, while the TR from the no-adaptor MOCS task was derived from single presentations. One potential outcome of this comparison is that if an adaptive process during the MOA results in artificially low TRs when adjusting the ITD and artificially high TRs when adjusting the ILD, the MOCS TR should fall between the MOA values. Another possibility is that the MOCS will agree with one of the TRs from the MOA. In the latter case, it is hypothesized the MOCS TR will be similar to the ILD_{adj}, consistent with the rationale the ITD_{adj} is influenced by adaptation.

Planned comparisons between the MOCS TR and the ITD_{adj} and ILD_{adj} TRs were carried out using bootstrapped paired-samples *t*-tests (10,000 replications). The comparison between the MOCS TR (40.8 µs/dB) and the ITD_{adj} TR (27.6 µs/dB) revealed a significant difference ($t(8) = 3.57$, 95% CI(−1.84, 1.87), $p < 0.01$, $d = 1.19$). The comparison between the MOCS TR and the ILD_{adj} (39.01 µs/dB) revealed no significant difference ($t(8) = 0.55$, 95% CI(−1.84, 1.9), *p* = n.s.). These results indicate TRs were similar when obtained using single presentations of cue combinations (MOCS) and when adjusting the ILD to offset fixed ITDs; however, the ITD_{adj} differed from both values. Thus the MOCS TR did not lie between the MOA TRs, but instead agreed with the ILD_{adj}. This relationship strongly implies that no adaptation occurred during the ILD_{adj} condition of the MOA, whereas adaptive effects did occur during the ITD_{adj} condition.

Two conclusions can be drawn from the findings thus far. The first is that the shallower slope of the ITD_{adj} could be due to an adapted ILD during adjustment. That is, repeated presentations of the ILD either (1) led to a reduction in sensitivity of the ILD, requiring less of an ITD to perceive the stimulus at midline, or (2) caused the ITD percept to shift away in the opposite direction, thereby requiring less of an ITD to center the auditory percept (e.g., as in the MOCS task). The second conclusion is that the ILD_{adj} appears unaffected by adaptation during the MOA task. This challenges the notion of cue-dependent TRs, in that only the ITD_{adj} appears to be affected by adaptation, while the ILD_{adj} is consistent across psychophysical methods, at least under the conditions used here.

6.1.3 Experiment 1 vs. Experiment 3

The final comparison concerns the ITD_{adj} and ILD_{adj} from Experiment 1, and the TRs from the adaptor conditions in Experiment 3. It bears repeating that the adaptor conditions from Experiment 3 can be somewhat counterintuitive. Because these conditions are referred to by the type of adaptor present, participants became more sensitive to the opposing (i.e., unadapted) cue. That is, for the ITDadaptor condition the ILD cue is more prominent, and for the ILDadaptor condition the ITD is dominant.

It is interesting to note the similarity between the ITD_{adj} and ILDadaptor conditions. The TRs from these conditions both favor the ITD, and do not differ statistically from each other ($t(3) = -1.18$, $p > 0.05$). This is important, because the ILDadaptor condition creates a scenario in which the ILD cue is adapted, rendering the ITD the dominant cue. In other words, TRs obtained when the ITD is the dominant cue are statistically similar, regardless of task. The similarity in TRs between ITD-dominant conditions arising from the MOA and the ILDadaptor tasks is direct evidence that suggests the comparatively smaller TR of the ITD_{adj} TR likely arises from adaptive effects.

The relationship between the ILD_{adj} and the ITDadaptor conditions is equally illuminating. These two conditions both favor the ILD, yet resulted in statistically different TRs ($t(3) = -4.28$, $p < 0.025$, $d = -2.14$). The ILD_{adj} is statistically similar to the no-adaptor condition, suggesting the ILD_{adj} TR was not susceptible to adaptation during adjustment. It is important to note, however, that overtly adapting the ITD produced a significantly larger TR compared with other ILD conditions. Thus there is evidence for the possibility of ILD adaptation, and evidence that this adaption did not occur during the ILD_{adj} condition of the MOA.

The ITDadaptor TR is also more consistent with the TR reported in Stecker (2010), who abolished the ITD and measured TRs with the dominant ILD (though the use of click trains could also account for the large TR). At this point, it can be said with some confidence that adapting the ITD can lead to increased perceptual sensitivity to the ILD, but this increase was not seen during the MOA task. These findings strongly imply the ITD_{adj} was biased by an adaptive process during the MOA, while the ILD_{adj} was not influenced, accounting for the "cue-dependent" TRs obtained using an MOA task.

6.2 Discussion

For the stimulus parameters used here, the results of these experiments indicate the ITD is affected by adaptation using the MOA, but the ILD is not. Specifically, the ILD becomes adapted only when it serves as the repeated cue (i.e., while adjusting the ITD), but is unaffected by adaptation when being actively adjusted. A reasonable assumption is that adaptation occurs during the repeated presentations, but the changing value of the ILD during adjustment provides no opportunity for adaptation. The result is a lower TR in the ITD_{adj} condition, due to either a decrease in the amount of ITD necessary to offset the perceptually weaker ILD, or a repulsive shift of the ITD percept away from the adapted ILD. Conversely, TRs derived from the ILD_{adj} condition reflect percepts based on the presence of both cues, as neither cue serves as an adaptor in this condition. Accordingly, the ILD_{adj} condition produces a TR similar to the no-adaptor MOCS task.

Extrapolating these findings to the dual images reported from pointing tasks using the MOA, it seems reasonable to conclude that the unadapted time image reflects the ability to perceive the ITD component of a mixed cue with training, and the time/intensity image produces TRs based on the presence of competing cues. Indeed, it was discussed in Experiment 1 that the time/intensity TRs from pointing tasks are consistent with those from the ILD_{adj} TRs from centering tasks. Both of these types of TRs are similar to the no-adaptor MOCS reported here.

From these observations, the following statements can be reasonably made: 1) the ILD_{adj} (centering), time/intensity image (pointing) and the no-adaptor MOCS (head-pointing) produce similar TRs when stimulus parameters are similar, and are based on cue interaction; 2) the time image (pointing) and ITD_{adj} (centering) produce TRs comparatively smaller than their task-respective counterparts; 3) the time image simply represents a listener's ability to identify the ITD when cues are in opposition; and 4) the ITD_{adj} TR is based on cue interaction, but is a product of spatial adaptation that overestimates the TR (i.e., produces a smaller TR) for a given cue combination.

These conclusions successfully account for the cue-dependent nature of TRs reported in the literature across a variety of methodologies. This study argues that binaural spatial adaptation should be considered as a major contributor to differences in TRs obtained using an MOA task. The next chapter (Chapter 7) considers the broader impact of the results of the experiments presented here, both in light of existing literature, and in terms of specific adaptive mechanisms. Study limitations and directions for future work are also discussed.

Chapter 7

GENERAL DISCUSSION

7.1 Adaptation, attention and regression

As discussed in the Introduction, Lang and Buchner (2009) recorded the final ITD and ILD cue values chosen by participants to center a 500 Hz pure tone using the MOA, and then played those same cue values as single presentations using a lateralization MOCS task. The results revealed that perceptions were no longer centered when MOA values were presented during the MOCS task. Instead, the perceived azimuth deviated away from midline, moving toward the perceived location of the fixed cue during adjustment. That is, the adjusted cue value was no longer sufficient to offset the complementary cue to midline, resulting in perceptions “shifted back” toward the fixed cue. Lang and Buchner (2009) explained this phenomenon as attentional upweighting of the cue being adjusted; increased perceptual salience of the adjusted cue rendered it more effective than the fixed cue.

The findings of the present study run counter to the idea of attention increasing the effectiveness of the adjusted cue; in fact, this study argues the opposite is true. The results presented here support the notion that (1) artificial salience of the adjusted cue is due to decreased salience from the fixed cue, or (2) the fixed cue becomes an adaptor through repeated presentations and causes perception of the adjusted cue to be pushed away from the adaptor. In essence, the attentional upweighting theory focuses on the adjusted cue and suggests an increase in its effectiveness, whereas the adaptation theory proposed here focuses on the fixed cue and argues its effectiveness is either decreased, or the fixed cue repulses the percept of the adjusted cue away.

While the present study was not designed to state definitively whether adaptation, attention (Lang & Buchner, 2009), or regression (Trahiotis & Kappauf, 1978) accounts for cue-dependent TRs, it is noteworthy that binaural spatial adaptation can account for both theories. Figure 21 shows data from the current study plotted according to the convention of Lang and Buchner (2008, 2009): as deviation in perceived azimuth from midline between MOA and MOCS tasks. Deviations in Figure 21 occur in the same direction as the Lang and Buchner investigations, reproducing their findings based on data clearly derived from adaptive effects. Specifically, deviations from perceived midline appear to "shift back" toward the value of the fixed cue. In general, when a fixed cue favored the left during adjustment, the MOCS task revealed a shift in perceived azimuth to the left, and vice versa. The very use of the term "shift" in this context suggests adaptation, as described in a number of adaptive localization aftereffect studies (e.g., Canévet & Meunier, 1996; Kashino & Nishida, 1998; Meunier et al., 1996; Thurlow & Jack, 1973). Moreover, Lang and Buchner (2008) separated stimuli during their MOA task with 500 ms of silence, which is within the range of adaptive spatial effects.

Not only does adaptation account for the "shift-back" effect attributed to attentional upweighting, it also reconciles heretofore unexplained findings relating attentional upweighting and the regression theory. In an extension of the work of Lang and Buchner described above, Ignaz et al. (2014) found that the shift-back effect increased in the presence of a reference tone. The authors attributed this phenomenon to the tendency of the adjusted cue to move toward the value of the same cue in the reference (i.e., the regression theory). Ignaz et al. (2014) postulated that both attentional upweighting and regression contributed to their results, though the details of such an interaction were uncertain. Spatial adaptation, however, offers a simpler solution to greater deviations in perceived azimuth in the presence of an adaptor. As described throughout this document, it has been well established that perceived azimuth shifts away from the location of a binaural adaptor. A repeated reference tone at midline (used in Ignaz et al., 2014) itself becomes an adaptor, causing a greater shift of the adjusted cue away from midline than when the reference tone is absent. This scenario requires even less of an ITD to center the auditory image than when the fixed cue in the opposite hemifield acts as the repulsing adaptor, and therefore leads to a greater shift-back effect as well.

The role of attention in the cue trading literature seems better suited to explaining the time image from pointing tasks. In these experiments, participants are overtly instructed to attend to one image or the other. Historically, this has been an extremely difficult task that only some listeners can perform after significant practice (e.g., Hafter & Jeffress, 1968). Parsing the stimulus for the presence of the ITD therefore requires practice and focused attention. In contrast, statistically differing TRs are easily obtained using the MOA, which might indicate the involvement of an automatic perceptual phenomenon that does not require attention to achieve. It is, however, important to reiterate that the current study design cannot exclude attentional upweighting as a possible influence.

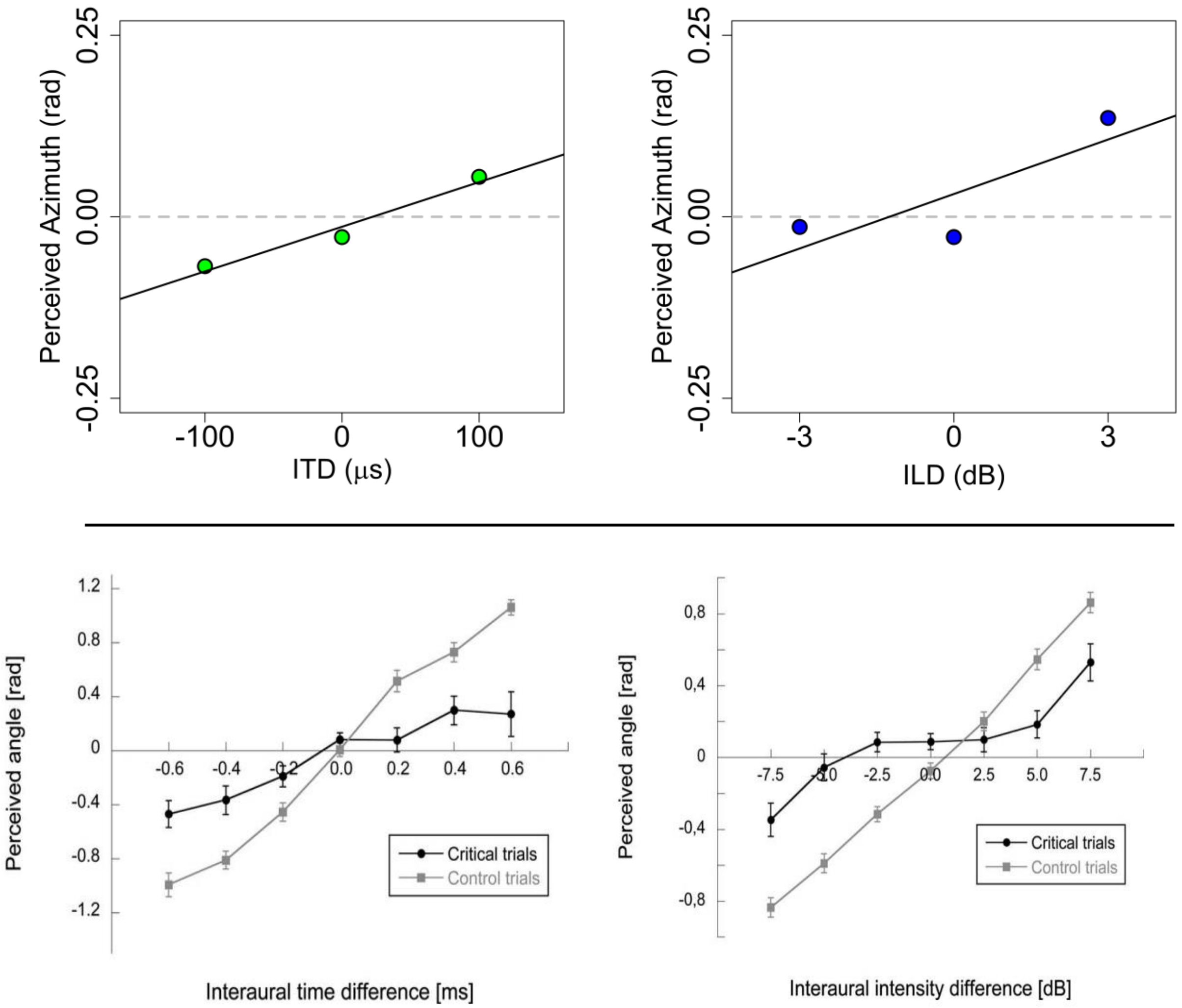


Figure 21. Data plotted to show the “shift-back” effect. Top panels: Data from the current study plotted following the convention of Lang and Buchner (2009) to show the “shift-back” effect. Top left: each point represents the perceived azimuth (ordinate) of ITD and ILD cue combinations from the MOCS that most closely matched the values obtained during the MOA task to center the percept. Deviations from the grey dotted line indicate leftward (negative values) or rightward (positive values) deviation from perception at midline. Top right: same as top left, but for fixed values of ILD. Bottom panels: corresponding plots from Lang and Buchner (2009). The black lines are the points of comparison with the top plots. From “Relative influence of interaural time and intensity differences on lateralization is modulated by attention to one or the other cue: 500-Hz sine tones,” by A. Lang and A. Buchner, 2009, *The Journal of the Acoustical Society of America*, *33*, pg. 2536–2542. Copyright 2009 by AIP Publishing. Reprinted with permission.

7.2 Implications

7.2.1 Past literature

The finding that the $\mathrm{ILD}_{\mathrm{adj}}$ TR is consistent across MOA and no-adaptor MOCS tasks affects the way in which past studies are interpreted. This finding goes against the conventional belief that differing TRs from an adjustment task are biased by each cue. Instead, it appears the $\mathrm{ITD}_{\mathrm{adj}}$ condition is the only scenario in which the TR is influenced by processes other than perception of the stimulus under certain conditions. The implication is that, insofar as MOCS tasks produce TRs more directly related to the interaction of the cues, the $\mathrm{ILD}_{\mathrm{adj}}$ TRs should produce TRs of equal “accuracy.” Unfortunately, the early literature investigating cue trading only measured TRs when the ITD was adjusted to offset a fixed ILD, which may have yielded TRs more sensitive to the stimulus parameters than would otherwise have been the case. Had past work explored the $\mathrm{ILD}_{\mathrm{adj}}$, there may have been more evidence to show consistent TRs across methodologies using the parameters so often employed in those studies. Early TRs likely overestimate (smaller TRs imply a more dominant ITD) the actual nature of perceptions caused by the opposition of binaural spatial cues, at least for 500 Hz pure tones. Going forward, it seems as though the general advice against using the MOA to obtain TRs (e.g., Young Jr & Levine, 1977) should be modified to apply only to the $\mathrm{ITD}_{\mathrm{adj}}$ condition when using parameters similar to those used here. It seems the $\mathrm{ILD}_{\mathrm{adj}}$ TR reflects contributions from both the ITD and ILD, and therefore can be considered a valid representation of cue trading. Furthermore, it seems reasonable to argue that the more intuitive head-pointing procedure, which yields TRs similar to the $\mathrm{ILD}_{\mathrm{adj}}$ condition, is likely the best method for quantifying cue interaction.

7.2.2 Relative effectiveness of adaptation

One of the most interesting findings of the present study is the relative sensitivity of each cue to adaptation. It was shown that while the ILD_{adj} TR was unaffected by adaptive effects, the ITD_{adj} condition produced a lower (more dominant) TR. In order to understand this phenomenon, it is useful to recall the parameters used in this study, and the known effects those parameters have on binaural cues. As discussed in Experiment 1, there are three main points to consider. First, the ITD is known to drive perception in wideband stimuli (e.g., Macpherson & Middlebrooks, 2002), giving rise to the general rule that the ITD is the dominant cue. Second, low frequency ILDs are contrived stimuli that exist only under headphones, and therefore may be more easily ignored by the auditory system. Third, Harris (1960) found that larger values of ILD created diffuse percepts that proved difficult to lateralize.

Taken together, a reasonable conclusion is that, under the conditions used in this study, the ITD was the more coherent cue, while the ILD was more diffuse and less effective. Therefore, in the ILD_{adj} condition, the ITD either remained dominant despite repeated presentations, or the ILD was too incoherent to shift away from the fixed ITD. Conversely, in the ITD_{adj} condition, the fixed ILD was either successfully weakened by adaptation, or the ITD was sufficiently coherent to shift away from the ILD adaptor. This logic is consistent with the ILDadaptor condition producing a TR similar to the ITD_{adj} condition (i.e., the ITD was dominant in both conditions). It is noteworthy that the only ILD-dominated TR occurred when the ITD was overtly adapted with a train of pure tones, suggesting the adaptive effects arising from the MOA were insufficient to adapt the ITD enough to allow the ILD to dominate.

This differential effectiveness of binaural cues depending on the stimulus parameters used is consistent with the wide range of TRs observed across tasks, stimulus type, level, and frequency. Thus, it is possible spatial adaptation may be highly context-sensitive, and depend to a great extent on the relative strength of the stimulus parameters of a given task.

7.2.3 Processing of binaural spatial cues

Psychophysical considerations

The specific nature of how ITDs and ILDs are represented in the auditory system remains unknown. Older models explained binaural percepts of laterality as originating solely from time or level differences via conversion from one cue to another (e.g., Deatherage & Hirsh, 1959; van Bergeijk, 1962). For example, level differences could be converted to time differences because more intense stimuli produce shorter latency firing of the auditory nerve. Conversely, time could be converted to a level difference because variations in onset time produce temporary level differences at the onset slopes. However, more recent work has largely shown these peripheral conversion processes to account for only a small portion of the overall perception of laterality (e.g., Joris et al., 2008; Joris et al., 1998).

Phillips and Hall (2005) commented on peripheral and central mechanisms of binaural cue processing. In a novel paradigm, they used two adaptors, each carrying either exclusively ITDs or exclusively ILDs, at two different frequencies. Further, each frequency carried binaural cues that caused perceived laterality to occur on opposite sides of the head. After exposure to the adaptors, participants were asked to indicate whether their perceptions of a diotic stimulus was shifted to the right or left. Psychometric functions plotting the number of right-shifted percepts revealed that in both ITD and ILD experiments, perceived laterality was shifted in opposite directions for the two frequencies. Due to the complex nature of the stimuli, Phillips and Hall argued against any bias in responses at the cognitive level.

Building on these findings, Phillips et al. (2006) used a similar paradigm, but tested ITD adaptors on ILD targets, and vice versa. The results were similar to their previous experiment, with psychometric functions showing displacement of the auditory percepts away from the adaptors. Phillips et al. (2006) offer two interpretations of their data. First, they postulated the existence of a single, central "spatial processor" that receives convergent inputs from ITD and ILD cues. Within this framework, the cue type becomes irrelevant, because either (1) the output of the spatial processor would not contain information about the adapted cue, or (2) the output would be the result of the high-level processor itself being adapted. This possibility seems unlikely for several reasons. One reason is because, with practice, listeners are able to identify the time image in a pointing task, which seems to imply the ability to focus on a single component of the otherwise cohesive auditory percept at will. Another reason is that cue trading is incomplete. In a model where cue types are irrelevant, inputs converging onto a single spatial processor would arguably trade completely, in a linear fashion.

Second, Phillips et al. (2006) speculated on the possibility of independent processing of ITDs and ILDs. For instance, they cite physiological evidence demonstrating the neural representations of time and level differences only partially overlap (e.g., Schröger, 1996), and that ILDs can alter the processing of ITDs through a subset of neurons in the inferior colliculus of the cat (e.g., Kuwada & Yin, 1983). (The discussion becomes circular when one asks whether the collicular modulation is inherent to the neurons, or achieved via converging inputs.) Phillips et al. (2006) and the current study show that binaural spatial cues interact, inasmuch as one cue can adapt the other. This strongly suggests at least enough independence for ITDs and ILDs to be coded as separate but related perceptual features, albeit perhaps within a shared, more general spatial processing channel. In other words, the spatial information carried by each cue provides similar information, but is not completely redundant.

Neuroanatomical considerations

As a point of comparison, it is worth briefly considering the neural processing of color vision. The visual system uses parallel color-opponent channels to carry information to the cortex, where is it integrated at a later stage (Chatterjee & Callaway, 2003). Similarly, the auditory system projects spatial information carried by ITDs and ILDs to the cortex for further interpretation (Yin, 2002). Primary auditory cortex contains patches of binaurally-selective neurons (Imig & Adrian, 1977; McLaughlin et al., 2016; Middlebrooks & Zook, 1983; Stecker et al., 2005), just as color might be processed using a similar distributed system of patch activation (e.g., Bird et al., 2014; Conway, 2009; Li et al., 2014). That is, in lieu of a topographical map, color and auditory space are both calculated quantities that arise from multiple perceptual channels.

Despite the similarities, there is an important difference in how color and auditory space are initially encoded. The visual system perceptual channels remain primarily segregated through the brainstem and thalamus. For instance, the lateral geniculate nucleus (LGN) of the visual system is composed of three classes of cell layers, each optimized for processing specific stimulus features from retinal inputs, and passes these segregated streams, including color, to primary visual cortex (for review see Kaufman et al., 2011).

In contrast, the binaural integration required to code auditory space leads to significant crossing of the left/right and ITD/ILD streams in the auditory brainstem. For example, the inferior colliculus (IC) receives excitatory and inhibitory inputs from diverse brainstem locations, such as the dorsal and ventral nuclei of the lateral lemniscus, the medial and lateral superior olives (MSO and LSO, respectively) and the cochlear nuclei (for review see Pickles, 2012). Loftus et al. (2004) used anterograde tracers in the cat to map the projections from the MSO (largely ITD sensitive) and LSO (largely ILD sensitive) to their terminations in the central nucleus of the IC (ICC). They found MSO inputs were segregated from contralateral LSO inputs, suggesting ITDs and ILDs are processed independently; however, they also showed that ipsilateral LSO projections, largely inhibitory in nature, converged with the MSO inputs, suggesting cue integration. The overall conclusion from Loftus et al. (2004) was that several neural circuits likely exist in the ICC that make use of different input combinations. This information is then projected from the ICC and terminates in patches in the medial geniculate body (MGB). Patches in the MGB are maintained at the level of primary auditory cortex and make up the binaurally sensitive patches mentioned above (Velenovsky et al., 2003).

The brief comparison between visual and auditory parallel processing seems to confirm the aptness of the term *color-opponency*, describing independent perceptual channels that are preserved through the visual thalamus. The comparison also reveals the processing of ITD and ILD is not as straightforward. Animal work has shown there is sufficient circuity in the integrative nuclei of the ICC and MGB to combine ITD and ILD information, yet cortical patches not only show general binaural sensitivity, but can also respond preferentially to cue type (Higgins et al., 2017).

In piecing together the psychophysical and neurophysiological findings, it is interesting to note that the perception of dual images described in early cue trading studies is a difficult task that does not resemble the phenomenon of auditory bistability: the spontaneous transition between different auditory percepts when presented with ambiguous stimuli (Hupé et al., 2008). Bistable stimuli capitalize on grouping principles (for review see Bregman, 1994), allowing for multiple perceptions of a constant stimulus. The lack of bistability in perceiving the time and time/intensity images suggests ITD and ILD cues are not auditory “objects” in themselves, and there is likely no local neuronal competition for perception (arguably what drives bistability; e.g., Hupé et al., 2008). The lack of bistable percepts suggests the relative coherence of cue-traded percepts stems from an integration of ITD and ILD information, which agrees with the central spatial processor proposed by Phillips et al. (2006). In contrast, the results of the present study show that perceived laterality is context-dependent, and that adaptation of a single cue type is possible. These findings imply independence of ITD and ILD processing, at least at early stages. Taken together, a simple model for binaural spatial processing emerges: (1) ITD and ILD cues are treated independently at early stages of processing, (2) the cues are later combined into a single percept at the cognitive level, and (3) cue-selectivity in the cortex suggests not all cue-specific information is lost during integration.

7.2.4 Technology

Assistive devices

Previous work has shown that sound localization performance can become degraded with the use of hearing aids (for review see Diedesch, 2016). More specifically, factors such as microphone placement, disruption of pinna cues, and dynamic range compression haven been shown to contribute to difficulty localizing (e.g., Diedesch, 2016; Diedesch et al., 2018; Hassager et al., 2017). Hassager et al. (2017) reported hearing aid processing can lead to auditory images that are broad, intracranial or split, which falls within the purview of the current study. Localization performance may be improved in these instances by carefully examining the cue interactions caused by hearing aid processing. It is possible that knowledge of the most disruptive cue combinations could inform hearing aid processing algorithms.

Furthermore, developing consistent, objective techniques to quantify binaural performance can impact diagnostic testing and equipment in the audiology clinic. Amassing normative data on binaural spatial performance can lead to a better understanding of how auditory spatial cues interact with hearing aids, and the types of real-world benefits spatial hearing can offer. In the hearing science laboratory, knowledge of binaural spatial cues could become a routine part of testing hearing aid features, similar to work that is currently done with directional microphones, speech intelligibility and listening effort (e.g., Picou et al., 2017; Simpson et al., 2018). TRs also provide a convenient way to quantify spatial hearing results from such tests, allowing comparison across hearing aid features, manufacturers, and listening conditions. For example, bilateral cochlear implant patients might perform a localization task that prompts the listener to orient toward a lateral cue, achieved by manipulating the ILD. Subsequently, various amounts of ITD can be introduced to the fixed ILD signal to determine the effectiveness (or ineffectiveness) of a particular processing strategy of incorporating ITD cues.

Virtual reality

An important aspect of this study was that all experiments were conducted in VR. The reliability of participant responses suggests the use of VR did not negatively affect the quality of the data. Furthermore, TRs obtained in this study were comparable to those obtained using a variety of technologies, ranging from analog circuits (e.g., Deatherage & Hirsh, 1959) to touch-screen tablets (Stecker, 2010).

Establishing the use of VR as a legitimate tool for psychophysical studies is useful for several reasons. For instance, VR allowed for the natural and intuitive response technique of simply orienting the head to a stimulus, and pulling a trigger to “shoot” balloons. Spontaneous participant comments revealed the more interactive tasks (i.e., MOCS) were noticeably more enjoyable than the less interactive tasks (i.e., MOA). These comments imply interactive virtual environments may lead to more engaged listeners, potentially delaying the effects of mental fatigue and reducing the overall number of visits required to complete data collection. VR tasks can also be shared easily. Virtual scenes can be saved as stand-alone files, and require only inexpensive and commercially available equipment (e.g., Oculus Rift) to implement. Not only does VR remove visual distractions from the laboratory surroundings, but virtual environments provide excellent consistency for comparing results across different laboratories.

Virtual reality is already being used by physical therapists in a clinical setting (for review see Sveistrup, 2004). Improved methods of customizing binaural spatial information on an individual basis could have a major impact on the realism and effectiveness of virtual environments. More realistic environments could be especially useful in training individuals who rely more heavily on auditory spatial awareness, such as those with visual and vestibular impairments. The ready availability and affordability of VR equipment could even allow patients to train at home. An early step toward such technology was introduced by Sechler et al. (2017), who created a portable VR system for objective sound localization testing for cochlear implant users.

The results of this study can also contribute to the virtual and augmented reality industries. Greater understanding of the individual differences in sensitivity to binaural cues will allow for finer control over spatial audio and more effective calibration techniques. For example, more people may be willing to take the time to calibrate their virtual audio if the task is made into a game similar to that used here. Better understanding of cue trading can lead to the capture of information about ITD and ILD effectiveness simultaneously, reducing the duration of the calibration process. Libraries of established TRs could be used in personalized calibration of the generic head-related transfer function (HRTF) that is commonly used in virtual audio software. That is, automatic software adjustments to the HRTF, based on the results of a brief cue-trading task, have the potential to provide more individualized information on the relative weighting of binaural cues under different conditions, thereby increasing realism in the virtual environment.

7.3 Study limitations

7.3.1 Fixed group sample size

The major drawback of splitting the participants into two groups was the subsequent reduction in sample size. Data from only 4 listeners resulted in a lack of statistical significance between the no-adaptor and ILDadaptor MOCS conditions. While the effect size was quite large and the data exhibited a trend toward significance, the findings of this study would be strengthened by a larger Fixed group sample size.

7.3.2 Restricted range of fixed ILD values

Another limitation was the failure to obtain responses at all levels of the fixed ILD during the MOA task. Part of the problem may have been due to the fact this study used cue values based on a centering task (i.e., Whitworth & Jeffress, 1961), rather than a pointing task. In retrospect, aligning the fixed ILDs in this study with those commonly used in centering tasks would likely have increased the range of useable ILDs. The fact remains, however, that even at ±6 dB participants consistently struggled to center the percept. Other studies using a centering task and the MOA have reported difficulty around this ILD value, but have typically been able to salvage sufficient data for reporting (e.g., Lang & Buchner, 2009; Moushegian & Jeffress, 1959).

Another potential cause for the truncated range of fixed ILDs may be the number of trials presented. The present study required 8 ILD and 8 ITD adjustments for each of either 14 conditions (first five listeners) or 12 conditions (last four listeners). The total number of judgments for the current experiment was therefore either [(8_{ITD} + 8_{ILD}) * $14_{\text{Condition}}$)] 224 judgments, or [(8_{ITD} + 8_{ILD}) * $12_{\text{Condition}}$] 192 judgments, respectively. In comparison, Whitworth and Jeffress (1961) collected two sets of 8 ILD judgments and 4 ITD judgments for each of 49 conditions, resulting in a total number of 590 responses. Harris (1960) reported a comparable 560 responses, while Deatherage and Hirsh (1959) reported a total of only 165 judgments per listener. More similar to the current study, Stecker (2010) obtained at least 240 responses per participant. Considering the literature in general, the current study falls into the lower end of the number of trials presented. It is possible more trials and therefore more experience with the task in general would have yielded responses across a greater range of fixed ILDs. It should be noted, however, that all participants completed practice sessions, and that split-half reliability and standard error of the data were good.

7.3.3 Generalizability

An important limitation inherent to cue trading literature as a whole is the ability to generalize findings to other stimuli and study designs. It has been shown throughout this document that TRs are highly sensitive to a variety of parameters, such as presentation level (David Jr et al., 1959), task (Lang & Buchner, 2008, 2009), spectral content (Harris, 1960), and the use of reference tones (Ignaz et al., 2013, 2014), to name a few. Therefore, the results of the experiments reported here have relatively little generalizability outside the conditions used in this study. The larger concept of adaptive effects in cue trading can, however, be taken as a general rule, and used to make predictions for other tasks and stimulus parameters.

7.3.4 Ecological validity

The rather poor ecological validity of this study is also a limitation. For instance, the experiments were all conducted using insert earphones, rather than in the sound field; the stimuli contained low-frequency ILDs, which do not occur in everyday life; and the effects are dependent on a strict set of conditions. It should be noted, however, that the use of virtual reality increased realism in these experiments, and the response technique of orienting the head in a direction of a sound is intuitive and common in daily life. It should also be noted that experiments conducted in the sound field support the findings of a shift in the auditory percept using adaptors and probes (e.g., Kopčo et al., 2007), increasing the relevance of spatial adaptation to more realistic environments.

7.4 Future directions

7.4.1 Differentiating adaptation and attention

The current study provided strong evidence for the contribution of binaural spatial adaptation to the cue-dependent nature of TRs. However, the study was not designed to demonstrate causation, and so adaptation can only be offered as a likely mechanism. The most immediate next step is to discover which process or processes contribute to the differential TRs obtained using the MOA, and to quantify those relationships. A psychophysical experiment might include measuring TRs over frequencies that produce a range of ITD effectiveness, to determine whether TRs reflect changes in the usefulness of the ITD as an adaptor, or whether

TRs remain constant due to the influence of attention. Conversely, an electrophysiological experiment might present stimuli in a passive listening scenario, effectively removing attentional effects, to examine changes in the EEG signal for signs of adaptation during cue adjustment. Such an experiment could be implemented using the EEG-based binaural spatial adaptation paradigm used by Magezi and Krumbholz (2010).

Another important piece of information that needs to be resolved is differentiating between the two possible types of adaptation proposed here; that is, whether the fixed cue decreased in potency, or caused the adjusted cue to shift away.

7.4.2 Next steps

Kawashima and Sato (2012) investigated timing differences in amplitude modulated tones at high frequencies (above 2 kHz), and determined the envelope ITD also contributes to the localization aftereffect in the expected direction (i.e., the test signal shifted away from the adaptor signal). Future work should examine the role of the envelope ITD in cue-trading experiments to determine whether the same adaptive effects are involved in broader-spectrum signals. Similarly, additional work is needed to examine TRs at frequencies where the ILD dominates the ITD, to determine whether the ITD dominance observed in this study can be reversed using sinusoidal tones. Another factor that should be studied is whether the use of generic cue values significantly affects the TRs observed using the MOA. That is, would ITDs and ILDs that were perceptually matched for each listener (e.g., Yost et al., 1975) show more sensitivity to the ILD using the same parameters as this study?

Future studies also need to define what constitutes an “accurate” TR. For instance, this study showed the ITD_{adj} condition produced a smaller (more dominant) TR than the ILD_{adj}

condition, and that the ILD_{adj} TR was consistent with the no-adaptor MOCS task. Which TR is more accurate? Was it the ILD_{adj} TR, that was not influenced by adaptation, or was it the ITD_{adj} TR that correctly revealed the ITD was the dominant cue? Carefully controlled studies investigating TRs under a variety of circumstances are needed to produce a richer data set than what is currently available. Hafter and Carrier (1972) interpreted the lack of a complete trade between time and intensity as a refutation of the existence of a single TR. A larger set of TRs may prove that a constant TR does exist, but changes in a predictable way based on the context in which it is measured.

Chapter 8

CONCLUSIONS

The series of experiments presented here strongly suggest binaural spatial adaptation contributes to the cue-dependent TRs obtained using the MOA. In fact, spatial adaptation accounts for the variation in TRs observed across a variety of tasks and conditions reported in the literature. In general, task conditions that favor the ITD appear susceptible to adaptive effects, while conditions favoring the ILD are only susceptible when the ITD is overtly suppressed. These claims arise from considering the TRs across all three experiments. The major findings include: (1) introducing adaptors to an MOCS task leads to greater sensitivity to the unadapted cue; (2) the unadapted MOCS TR is similar to the $\mathrm{ILD}_{\mathrm{adj}}$ TR, suggesting adaptation does not occur when adjusting the ILD to offset a fixed ITD; (3) the MOCS ILDadaptor condition (ITD dominant) yields a TR similar to the $\mathrm{ITD}_{\mathrm{adj}}$ TR, suggesting the ITD is affected by adaptation during an MOA task; (4) the notion of cue-dependent TRs is too broad, as only the $\mathrm{ITD}_{\mathrm{adj}}$ differed from the no-adaptor MOCS TR under the conditions used here; and (5) at least at early stages, the auditory system may process binaural cues as distinct features of the same perceptual channel, allowing for interaction between cue types. While further work in the area of cue trading is needed, the results of this study offer new insights and promising directions for future research.

REFERENCES

Babkoff, H., Sutton, S., & Barris, M. (1973). Binaural interaction of transients: Interaural time and intensity asymmetry. *J Acoust Soc Am, 53*(4), 1028-1036.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods, 37*(3), 379-384.

Banister, H. (1926). Three experiments on the localization of tones. *British Journal of Psychology, 16*(4), 265-292.

Bartlett, F., & Mark, H. (1922). A note on local fatigue in the auditory system. *British Journal of Psychology, 13*(2), 215-218.

Bird, C. M., Berens, S. C., Horner, A. J., et al. (2014). Categorical encoding of color in the brain. *Proc Natl Acad Sci, 111*(12), 4590-4595.

Blauert, J. (1997). Spatial Hearing (revised edition). *Cambridge, MA: Massachusetts Institute of Technology*.

Braasch, J., & Hartung, K. (2002). Localization in the presence of a distracter and reverberation in the frontal horizontal plane. I. Psychoacoustical data. *Acta Acust United Acust, 88*(6), 942-955.

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*: MIT press.

Butler, R., & Naunton, R. (1962). Some effects of unilateral auditory masking upon the localization of sound in space. *J Acoust Soc Am, 34*(8), 1100-1107.

Canévet, G., & Meunier, S. (1994). Auditory adaptation and localization. *Acustica, 80*(3), 311-314.

Canévet, G., & Meunier, S. (1996). Effect of adaptation on auditory localization and lateralization. *Acta Acust United Acust, 82*(1), 149-157.

Carlile, S., Hyams, S., & Delaney, S. (2001). Systematic distortions of auditory space perception following prolonged exposure to broadband noise. *J Acoust Soc Am, 110*(1), 416-424.

Chatterjee, S., & Callaway, E. M. (2003). Parallel colour-opponent pathways to primary visual cortex. *Nat, 426*(6967), 668-671.

Conway, B. R. (2009). Color vision, cones, and color-coding in the cortex. *Neuroscientist, 15*(3), 274-290.

David Jr, E., Guttman, N., & Van Bergeijk, W. (1959). Binaural Interaction of High‐Frequency Complex Stimuli. *J Acoust Soc Am, 31*(6), 774-782.

Deatherage, B. H., & Hirsh, I. J. (1959). Auditory localization of clicks. *J Acoust Soc Am, 31*(4), 486-492.

Diedesch, A. C. (2016). *Binaural-cue weighting in sound localization with open-fit hearing aids and in simulated reverberation.* Vanderbilt University.

Diedesch, A. C., Stecker, G. C., & Gallun, F. J. (2018). Evaluating sound localization cues for premium hearing aids across multiple manufacturers. *J Acoust Soc Am, 143*(3), 1939-1939.

Domnitz, R., & Colburn, H. (1977). Lateral position and interaural discrimination. *J Acoust Soc Am, 61*(6), 1586-1598.

Feddersen, W., Sandel, T., Teas, D., et al. (1957). Localization of high‐frequency tones. *J Acoust Soc Am, 29*(9), 988-991.

Flügel, J. (1920). On local fatigue in the auditory system. *British Journal of Psychology, 11*(1), 105-134.

Freyman, R. L., Zurek, P. M., Balakrishnan, U., et al. (1997). Onset dominance in lateralization. *J Acoust Soc Am, 101*(3), 1649-1659.

Gaik, W. (1993). Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling. *J Acoust Soc Am, 94*(1), 98-110.

Gilkey, R. H., Good, M. D., Ericson, M. A., et al. (1995). A pointing technique for rapidly collecting localization responses in auditory research. *Behavior Research Methods, 27*(1), 1-11.

Hafter, E., & Dye Jr, R. (1983). Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number. *J Acoust Soc Am, 73*(2), 644-651.

Hafter, E. R., & Carrier, S. C. (1972). Binaural Interaction in Low-Frequency Stimuli: The Inability to Trade Time and Intensity Completely. *J Acoust Soc Am, 51*(6B), 1852-1862.

Hafter, E. R., & Jeffress, L. A. (1968). Two-Image Lateralization of Tones and Clicks. *J Acoust Soc Am, 44*(2), 563-569.

Harris, G. G. (1960). Binaural interactions of impulsive stimuli and pure tones. *J Acoust Soc Am, 32*(6), 685-692.

Hassager, H. G., May, T., Wiinberg, A., et al. (2017). Preserving spatial perception in rooms using direct-sound driven dynamic range compression. *J Acoust Soc Am, 141*(6), 4556-4566.

Higgins, N. C., McLaughlin, S. A., Rinne, T., et al. (2017). Evidence for cue-independent spatial representation in the human auditory cortex during active listening. *Proc Natl Acad Sci, 114*(36), E7602-E7611.

Hofman, P. M., Van Riswick, J., & Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nat Neurosci, 1*(5), 417-421.

Hupé, J.-M., Joffo, L.-M., & Pressnitzer, D. (2008). Bistability for audiovisual stimuli: perceptual decision is modality specific. *J Vis, 8*(7), 1-1.

Ignaz, A., Lang, A.-G., & Buchner, A. (2013). The impact of practice on the adjustment of interaural cues in a lateralization task. *J Acoust Soc Am, 134*(2), 901-904.

Ignaz, A., Lang, A.-G., & Buchner, A. (2014). The impact of reference tones on the adjustment of interaural cues. *J Acoust Soc Am, 135*(4), 1986-1992.

Imig, T., & Adrian, H. (1977). Binaural columns in the primary field (A1) of cat auditory cortex. *Brain Res, 138*(2), 241-257.

Jeffress, L. A., & Taylor, R. W. (1961). Lateralization vs localization. *J Acoust Soc Am, 33*(4), 482-483.

Jones, F., & Bunting, E. (1949). Displacement after-effect in auditory localization. *Amer. Psychologist, 4,* 389.

Joris, P. X., Michelet, P., Franken, T. P., et al. (2008). Variations on a Dexterous theme: Peripheral time–intensity trading. *Hear Res, 238*(1), 49-57.

Joris, P. X., Smith, P. H., & Yin, T. C. (1998). Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron, 21*(6), 1235-1238.

Kappauf, W. E. (1975). Regression effect in judgments to determine equal response contours. *Atten Percept Psychophys, 17*(4), 405-410.

Kashino, M., & Nishida, S. y. (1998). Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect. *J Acoust Soc Am, 103*(6), 3597-3604.

Kaufman, P. L., Adler, F. H., Levin, L. A., et al. (2011). *Adler's Physiology of the Eye*: Elsevier Health Sciences.

Kawashima, T., & Sato, T. (2012). Adaptation in sound localization processing induced by interaural time difference in amplitude envelope at high frequencies. *PLoS One, 7*(7), e41328.

Klemm, O. (1920). Untersuchungen über die Lokalisation von Schallreizen IV: über den Einfluss des binauralen Zeitunterschieds auf die Lokalisation. *Arch ges Psychol, 40,* 117-145.

Kopčo, N., Best, V., & Shinn-Cunningham, B. G. (2007). Sound localization with a preceding distractor. *J Acoust Soc Am, 121*(1), 420-432.

Kuwada, S., & Yin, T. C. (1983). Binaural interaction in low-frequency neurons in inferior colliculus of the cat. I. Effects of long interaural delays, intensity, and repetition rate on interaural delay function. *J Neurophysiol, 50*(4), 981-999.

Lang, A.-G., & Buchner, A. (2008). Relative influence of interaural time and intensity differences on lateralization is modulated by attention to one or the other cue. *J Acoust Soc Am, 124*(5), 3120-3131.

Lang, A.-G., & Buchner, A. (2009). Relative influence of interaural time and intensity differences on lateralization is modulated by attention to one or the other cue: 500-Hz sine tones. *J Acoust Soc Am, 126*(5), 2536-2542.

Leys, C., Ley, C., Klein, O., et al. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J Exp Soc Psychol, 49*(4), 764-766.

Li, M., Liu, F., Juusola, M., et al. (2014). Perceptual color map in macaque visual area V4. *J Neurosci, 34*(1), 202-217.

Loftus, W. C., Bishop, D. C., Saint Marie, R. L., et al. (2004). Organization of binaural excitatory and inhibitory inputs to the inferior colliculus from the superior olive. *J Comp Neurol, 472*(3), 330-344.

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *J Acoust Soc Am, 111*(5), 2219-2236.

Magezi, D. A., & Krumbholz, K. (2010). Evidence for opponent-channel coding of interaural time differences in human auditory cortex. *J Neurophysiol, 104*(4), 1997-2007.

Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *J Acoust Soc Am, 87*(5), 2188-2200.

McFadden, D., Jeffress, L. A., & Lakey, J. R. (1972). Differences of interaural phase and level in detection and lateralization: 1000 and 2000 Hz. *J Acoust Soc Am, 52*(4B), 1197-1206.

McFadden, D., & Pasanen, E. G. (1976). Lateralization at high frequencies based on interaural time differences. *J Acoust Soc Am, 59*(3), 634-639.

McLaughlin, S. A., Higgins, N. C., & Stecker, G. C. (2016). Tuning to binaural cues in human auditory cortex. *J Assoc Res Otolaryngol, 17*(1), 37-53.

Meunier, S., Bodden, M., Canévet, G., et al. (1996). Auditory adaptation and localization: Effect of frequency and bandwith. *Acustica, 82,* S215-S215.

Middlebrooks, J. C., & Zook, J. M. (1983). Intrinsic organization of the cat's medial geniculate body identified by projections to binaural response-specific bands in the primary auditory cortex. *J Neurosci, 3*(1), 203-224.

Mills, A. W. (1960). Lateralization of High‐Frequency Tones. *J Acoust Soc Am, 32*(1), 132-134.

Moushegian, G., & Jeffress, L. A. (1959). Role of Interaural Time and Intensity Differences in the Lateralization of Low‐Frequency Tones. *J Acoust Soc Am, 31*(11), 1441-1445.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological methods, 8*(4), 434.

Phillips, D. P., Carmichael, M. E., & Hall, S. E. (2006). Interaction in the perceptual processing of interaural time and level differences. *Hear Res, 211*(1), 96-102.

Phillips, D. P., & Hall, S. E. (2005). Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level. *Hear Res, 202*(1), 188-199.

Pickles, J. O. (2012). *An introduction to the physiology of hearing*: Brill.

Picou, E. M., Moore, T. M., & Ricketts, T. A. (2017). The Effects of Directional Processing on Objective and Subjective Listening Effort. *J Speech Lang Hear Res, 60*, 199-211. doi: https://doi.org/10.1044/2016_jslhr-h-15-0416

Rayleigh, L. (1875). On our perception of the direction of a source of sound. *Proceedings of the Musical Association, 2*, 75-84.

Rayleigh, L. (1907). XII. On our perception of sound direction. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 13*(74), 214-232.

Röser, D. (1965). Schallrichtungsbestimmung bei krankhaft veraendertem Gehoer [The determination of the direction of sound with defective hearing]. *Dissertation, Technische Hochschule*.

Schröger, E. (1996). Interaural time and level differences: integrated or separated processing? *Hear Res, 96*(1), 191-198.

Sechler, S. D., Valdes, A. L., Waechter, S. M., et al. (2017). *Virtual reality sound localization testing in cochlear implant users.* Paper presented at the Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on.

Shaxby, J., & Gage, F. (1932). Studies in the Localization of Sound. *Med. Research Council (Brit.) Spec. Rept, Ser., 166*, 1-32.

Sheldon, P. E. (1973). Equal-onset contours of vibrotactile stimuli. *Atten Percept Psychophys, 13*(3), 403-407.

Shinn-Cunningham, B. G., Santarelli, S., & Kopco, N. (2000). Tori of confusion: Binaural localization cues for sources within reach of a listener. *J Acoust Soc Am, 107*(3), 1627-1636.

Simpson, A., Bond, A., Loeliger, M., et al. (2018). Speech intelligibility benefits of frequency-lowering algorithms in adult hearing aid users: a systematic review and meta-analysis. *Int J Audiol, 57*(4), 249-261.

Stecker, G. C. (2010). Trading of interaural differences in high-rate Gabor click trains. *Hear Res, 268*(1), 202-212.

Stecker, G. C., Harrington, I. A., & Middlebrooks, J. C. (2005). Location coding by opponent neural populations in the auditory cortex. *PLoS Biol, 3*(3), e78.

Stevens, S. S., & Newman, E. B. (1936). The localization of actual sources of sound. *The American Journal of Psychology, 48*(2), 297-306.

Sveistrup, H. (2004). Motor rehabilitation using virtual reality. *J Neuroeng Rehabil, 1*(1), 10.

Teas, D. C. (1962). Lateralization of acoustic transients. *J Acoust Soc Am, 34*(9B), 1460-1465.

Thurlow, W. R., & Jack, C. E. (1973). Some determinants of localization‐adaptation effects for successive auditory stimuli. *J Acoust Soc Am, 53*(6), 1573-1577.

Trahiotis, C., & Kappauf, W. E. (1978). Regression interpretation of differences in time‐intensity trading ratios obtained in studies of laterality using the method of adjustment. *J Acoust Soc Am, 64*(4), 1041-1047.

van Bergeijk, W. A. (1962). Variation on a theme of Bekesy: a model of binaural interaction. *J Acoust Soc Am, 34*(9B), 1431-1437.

Van Veen, H. A., Distler, H. K., Braun, S. J., et al. (1998). Navigating through a virtual city: Using virtual reality technology to study human action and perception. *Future Generation Computer Systems, 14*(3-4), 231-242.

Velenovsky, D. S., Cetas, J. S., Price, R. O., et al. (2003). Functional subregions in primary auditory cortex defined by thalamocortical terminal arbors: an electrophysiological and anterograde labeling study. *J Neurosci, 23*(1), 308-316.

Whitworth, R. H., & Jeffress, L. A. (1961). Time vs intensity in the localization of tones. *J Acoust Soc Am, 33*(7), 925-929.

Wightman, F. L., & Kistler, D. J. (1989). Headphone simulation of free-field listening. I: stimulus synthesis. *J Acoust Soc Am, 85*(2), 858-867.

Wightman, F. L., & Kistler, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am, 91*(3), 1648-1661.

Wittmann, J. (1925). Beiträge zur Analyse des Hörens bei dichotischer Reizaufnahme. *Arch. ges. Psychol, 51*, 21-122.

Yin, T. C. (2002). Neural mechanisms of encoding binaural localization cues in the auditory brainstem *Integrative functions in the mammalian auditory pathway* (pp. 99-159): Springer.

Yost, W. A., Tanis, D. C., Nielsen, D. W., et al. (1975). Interaural time vs. interaural intensity in a lateralization paradigm. *Percept Psychophys, 18*(6), 433-440.

Young Jr, L. L., & Levine, J. (1977). Time-intensity trades revisited. *J Acoust Soc Am, 61*(2), 607-609.

Young, L. L. (1976). Time-intensity trading functions for selected pure tones. *J Speech Lang Hear Res, 19*(1), 55-67.

Zwislocki, J., & Feldman, R. (1956). Just noticeable differences in dichotic phase. *J Acoust Soc Am, 28*(5), 860-864.