VISUAL LOCALIZATION ACCURACY DETERMINES THE BIAS OF

AUDITORY TARGETS IN AZIMUTH AND DEPTH

By

AMANDA BOWEN

A Thesis Submitted to the Graduate Faculty of

WAKE FOREST UNIVERSITY

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF ARTS

in the Department of Psychology

May, 2010

Winston-Salem, North Carolina

Approved By:

James A. Schirillo, Ph.D., Advisor ______________________

Examining Committee

Dale Dagenbach, Ph.D. ______________________

Wayne E. Pratt, Ph.D. ______________________

Ramnarayan Ramachandran, Ph.D. ______________________

ACKNOWLEDGMENTS

I would like to acknowledge my advisor, Jim Schirillo, for his patience and persistence during my introduction to and education of this field, as well as his invaluable guidance during the past 2 years. I would also like to thank my committee members, Dale Dagenbach, Wayne Pratt, and Ramnarayan “Ram” Ramachandran, for their helpful and insightful suggestions and critiques of this project.

Additionally, I would like to extend a special thank you to Jeff Muday, who was generous with his already limited time by programming my thesis project, as well as my first year project. And, of course, I deeply appreciate, and have great sympathy for, all of my participants over the past 2 years who spent hours upon hours sitting in a dark room localizing sounds and lights.

I am especially grateful to Wake Forest University for funding my continued education and to the psychology department staff, faculty, and fellow graduate students. This was truly a supportive and enriching environment of which to be a part.

Finally, I would like to acknowledge my parents, without whose unconditional love and support I would most certainly not be where I am today.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Amanda Bowen

VISUAL LOCALIZATION ACCURACY DETERMINES THE BIAS OF AUDITORY TARGETS IN AZIMUTH AND DEPTH

Thesis under the direction of James A. Schirillo, Ph.D., Professor of Psychology.

Visual bias is a psychophysical phenomenon where an accurately localized signal, such as a light, will bias a spatially discrepant signal that is localized with less accuracy, such as a sound, when the two signals are perceived as unified. Many previous studies have demonstrated visual bias in azimuth, but none have directly tested, or found, this effect in depth. By increasing the saliency of the auditory and visual signals from a previous experiment, the current study was able to produce more accurate depth discrimination. Visual localization error was smaller than that of auditory localization error in both azimuth and depth, and visual bias was demonstrated in both dimensions. Additionally, visual localization accuracy was positively correlated to visual bias on perceived unified trials.

INTRODUCTION

The accurate perception of our everyday world not only involves the ability to localize auditory and visual stimuli in space, but also our ability to combine information from the two modalities. Oftentimes, our ability to localize an auditory signal is poor, in terms of accuracy and variability, and for this reason we typically rely more on visual signals to inform us about our environment. Visual signals tend to be localized more accurately and with less variability, and tend to bias the less reliable auditory signals towards them when they are presented in both spatial and temporal proximity (Bertelson & Aschersleben, 1998; Bertelson & Radeau, 1981; Hairston et al., 2003; Wallace et al., 2004; Warren, 1979; Warren, Welch, & McCarthy, 1981). This perceptual phenomenon is referred to as the *ventriloquist effect*. A practical example of this effect is when one is watching a movie in a theater and the voices of the actors, despite coming from the surrounding area, appear to be coming from the screen.

Whereas some have theorized that visual bias is due to a complete capture of the weaker auditory signal by the stronger visual signal (Bertelson & Radeau, 1981), a more supported theory states that multiple modalities are combined in a statistically optimal fashion. Maximum Likelihood Estimation (MLE) suggests that we weight auditory and visual signals according to their reliability in order to make the most precise localization estimates possible given the information present (Alais & Burr, 2004; Battaglia, Jacobs, & Aslin, 2003; Ernst & Banks, 2002; Godfroy, Welch, Sandor, & Roumes, 2008). The following thesis will examine both the localization of auditory and visual signals in two dimensions, azimuth and depth, as well as calculate the visual bias of a cross-modal auditory-visual signal and examine the relationship between bias and the variance and accuracy of the unimodal localizations.

Visual & Auditory Localization

The localization of auditory and visual targets in azimuth (i.e., lateral space) and depth depend on our ability to integrate several cues within each modality. Such auditory cues include the spectral signature created by the shape of the outer ear, or pinna (Blake & Sekuler, 2006), and binaural cues, or the differences in timing (ITDs) and intensity (IIDs) at which sounds hit the right and left ear. ITDs tend to be particularly important for judging the direction of sounds and are most distinguishable in low frequency sounds since longer wavelengths can travel around the head. IIDs are helpful in judging the distance of sounds and are typically more distinguishable in high frequency sounds as sound waves are diffracted when they are shorter than the diameter of the head (Brungart & Rabinowitz, 1999; Hartmann, 1999). Errors in distance perception occur most often for sounds directly in front and behind of the individual, as that is when interaural differences are most ambiguous.

Brungart and Rabinowitz (1999) demonstrated the importance of interaural differences in auditory depth perception. In an anechoic chamber using a Kenmar manikin with a microphone placed at the eardrum locations, head-related transfer functions (HRTFs), or synthetic sound recordings that take binaural and pinna differences into account, were measured from sounds produced within a relatively small distance (0.12 m to 1 m). An acoustic point source produced a moving periodic chirp signal ranging from 200 Hz to 15 kHz and the HRTF recordings showed that differences in the timing and intensity of the sound hitting each ear increased as the sound moved closer to the manikin. ITDs increased only slightly as distance decreased, and the majority of this increase happened when signals were in the median plane. IIDs were near 0 dB in the median plane (right in front of the head), and increased as sources moved lateral to the head.

When binaural cues are weak (especially in the median plane), distance judgments greatly depend on amplitude, which decreases as the distance between an individual and a sound source increases (Ashmead, LeRoy, & Odom, 1990; Brungart, 1999). Brungart (1999) found superior localization performance in experimental conditions containing fixed-amplitude sounds, meaning loudness varied only with changes in distance, allowing individuals to derive sound distance information. Additionally, sound intensity, or pressure, varies as the inverse of the distance to a sound. Altering pressure cues so as to render them useless for judging a sound source's distance (i.e., making them equal at the participants' ears despite the sound's increasing distance) causes the discrimination threshold for distance to increase significantly (Ashmead, LeRoy, & Odom, 1990).

The reverberations from a sound source also affect the way in which auditory signals are perceived. The ratio of direct-to-reverberant energy (*v*) varies with sound source distance; a larger ratio of direct-to-reverberant energy typically signifies a closer sound source (Butler, Levy, & Neff, 1980). In this way, sounds with more reverberations are perceived as being further away than those with less. However, when only reverberations differ among auditory signals, localizations are somewhat variable (Zahorik, 2002b). Additionally, sound frequency is helpful for localization (Coleman, 1968), and, as low frequency sounds contain more reverberations than high frequency sounds, they are typically perceived as being further away. Butler, Levy, and Neff (1980) demonstrated the aforementioned finding by presenting participants with 5 s trains of broadband noise bursts that were either high pass or low pass filtered. Localization responses demonstrated that low frequency sounds were always perceived as being further from the participant than high frequency sounds, regardless of room type. Overall, with many auditory cues available for localization of auditory signals, individuals typically gain information about a sound source according to what features are most salient about that particular sound, and this saliency affected by sound type, distance, and direction (Zahorik, 2002a).

Visual cues, including stereopsis and head/eye orientations, provide participants with the ability to judge light locations in two dimensions. Stereo vision enables individuals to see objects in depth in near space as binocular disparity, which is produced by a slight disparity between the left and right eye images when viewing an object (Blake & Sekuler, 2006; Schiffman, 2000). Stereo acuity, or the ability to gauge the depth between objects, is found to be better for objects that are closer to a fixation point rather than farther from a fixation point, because the disparity threshold (or the threshold at which you can resolve small disparities) decreases as the distance between a fixation point and a target point decreases (McKee, Levi, & Bowne, 1990). Objects that are further than a fixation point appear at an uncrossed disparity, whereas objects closer than a fixation point are at a crossed disparity. Because these objects are not at fixation, they project onto non-corresponding points on the back of the retina which provides a relative distance cue, albeit an unconscious one (Schiffman, 2000).

The angle between the line from one eye to another and from one eye to the target represents the ocular vergence angle (Blohm, Khan, Ren, Schreiber, & Crawford, 2008).

The angles differ for objects of different distances, serving as a cue to the absolute distance of an object (Viguier, Clement, & Trotter, 2001). Neurons, which are activated by specified regions in space, produce movements in eyes, ears (in animals with moveable ears), and the head to focus on locations of sounds and lights, providing localization information (Stein & Meredith, 1993). Convergence occurs when the eyes turn toward each other to view an object that is close to an individual, indicating that the muscles used to control eye movements may provide an indication of target distance (Schiffman, 2000).

Our ability to use visual cues to localize stimuli is typically more accurate than our ability to use auditory cues to localize sounds. Indeed, previous research has found that participants are less variable and more accurate in their visual target localizations than auditory target localizations (Alais & Burr, 2004; Battaglia, Jacobs, & Aslin, 2003; Godfroy et al., 2008; Hairston et al., 2003). Results from a previous study also showed that visual localization was more accurate than auditory, but only in the azimuth dimension (Bowen, 2009). In addition, temporally and spatially aligning visual and auditory stimuli result in localization performance that is less error-ridden and less variable than the best modality alone (Godfroy et al., 2008). Moreover, giving individuals signals from multiple sensory modalities also typically decreases reaction time, as was found by Barutchu et al. (2010) and Bowen (2009).

However, integrating information from multiple sensory modalities is not entirely straightforward. The superior colliculus is involved with integrating information from both visual and auditory inputs. The superior colliculus is organized in a map-like fashion with specified regions that are responsive to stimuli contained within certain areas in space, and these regions are called receptive fields. Parts of the receptive fields of multisensory neurons overlap but other parts do not. Thus, stimulus combinations only produce stronger responses when two stimuli are in spatial or temporal proximity, and when the stimuli do not produce a strong effect on their own (Anastasio, Patton, & Belkacem-Boussaid, 2000; Stein & Meredith, 1993). As such, it is expected in the current study that visual localization will be more accurate and less variable than auditory localization and that auditory stimuli, when paired with spatially coincident visual stimuli, will be localized with more accuracy and less variability on perceived unified trials than all other multisensory conditions. I anticipate replicating this effect in azimuth, as well as producing the effect in depth by increasing the discriminability of the signals in depth.

Visual Bias

More precise localization in the visual modality tends to bias auditory signals in azimuth. This robust finding is known as the ventriloquist effect, much as a ventriloquist can deceive their audience into perceiving words as coming from a dummy instead of the ventriloquist. For example, Bertelson and Radeau (1981) found that competing visual signals from 7˚ to 25˚ in azimuth will pull the perceived location of a centrally located auditory target to the left or right. Similarly, Bertelson and Aschersleben (1998) demonstrated that a central light will cause participants to localize auditory targets that are moving from the right or left toward a center point to be localized on the opposite side of their origination, showing that visual bias is a perceptual process.

Visual bias is also affected by cognitive processes, specifically, the judgment of unity. For example, in a follow-up to the first study by Bertelson and Radeau (1981), when participants were asked whether they perceived the light and sound to be fused, there was an overall reduction in the number of times they incorrectly identified an auditory target as coming from the visual signal's location and bias occurred only on trials indicated as fused. Additionally, Warren (1979) presented participants with visual and auditory or visual and proprioceptive (touch) information together while displacing the visual information using a visually distorting prism. Participants biased auditory and haptic targets to the visual signals in both instances, but bias was reduced when the researchers informed participants that the visual information they were receiving may be unreliable. Warren, Welch, and McCarthy (1981) explained that only when there is compelling evidence that two cues are originating from the same event and participants make a "unity assumption" do they bias. Choe, Welch, Gilford, and Juola (1975) found that participants are more likely to perceive lights and sounds as being unified when they occur simultaneously than when they are separated by even a short time period. Thus, the co-occurrence of visual and auditory stimuli aids in the perception of unity, which is oftentimes required for visual signals to bias auditory ones.

Other literature in the extensive research on visual bias in azimuth has indicated the importance of unification judgments. Hairston et al. (2003) had participants localize either visual or auditory targets only in unimodal trials, or auditory targets in the presence of distracting (irrelevant) visual signals that were either coincident or of increasing spatial disparities in cross-modal trials. Participants also made unity judgments in each cross-modal trial, which indicated whether they perceived the lights and sounds as coming from the same or different locations. Variability was smaller for visual-only and spatially coincident multisensory localizations than for auditory-only localizations. Thus, in trials where visual and auditory stimuli were spatially disparate, there was a bias of auditory stimuli towards the light and bias was positively correlated to unity judgments (i.e., more bias was found for trials that participants reported as being unified).

In a follow-up study to Hairston et al. (2003), Wallace et al. (2004) examined the relationship between bias, perceptual unity, and localization variability using both spatially disparate and temporally disparate lights and sounds. They presented participants with a sound in conjunction with either a spatially disparate or temporally disparate light flash. The participants' task was to localize the sound and then make a unification judgment. When participants judged the sound and light as unified, they almost completely biased the sound to both spatially and temporally disparate lights. However, on trials where the visual and auditory stimuli were close but perceived as being nonunified, the participants demonstrated a negative bias, indicating that they perceived the sound as repulsed *away* from the light. This negative bias was also found in a study by Körding et al. (2007). They showed that when participants perceived a sound and light, situated lateral to each other in space, to be nonunified, they pushed the auditory signal away from the visual signal.

While there have been numerous multisensory localization studies in azimuth, very few studies have examined the localization of auditory and visual stimuli in depth. In the auditory-only domain, Mershon and Bowers (1979) produced sound signals that were 1, 2, and 6 m away from a participant who was asked to report the perceived distance of the sound in feet and centimeters. Brungart, Durlach, and Rabinowitz (1999) similarly had participants judge auditory distance by moving a response sensor to the perceived origination of sounds that were presented from 10 cm to 1 m away. Zahorik's (2002a) participants also judged the distance of sounds (including white noise and speech samples) presented from 0.3 m to 13.8 m away. Results from these studies demonstrated that participants were somewhat inaccurate at auditory depth localization, whereas auditory localization in azimuth tends to be fairly correlated with actual sound locations (Brungart, Durlach, & Rabinowitz, 1999). In particular, participants tend to overestimate close sounds and underestimate far sounds, with estimates becoming more error ridden as sounds are presented further away. Likewise, visual localization in depth, unlike visual localization in azimuth, is not extremely accurate, as Viguier, Clement, and Trotter (2001) found that participants' localizations became worse as light sources increased in distance from 20 cm to 120 cm.

Similar to unimodal localization studies in depth, there have been relatively few studies devoted to multisensory localization in depth. Gardener (1968) found that by placing speakers in a row (ranging from 3 to 30 feet) in an ancheoic chamber, such that only the front speaker was visible, participants localized all auditory stimuli at the nearest speaker. Mershon, Desaulniers, Amerson, and Keifer (1980) expanded on Gardener's (1968) finding by using reverberant rooms and demonstrating that participants will bias a sound towards a visible "dummy" loudspeaker that is either near or far from the sound, thus demonstrating 'visual capture.' While Zahorik (2001) employed an experimental method very similar to Gardener (1968) and Mershon et al. (1980), he found results that conflicted with both of these studies. Zahorik (2001) used sounds recorded in an ancheoic chamber and reproduced them from 5 loudspeakers ranging from 1 m to 5 m away (with only the nearest loudspeaker being visible) in a semi-reverberant room. The participants were either blindfolded or had full use of their eyesight. When participants were able to see the loudspeakers, they were much more accurate at judging the distance of each sound presented, and, for this reason, they did not show visual capture. The results of these studies clearly demonstrate that the only few multisensory localization studies in depth have produced confusing findings regarding bias.

The findings of visual bias have been particularly robust in the azimuth dimension, and, as such, are expected to be reproduced in the current study. However, this finding has not been reliably extended to the depth dimension. A previous study exploring this phenomenon was unable to produce visual bias in depth, perhaps due to the experimental set-up (Bowen, 2009). Consequently, the current study employed techniques to increase the saliency of the visual and auditory signals in order to increase depth discrimination. If participants are better able to discriminate the signals in depth, then they should show visual bias on perceived unified trials in depth as well as azimuth.

Maximum Likelihood Estimation

There have been competing reasons suggested for why visual signals tend to bias auditory ones. Some theorists have suggested that the stronger visual signal completely dominates the auditory signal, known as visual capture. On a related note, Bertelson and Radeau (1981) proposed a criterion-shift theory, stating that whenever a participant fuses two signals, the location of the two signals is based on only one modality while the other modality is ignored. However, studies of visual bias do not typically produce 100% bias, as would be expected if localization is based on the biasing modality alone. An alternative theory to visual capture may be a better fit with the previous literature on bias. A Maximum Likelihood Estimation (MLE) theory has been proposed to explain visual bias; specifically, visual signals are more reliable than auditory (i.e., they show less variability and error) and, therefore, pull auditory targets toward them. Ernst and Banks

(2002) found that visual-only localization was less variable than haptic-only localization. Because of the smaller variance (i.e., higher reliability of visual information), participants assigned more weight to the visual, as opposed to the haptic signal, in cross-modal trials. Ernst and Banks (2002) were able to accurately predict multisensory localizations using the MLE model. Additionally, through increasing degradations to the visual stimuli, they were able to show how participants continue to weight those signals less in cross-modal trials, in-line with MLE predictions. Other researchers have produced similar results using visual and auditory signals (Alais & Burr, 2004; Battaglia, Jacobs, & Aslin, 2003).

Because the distribution of multisensory localization responses tend to fit an MLE model, this suggests that participants are combining visual and auditory signals in a statistically optimal fashion. The MLE model estimates the weight of each modality based on the variance of each signal; the more variance a modality has compared to the other modality, the less weight it is assigned. The variances of the localization of visual and auditory signals in the unimodal trials are used to predict the amount of reliance that participants will place on the two signals when they are presented with both in cross-modal trials. Moreover, these weights are used to produce an estimate of stimulus localization that is more reliable than the best unisensory estimate (Battaglia, Jacobs, & Aslin, 2003; Ernst & Banks, 2002). Therefore, the variance of the bimodal distribution should be smaller than either single modality alone (Witten & Knudsen, 2005). Based on the MLE model's predictions, it is hypothesized in the current study that multisensory localization on trials in which the light and sound are spatially coincident and perceived as being unified by the participant should be more accurate and less variable than the localizations made in the best unimodal condition (i.e., most likely, vision).

Godfroy et al. (2008) demonstrated that localizations made in unimodal visual trials were superior to those of auditory trials, but the most precise localizations were made on trials where participants were supplied with both signals, suggesting that congruent stimuli produce localization that is just as or more precise than that of the best modality alone. Participants made localizations on visual- or auditory-only trials, as well as bimodal trials in which the visual and auditory signals are spatially and temporally coincident. Visual localization was more accurate and less variable than auditory localization. Most importantly, localization was least variable when participants were presented with both auditory and visual information. In the current study, participants were presented with both spatially coincident and noncoincident signals, but it has been shown that the perception of unity is more important for finding visual bias than the actual spatial coincidence of the signals (Hariston et al., 2003; Körding et al., 2007; Wallace et al., 2004; Warren, 1979; Warren, Welch, & McCarthy, 1981), and for this reason, visual bias is expected in all trials perceived as unified. However, only auditory targets from cross-modal trials that are actually spatially coincident and perceived as unified should be localized with superior accuracy and less variability.

Despite the poor accuracy with which auditory signals are localized compared to visual signals, auditory localization in azimuth has been shown to be superior to auditory localization in depth. In a study by Brungart, Durlach, and Rabinowitz (1999), judgments of distance became more error-ridden with increasing distance and when sounds were closer to the median (versus lateral) plane, but azimuth localizations were highly correlated with the actual locations of the sounds in azimuth. For this reason, it is reasonable to expect that error should be higher for both auditory and visual signals in depth than in azimuth, and localization should be less variable in azimuth versus depth in cross-modal trials. However, it is expected that in order for participants to demonstrate bias in both azimuth and depth, visual localization must be, to some extent, accurate with low variability, in the unimodal trials of both dimensions. This leads to the hypothesis that visual accuracy will be predictive of visual bias on perceived unified trials; i.e., if participants demonstrate a very small amount of error in visual unisensory localization, they will be more likely to bias to the visual stimuli in cross-modal trials.

It is unclear whether it is visual accuracy or the auditory to visual variance, or both, that is related to bias. It is possible that a higher auditory to visual variance ratio is related to bias on perceived unified trials. Specifically, if auditory localization precision is low and visual localization precision is high in unimodal trials, one would expect participants to rely more on the visual signal than auditory signal in multimodal trials. Therefore, there should be a positive correlation between the auditory-to-visual variance ratio and bias on perceived unified trials. This would be supportive of a MLE model's predictions that unimodal variance can be used to predict the weight participants place on the two signals in multimodal trials, and, thus, the distribution of responses in said trials.

Similarly to the MLE model, a Bayesian model also makes estimations about outcomes using the variance models. However, a Bayesian model makes a different assumption than the MLE model. Specifically, the MLE model does not use prior information regarding the outcome of the data for its estimations. The MLE assumes a prior of one, meaning all instances are equally likely to occur (i.e., no previous information). However, a Bayesian theorem uses prior information, which can be anything from previous statistical information (e.g., believing a fair coin will end up on heads 50% of the time) to the fact that the human visual system is biased due to its greater reliability than other sensory systems. This prior information is then used along with the likelihood distribution, which is the probability of the data given the population mean or the stimulus source. The result is a posterior probability, which is the probability of the stimulus or population mean given the data, which, in turn, updates the prior. In this way, the prior in a Bayesian model is not a fixed value but a distribution that is constantly being updated according to the data (Jacobs, 2008; Witten & Knudsen, 2005).

The Bayesian theorem also accounts for the enhancement of localization when an auditory and a visual signal are presented at the same time and place. With increasing neural input from a signal, a target will go from most likely spontaneous (the rate of detection of a signal when no signal is present due to random neural firing) to most likely driven (the rate at which one is detecting a signal that is actually present). When two signals are present and one signal has a strong neural input (e.g., the visual signal), the detection of that signal is not greatly enhanced by a second signal (e.g., the auditory signal). However, because the auditory signal is weak, the detection of the signal becomes much more reliable when accompanied by a visual signal. Overall, pairing two signals together will result in an enhancement of localization (i.e., result in less error and variability), but the enhancement is greater for weak stimuli (Anastasio, Patton, & Belkacem-Boussaid, 2000). Therefore, in the current study, the more variable and less accurate auditory-only localizations are, the more the visual signal will influence their localization on spatially coincident trials.

Given what is known about visual and auditory localization, an MLE theorem can explain how individuals bias auditory stimuli towards visual stimuli, but an MLE theorem cannot be continuously updated with information from the data to better estimate localization on cross-modal trials. This may explain why some researchers have found that a Bayesian model better fits their data than an MLE model. For example, both Alais and Burr (2004) and Battaglia, Jacobs, and Aslin (2003) found that the visual system receives a disproportionate amount of weight compared to the auditory system, and, even when the visual stimuli is extremely degraded, individuals still rely on both auditory and visual cues relatively equally on cross-modal trials. This means that even when the visual stimulus is poor, individuals still rely on that system to a relatively large degree compared to an intact auditory signal. A Bayesian theorem can account for the preference of the visual system through the use of a prior, whereas the MLE model cannot. Examination of the comparison of the MLE and Bayesian models are outside the scope of the current experiment, but it is worth noting the similarities of the variance model proposed by the two theories as well as the advantage the Bayesian model has in estimating the distribution of responses in multisensory localization.

The Current Study

Overall, auditory and visual localizations made in azimuth tend to be more accurate and less variable than auditory and visual localizations made in depth. In a previous study, participants localized stimuli that were 200 msec in duration. They made localizations of either a visual or auditory stimulus in unimodal sessions or of an auditory stimulus in the presence of a visual distracter in cross-modal sessions, after which they made a unification judgment. As expected, participants biased their auditory localizations to the visual distracters on perceived unified trials, but not on perceived nonunified trials. However, this effect occurred only in azimuth, not in depth, likely due to the poor accuracy and high variability of both visual and auditory localizations made in depth on unimodal trials (Bowen, 2009).

The current study employed methods to increase the saliency of both visual and auditory targets in order to increase depth discrimination among participants. These methods to increase the saliency of visual information include using multiple fixation planes composed of noius points that were present for the duration of the experimental trials and signals that were Gaussian envelopes and presented for a long (i.e., 1000 msec) duration. Additionally, sound recordings with a flatter spectral function so all frequencies are equally represented, and a long duration sound (i.e., a 700 msec signal with 300 msec reverberation time) were used to increase the saliency of the auditory targets. If target saliency was increased adequately enough for participants to gain information about depth it is expected that visual and auditory unimodal localizations will follow the same pattern in depth as in azimuth, and visual bias should also be demonstrated in depth. However, all things being equal, individuals should still be better at localizing visual and auditory signals in azimuth than visual and auditory signals in depth. What is unknown is whether auditory localization variance will exceed visual localization variance to a greater extent in azimuth versus depth. If the ratio is larger in azimuth, then it is reasonable to expect individuals to demonstrate more visual bias in azimuth than in depth; this may be supportive of the MLE model's prediction that unimodal variances are in-line with the weights participants give the two signals in cross-modal trials. Consequently, this work will expand on the robust effect of visual bias by using unimodal variance to predict the extent of bias of cross-modal visual-auditory stimuli in azimuth and depth.

The current study aims to garner support for four specific hypotheses. First, multisensory localizations on coincident trials that are also perceived as unified are expected to be more accurate than all other cross-modal conditions. Additionally, localizations made on cross-modal trials that are coincident and perceived as unified are expected to be more accurate and less variable than localizations made in the most accurate sensory modality. Furthermore, if visual localization is more accurate than auditory localizations in unimodal trials in azimuth and depth, then visual bias is expected to be apparent in both dimensions for spatially noncoincident cases perceived as unified. Finally, it is expected that there will be a positive correlation between bias on perceived unified trials and visual accuracy, as well as the ratio of auditory to visual localization variance.

METHODS

Participants

Five experimentally naïve work-study students from the Wake Forest University psychology department participated in the current study. These students received payment from the department for their participation in the study. Additionally, two experimentally naïve individuals from the community agreed to participate in the study. All procedures were approved by the Institutional Review Board of Wake Forest University and were performed in accordance with the ethical standards established by the 1964 Declaration of Helsinki.

Materials

The experiment was conducted in a 4' x 10' non-anechoic psychophysics laboratory. Before the sound recordings and experimental trials took place, the visual and auditory stimuli were calibrated. This ensured that the auditory and visual stimuli were beginning and ending at the same time. This calibration was conducted using an Agilent 54641A two-channel Oscilloscope 350 MHz, 2 GSa/s. Sound reference measurements were also made to ensure that when a sound was played, the amplitude of that signal was known. These reference measurements were made at the location of each observer's head by placing a ½'' condenser microphone from Brüel & Kjær (omni-directional cartridge type 4191, 200V, free-field, 3Hz – 40kHz) enhanced by a ½'' microphone preamplifier (Type 2669-B), and measuring amplifier (Type 2609) at the location of the chin rest.

During an initial setting, binaural-related impulse responses (BRIRS) were constructed for each participant by positioning a loudspeaker (Infinity Reference 3002cf 3.5 Speaker System Frequency Response 85Hz – 21kHz; 75 Watts; 4 Ohms SPL 92 dB) at each of 9 different locations. Two Sennhesier (KE4-211-2) miniature-ear capsule microphones enclosed by Etymotic Research ER-13R-2 ring seals were placed in each observer's ears and sounds were recorded by the loudspeaker from each of the locations. These recordings were then played back as virtual sounds, and the BRIRs accounted for the unique transfer characteristics of the observer's pinna, outer ear canal, head and torso, as well as the experimental room reverberations just as accurately as synthetic HRTFs (Savioja, Huopaniemi, Lokki, & Väänänen, 1990; Vesa & Lokki, 2006).

Wightman and Kistler (1989a) produced a veridical experience of free field sounds from stimuli presented through headphones. They recorded HRTFs from both ears of participants by placing a miniature microphone into the ears. There was a large amount of interindividual variability among HRTFs, as would be expected among different individuals. Additionally, waveforms produced from ear-canal recordings closely resembled those of waveforms produced from free field sounds. In an additional study, Wightman and Kistler (1989b) had participants localize acoustics that they heard from the free field or over headphones. The HRTFs were recorded to sound as if they were originating from the same location as the free field acoustics and participants localized the acoustics in virtually the same locations in 3 dimensions. In the current experiment, sounds were recorded from actual areas in space, and, as Plenge (1974) demonstrated, sounds containing individualized interaural differences will produce a veridical external projection. We are confident that these headphone sounds were externalized by the participants, as Levy and Butler (1978) further demonstrated that the ability to externalize sounds is strong; participants in Levy and Butler's study (1978) were even able to externalize sounds lacking interaural differences in time and intensity. An example of a recorded BRIR from location 3 (i.e., the far right) can be seen in Figure 1.

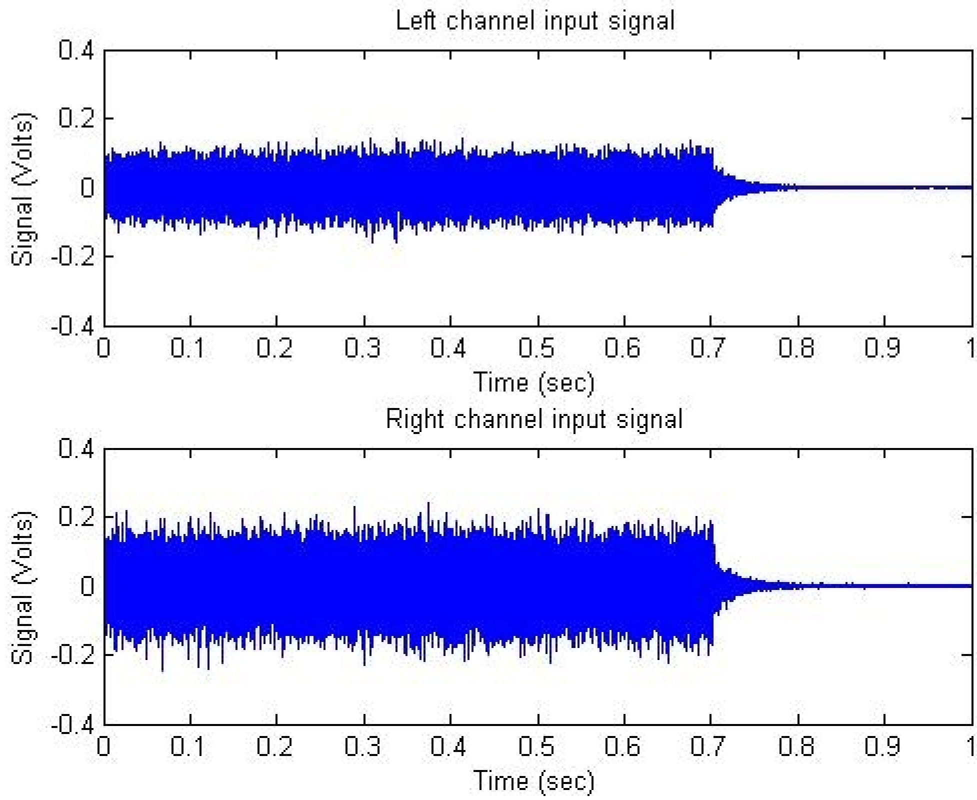

Figure 1. Sound Recording: Recording of sound from location 3. Notice the IID (span of signal) and ITD (onset of signal) differences between the left (top) and right (bottom) ears. This sound signal lasts for 700 msec with 300 msec of reverberations.

The sound signal consisted of a 700 msec impulse response using a maximum length sequence followed by 300 msec of reverberations; the reverberations were generated by the room while the sound signal was generated by Tucker-Davis Technologies software (System 3: 2 DSP Piranha Multiprocessor for 2 channel D/A A/D conversion). The sounds were played to participants through dynamic, open, diffuse-field studio headphones from Beyerdynamic (DT 990 Pro; Nominal frequency response 5Hz – 35kHz).

The visual stimuli were produced using a 20'' flat-profile, fixed-frequency Clinton Monoray monitor (640 x 480 spatial resolution) at 200 Hz to provide flicker-free stereo-fusion (100 Hz to each eye on alternate frames). The yellow (CIE x = 0.43, y = 0.54, 200 cd/m^2 calibrated luminance) DP104 phosphor CRT is designed for stereoscopic applications decaying to 0.1% peak intensity within 0.6 msec. The CRT monitor was situated 70 cm away from the participant and FE-1 ferroelectric Shutter Goggles driven by Cambridge Research Systems Visual Stimulus Generator VSG 2/5 were used to view the visual stimuli (75 cd/m^2 Gaussian envelope 1° visual angle light flash for 1000 msec) and target for localization (60 cd/m^2 0.1° square visual probe). These stimuli looked like a yellow dot or square flashed on the screen as stereoimages. A joystick was used to make localizations.

Design and Analyses

The sound and light locations consisted of three arcs of increasing depth from the participant to a CRT monitor screen (situated 70 cm away from the participant), with three sounds situated lateral to each other in each arc (see Figure 2). All center stimuli were located at 0° azimuth, while the locations to the right and left of each central point were located at 8.78° visual angle (i.e., stimuli 1 and 3 were 8.78° from 2, 4 and 6 were 8.78° from 5, 7 and 9 were 8.78° from 8). The center stimuli 2, 5, and 8, were located at 83.30 cm, 69.30 cm, and 54.95 cm in depth, respectively. Stimuli 1 and 3 were located at 83.23 cm, targets 4 and 6 at 69.06 cm, and stimuli 7 and 9 at 54.89 cm away from the participants. Fixation was established through the use of multiple noius points that were present on the screen for the entire duration of the experimental session that created multiple fixation planes. The back arc of signals were at an uncrossed disparity, meaning they appeared to be located "within" the computer monitor and behind fixation, whereas the front row appeared at a crossed disparity, or closer to the participant than fixation (Schiffman, 2000).

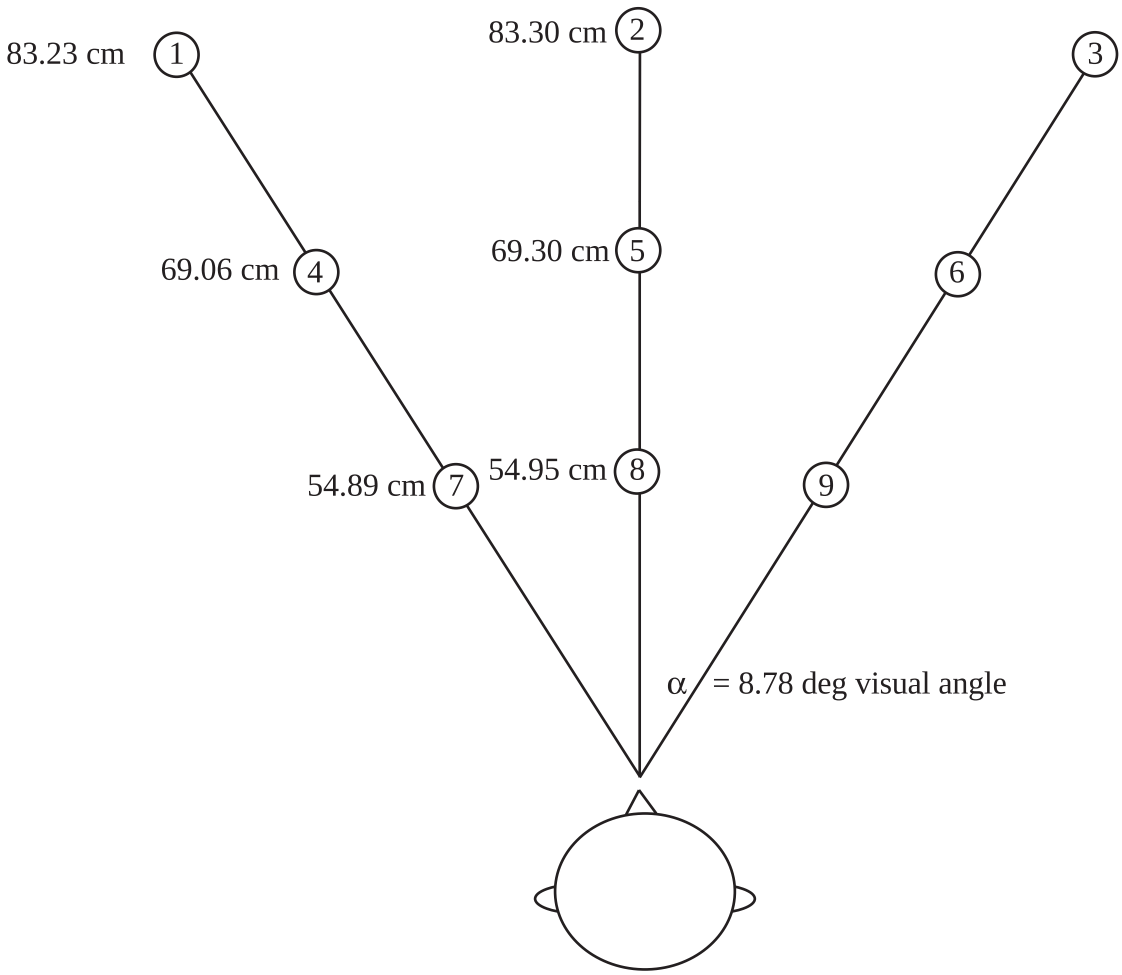

Figure 2. Sound and Light Locations: Sounds and lights were produced from 9 different locations during experimental trials. Signals 1, 3, 4, 6, 7, and 9 are located at 8.78° visual angle from central signals, and each arc is of increasing depth, from 54.89 cm to 83.30 cm, away from the participant.

The participants completed two unimodal sessions, one at the beginning of the series of experimental sessions, and a second at the completion of all experimental sessions. In this way, the first unimodal session could be compared to the second unimodal session for practice effects. Participants completed eight cross-modal sessions in between their completion of the two unimodal sessions. In each unimodal session, the participants completed 360 trials where they were presented with a 1000 msec sound burst, consisting of 700 msec direct sound and 300 msec reverberation or a Gaussian envelope 1° visual angle light flash for 1000 msec in random sequence. The sounds and lights were presented at each of the nine locations once per block (18 stimuli) in a random order and each block was repeated 20 times. The participant's task was to move a small (i.e., 0.1°) square visual probe in stereo-depth that initially appeared in the center of the screen, after the visual or auditory stimulus was presented, to where they perceived the stimulus to have originated. A joystick was used to move the probe side to side in azimuth and back and forth in depth before clicking a button on the joystick to indicate their localization response. This direct localization method has been shown to produce more accurate localization estimates than other methods, such as pointing or verbally estimating distance and direction (Brungart, Rabinowitz, & Durlach, 2000).

In each cross-modal session, the light and sound were presented simultaneously from either the same location (72 trials) or from separate locations (72 trials) randomly intermixed, and this was completed twice for a total of 288 trials per session. Again, the small, moveable visual probe appeared in the center of the screen and participants moved this probe with the joystick, clicking the trigger after positioning it at their localization response to indicate their localization. Participants were instructed to localize the auditory stimuli only, not the visual stimulus. They were made aware of the fact that on some trials the lights and sounds would be spatially coincident, while on other trials they would be noncoincident. They were instructed to keep their eyes fixated on the noius points during the stimulus presentation and in between trials and to try to not move their eyes toward the presented stimulus until the moveable probe appeared in the center of the screen. The participants were also asked to remain as accurate as possible while localizing as quickly as possible. They were instructed to always localize the stimulus where they first perceived the stimulus as originating from. After making a localization response, the participants indicated whether they perceived the two stimuli as being unified or nonunified by either moving the joystick forward to indicate unification or backward to indicate nonunification and clicking the joystick button a second time to record their response.

Upon completion of the experimental trials, the data for each participant were aggregated which constituted a single multisensory data set and two separate unisensory data sets for the first and second unimodal sessions. The localization responses were converted from pixel units to cm. The distance of localizations in depth were computed by taking the square root of the sum of the squared cm values of the azimuth localizations and depth localizations for each trial. The localizations made in azimuth were converted into degrees of visual angle. All trials in which localizations were made closer than 0 cm in depth (at the participant's nose) or further than 160 cm in depth (double the distance of the furthest signal) were removed from all data sets. These data sets were used for all analyses.

RESULTS

Unimodal Trials

In unimodal trials, participants localized either a visual or auditory target. Response time, accuracy, and variability of localizations were computed.

Response Time

Response time, measured in msec, is defined as how long participants took to make a two-dimensional localization. It is the time between when the sound or light was flashed and when participants moved the joystick to the perceived origin of the stimulus and pressed a button on the joystick to make their response. Response time did not differ between visual localization (M = 3202.97, SD = 268.06) and auditory localization (M = 3278.45, SD = 424.17), $F(1, 6) = .14$, $p = .72$, in the first unimodal session. However, the difference in response time between visual (M = 2036.36, SD = 469.51) and auditory localization (M = 2511.65, SD = 866.97) in the second unimodal session approached significance, $F(1, 6) = 5.62$, $p = .06$, $\eta^2 = .48$ (see Figure 3).

Notably, participants were faster at responding to visual targets in unimodal session two than in session one, $F(1, 6) = 27.96$, $p = .002$, $\eta^2 = .82$, and while the same was true for auditory target response, $F(1, 6) = 6.69$, $p = .04$, $\eta^2 = .53$, this effect was smaller. This constitutes a practice effect for response time from the first to the second unimodal sessions, which seems to be larger for visual localization compared to auditory localization.

Accuracy

Accuracy was measured by the amount of error (in cm in depth localization and in degrees of visual angle in azimuth localization) between participants' localization of a

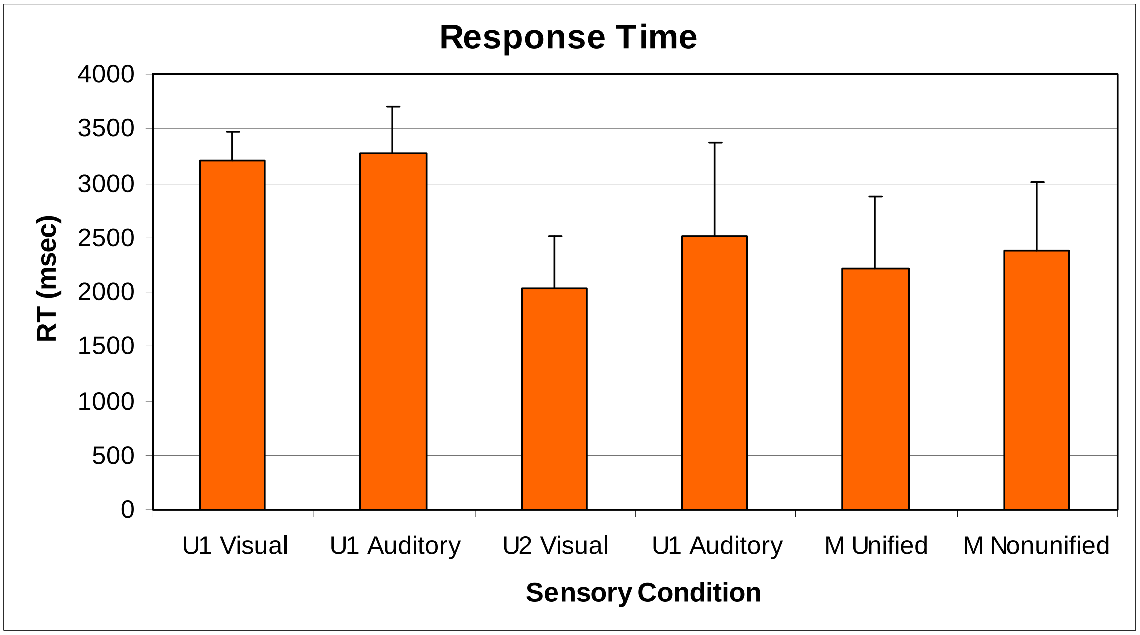

Figure 3. Response Time: Response time, measured in msec and plotted on the y-axis, is the time between the appearance of the movable probe on the screen and the when the participant clicked the joystick to make their localization. Sensory conditions are plotted on the x-axis. Error bars are standard deviations.

visual or auditory target and the target's actual location in each trial. For the depth dimension, visual localization in the first unimodal session (M = 5.87, SD = 3.99) was more accurate than auditory localization (M = 21.91, SD = 9.28), $F(1, 6) = 18.43$, $p = .005$, $\eta^2 = .75$. This finding was also true for visual (M = 6.00, SD = 4.22) and auditory localization (M = 10.2, SD = 5.03) in depth in the second unimodal session, $F(1, 6) = 17.81$, $p = .006$, $\eta^2 = .75$. Like much of the literature on visual localization in azimuth, accuracy for visual localization (M = 0.87, SD = 0.47, session one; M = 0.90, SD = 0.55, session two) was better than auditory localization (M = 9.51, SD = 5.71, session one; M = 4.84, SD = 2.12, session two), for session one, $F(1, 6) = 15.91$, $p = .007$, $\eta^2 = .73$, and session two, $F(1, 6) = 38.88$, $p = .001$, $\eta^2 = .87$ (see Figure 4). The accuracy of all participants followed this pattern in both depth and azimuth.

Additionally, there was a practice effect from the first to second unimodal session, but only in the auditory modality. Auditory depth localization in session two was more accurate than auditory depth localization in session one, $F(1, 6) = 13.42$, $p = .01$, $\eta^2 = .69$. This practice effect did not quite reach significance for auditory localizations made in azimuth, $F(1, 6) = 5.43$, $p = .06$, $\eta^2 = .48$. Due to the effect of practice on auditory localizations, results of both unimodal sessions will be used for all further analyses and comparisons involving the unimodal localizations.

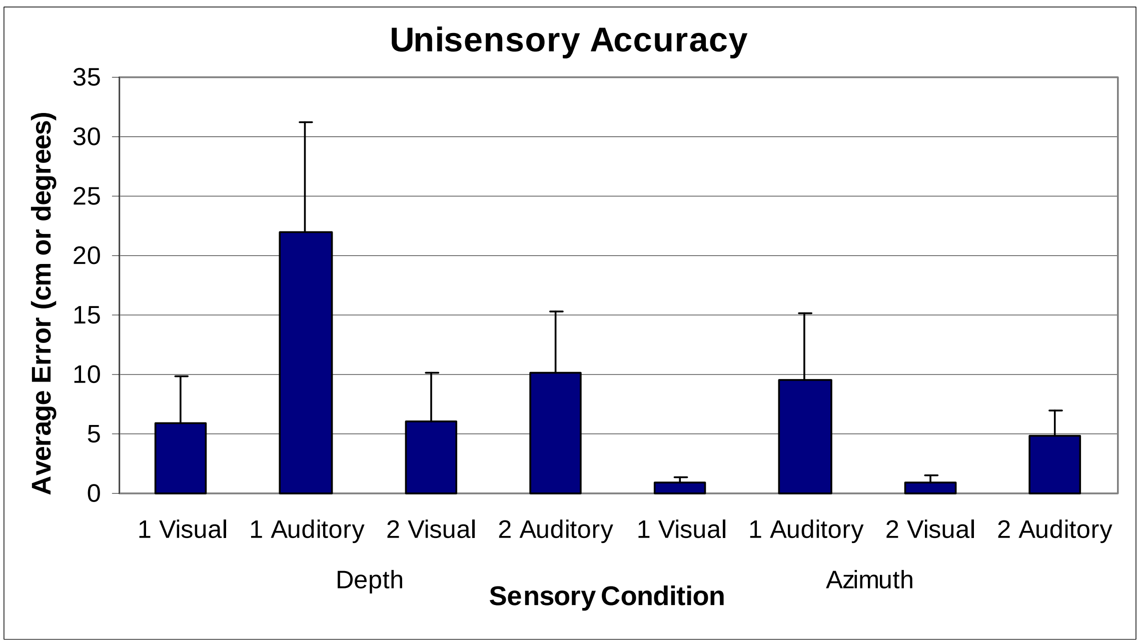

Figure 4. Unisensory Accuracy: Error, in cm or degrees and plotted on the y-axis, is the amount that participants' localizations of targets deviated from the actual location of the targets. Sensory conditions for unimodal session one and unimodal session two are plotted in depth and azimuth on the x-axis. Error bars are standard deviations.

Variability

Variability, measured as the standard deviation of accuracy, was smaller for visual depth localization in the first unimodal session ($M = 5.84$, $SD = 3.20$) than auditory depth localization ($M = 14.38$, $SD = 4.03$), $F(1, 6) = 21.04$, $p = .004$, $\eta^2 = .78$. This finding was also true for visual ($M = 5.49$, $SD = 3.88$) and auditory localization ($M = 8.41$, *SD* = 8.41) in depth in the second unimodal session, $F(1, 6) = 14.60$, $p = .009$, $\eta^2 = .71$. Like many other studies done in azimuth localizations, the variability of visual localization ($M = 1.65$, $SD = 1.83$, session one; $M = 1.40$, $SD = 1.34$, session two) was smaller than that of auditory localization ($M = 12.15$, $SD = 7.11$, session one; $M = 5.47$, $SD = 4.23$, session two) for unimodal session one, $F(1, 6) = 18.51$, $p = .005$, $\eta^2 = .76$, and session two, $F(1, 6) = 12.88$, $p = .01$, $\eta^2 = .68$ (see Figure 5). The variability of all participants' localizations followed this pattern in depth and azimuth. Additionally, the variability decreased from unimodal session one to unimodal session two, but only for auditory localizations in depth, $F(1, 6) = 19.55$, $p = .004$, $\eta^2 = .77$, and in azimuth, $F(1, 6) = 6.94$, $p = .04$, $\eta^2 = .54$.

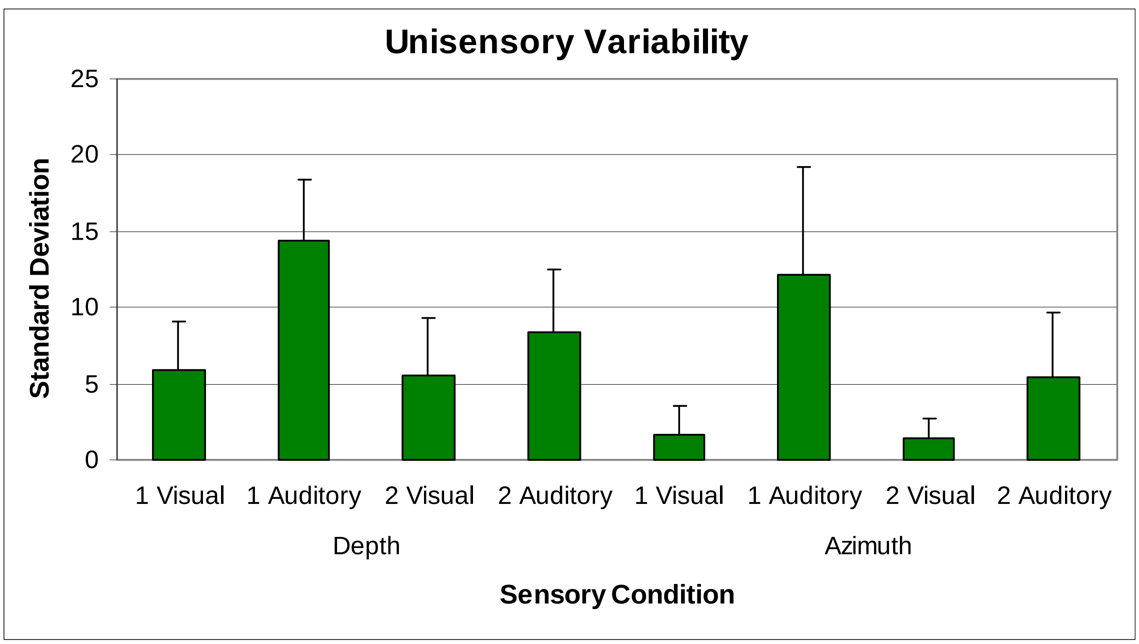

Figure 5. Unisensory Variability: Variability is the standard deviation of the error scores and plotted on the y-axis. Sensory conditions for unimodal session one and unimodal session two are plotted in depth and azimuth on the x-axis. Error bars are standard deviations.

Cross-modal Trials

In cross-modal stimulus conditions, participants always localized the auditory target in the presence of a distracting (irrelevant) visual signal originating from either the same or a different location as the auditory target. Response time, accuracy, variability, and visual bias of localizations were computed. The relationship between multisensory visual bias and visual localization accuracy, as well as the relationship between multisensory visual bias and the ratio of auditory to visual localization variance in unimodal trials, were also examined.

Response Time

Participants did not differ in response time on trials they perceived as being unified (*M* = 2218.01, *SD* = 658.26) versus nonunified (*M* = 2372.01, *SD* = 883.54), $F(1, 6) = .91$, $p = .38$, $\eta^2 = .13$. However, participants were faster at making responses in cross-modal conditions than in auditory only or visual only conditions, but only for unimodal session one. In this session, participants were slower to respond to auditory-only information than in cross-modal trials they perceived as being unified, $F(1, 6) = 40.58$, $p = .001$, $\eta^2 = .87$, and those they perceived as being nonunified, $F(1, 6) = 10.97$, $p = .02$, $\eta^2 = .65$. In unimodal session one, participants were also slower to respond to visual-only trials than to cross-modal trials, but only to ones they perceived as unified, $F(1, 6) = 11.38$, $p = .02$, $\eta^2 = .66$ (see Figure 3). However, it seems that participants are equally fast at responding to unimodal trials as cross-modal trials by session two, after completion of the eight cross-modal sessions, suggesting there is a practice effect in regards to response time. This concurs with the unimodal finding that visual and auditory stimuli were responded to more quickly from the first to second unimodal session.

Accuracy

Participants made more accurate auditory localizations in the presence of distracting visual stimuli in azimuth on all trials that they perceived as being unified (*M* = 2.98, *SD* = 1.17) than on ones they perceived as being nonunified (*M* = 5.03, *SD* = 1.47), $F(1, 6) = 14.75$, $p = .009$, $\eta^2 = .71$. This was also true for all participants' auditory localizations in depth that were perceived as unified (*M* = 8.39, *SD* = 5.39) versus those that were perceived as nonunified (*M* = 9.65, *SD* = 4.90), $F(1, 6) = 8.94$, $p = .02$, $\eta^2 = .60$.

Cross-modal trials were either spatially coincident or noncoincident and could be perceived as unifed or nonunified. This created four conditions that were entered into a repeated measures analysis of variance, separately for azimuth and for depth. The overall analysis revealed there was a difference among the accuracy of the four multimodal conditions in azimuth, $F(3, 18) = 15.61$, $p = .01$, $\eta^2 = .92$, and in depth, $F(3, 18) = 11.54$, $p = .02$, $\eta^2 = .90$. A within-subjects contrast revealed that localizations made in the spatially coincident and perceived unified condition (*M* = 1.35, *SD* = 1.40) were more accurate than all other conditions in azimuth, $F(1, 6) = 46.73$, $p < .001$, $\eta^2 = .89$, and the same was true for localizations made in coincident and perceived unified trials (*M* = 6.98, *SD* = 5.80) in depth, $F(1, 6) = 25.47$, $p = .002$, $\eta^2 = .81$ (see Figures 6 & 7). Unexpectedly, participants' accuracy of azimuth localizations made in coincident and perceived unified trials were not different than their accuracy of visual-only trials in unimodal session one or session two. This finding was the same for the depth dimension.

Variability

Variability was lower for auditory localizations of perceived unified (*M* = 4.44, *SD* = 1.19) than perceived nonunified stimuli (*M* = 5.14, *SD* = 1.24) in azimuth,

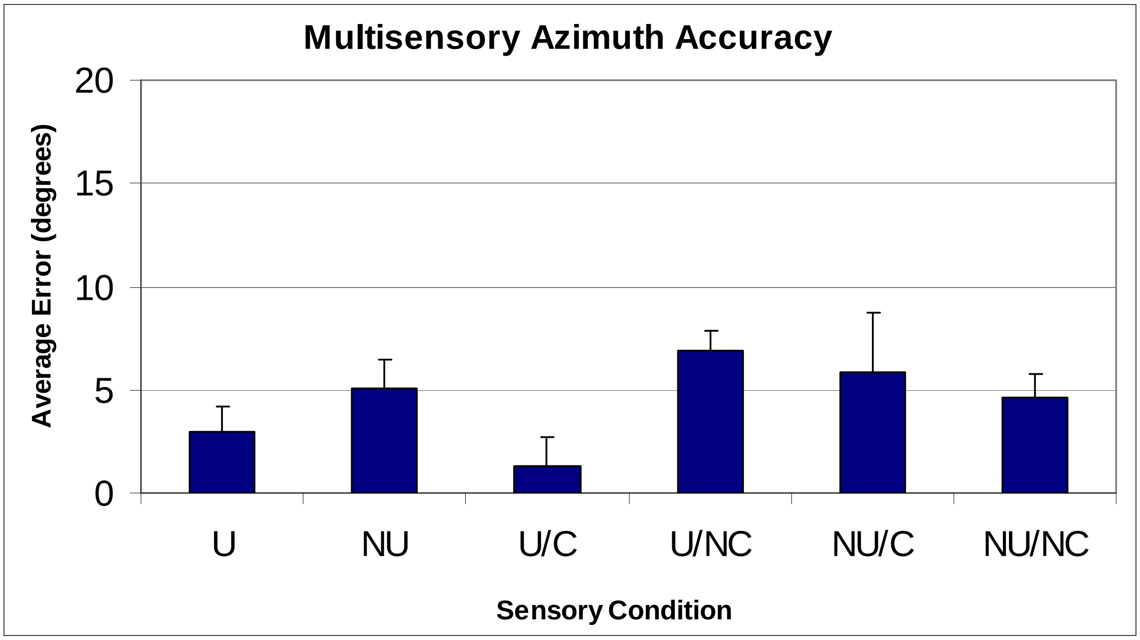

Figure 6. Multisensory Azimuth Accuracy: Error, in degrees and plotted on the y-axis, is the amount that participants' auditory localizations deviated from the actual location of the auditory targets. Sensory conditions including overall perceived unified and nonunified, as well as the four conditions in azimuth, are plotted on the x-axis. Error bars are standard deviations.

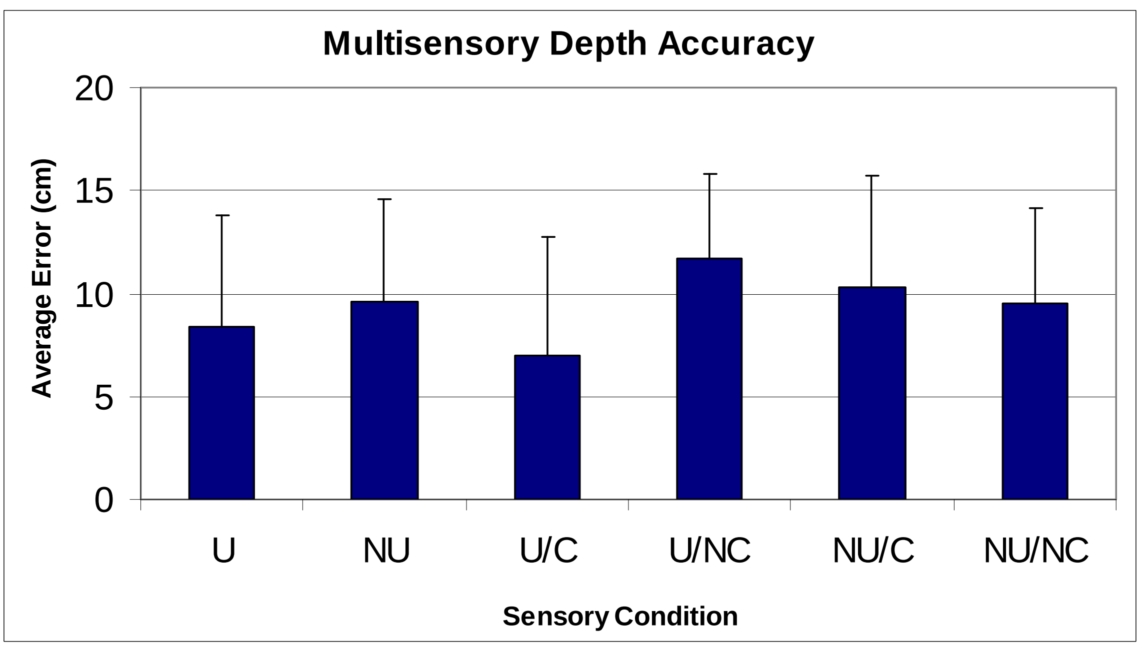

Figure 7. Multisensory Depth Accuracy: Error, in cm and plotted on the y-axis, is the amount that participants' auditory localizations deviated from the actual location of the auditory targets. Sensory conditions including overall perceived unified and nonunified, as well as the four conditions in azimuth, are plotted on the x-axis. Error bars are standard deviations.

$F(1, 6) = 6.49$, $p = .04$, $\eta^2 = .52$, but not between perceived unified ($M = 7.60$, $SD = 3.37$) and perceived nonunified stimuli ($M = 8.13$, $SD = 2.99$) in depth, $F(1, 6) = 3.39$, $p = .12$, $\eta^2 = .36$. Again, an overall analysis of the four cross-modal conditions revealed a difference among variability in azimuth, $F(3, 18) = 20.21$, $p = .007$, $\eta^2 = .94$, but not in depth, $F(3, 18) = 5.03$, $p = .08$, $\eta^2 = .79$.

Similar to accuracy, a within-subjects contrast revealed that localizations made in the spatially coincident and perceived unified condition ($M = 1.77$, $SD = 1.87$) were less variable than all other conditions in azimuth, $F(1, 6) = 54.21$, $p < .001$, $\eta^2 = .90$. And, despite the overall analysis not quite reaching statistical significance, the contrast revealed that localizations made in coincident and perceived unified trials in depth ($M = 6.98$, $SD = 5.80$) were the least variable compared to those of all other conditions, $F(1, 6) = 10.67$, $p < .001$, $\eta^2 = .64$ (see Figures 8 & 9). Additionally, while it was hypothesized that variability would be lower for localizations made on coincident and perceived unified trials than for the best unimodal condition in azimuth and depth, this did not turn out to be the case. This was likely because the variability for both visual and coincident/unified localizations were both small.

Visual Bias

Bias was computed using the following formula:

$$\text{Bias} = [C - N] / \text{Actual Disparity} \times 100$$

Where C is the localization in coincident trials, N the localization in noncoincident trials, and actual disparity is the physical disparity between the light and sound. Bias was computed for depth by holding azimuth constant, meaning, for example, that if an auditory signal was located at position 1, the bias in depth towards a visual signal coming

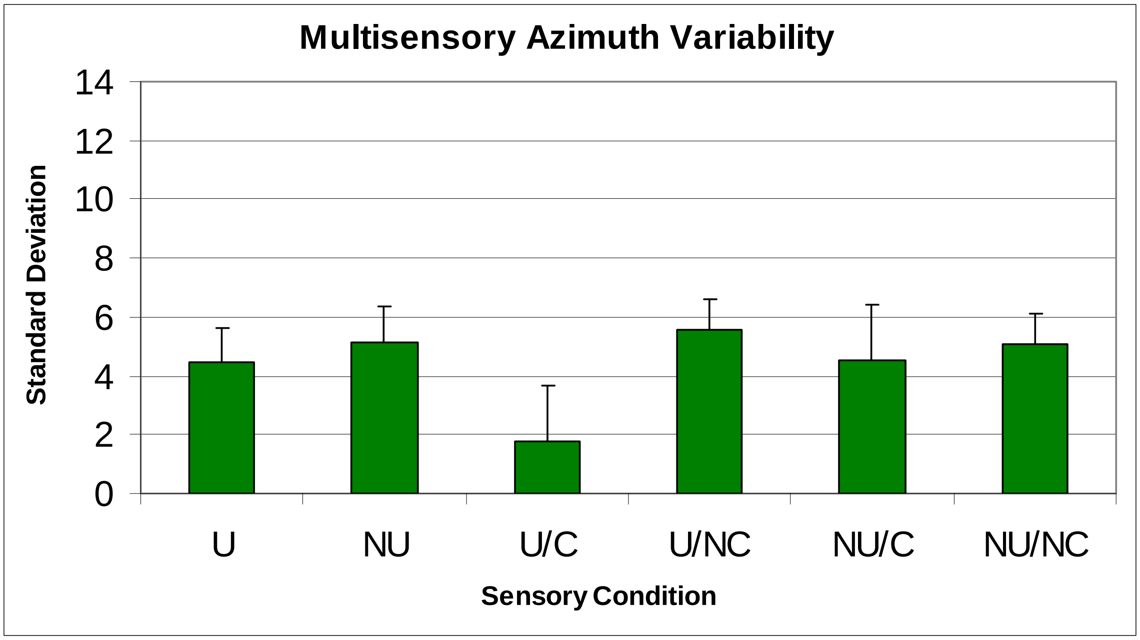

Figure 8. Multisensory Azimuth Variability: Variability was the standard deviation of the azimuth error scores and is plotted on the y-axis. Sensory conditions including overall perceived unified and nonunified, as well as the four conditions in azimuth, are plotted on the x-axis. Error bars are standard deviations.

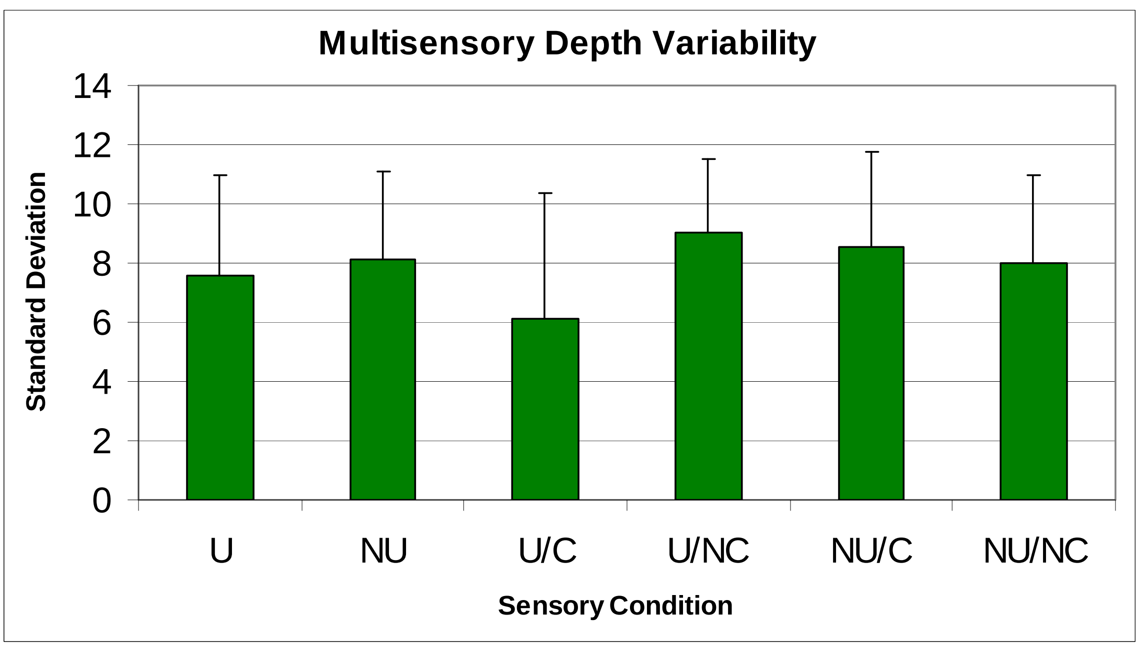

Figure 9. Multisensory Depth Variability: Variability was the standard deviation of the depth error scores and is plotted on the y-axis. Sensory conditions including overall perceived unified and nonunified, as well as the four conditions in depth, are plotted on the x-axis. Error bars are standard deviations.

from position 4 and position 7 were computed, as position 1, 4, and 7 are all located 8.78° from midline. The same was true for bias in azimuth; depth was held constant such that if an auditory signal was originating from position 1, bias in azimuth to visual signals located at position 2 and position 3 were computed, as positions 1, 2, and 3, were all located approximately 83 cm from the participant.

It was expected that participants would bias their auditory localization toward the visual stimuli in trials that they perceived as being unified, regardless of whether the two signals were actually coincident or noncoincident. This effect, which is robust in the literature and found in a previous study by Bowen (2009), was demonstrated in azimuth, where bias in perceived unified trials (M = 84.14%, SD = 22.99) was greater than bias in perceived nonunified trials (M = 1.40%, SD = 19.99), $F(1, 6) = 72.21$, $p < .001$, $\eta^2 = .92$. Bias in perceived unified trials did not differ from 100%, $F(1, 6) = 3.33$, $p = .12$, $\eta^2 = .36$, and in perceived nonunified trials, from 0% bias, $F(1, 6) = .03$, $p = .86$, $\eta^2 = .01$.

There was also a large amount of bias in depth in perceived unified trials (M = 72.19%, SD = 33.44) compared to perceived nonunified trials (M = -3.52%, SD = 6.69), $F(1, 6) = 28.32$, $p = .002$, $\eta^2 = .83$. Unexpectedly, bias in depth was large in perceived unified trials, not quite differing from 100% bias, $F(1, 6) = 4.84$, $p = .07$, $\eta^2 = .45$, but this was approaching significance. Bias in perceived nonunified trials did not differ from 0% bias, $F(1, 6) = 1.95$, $p = .21$, $\eta^2 = .25$. Additionally, bias in unified trials did not differ statistically between depth and azimuth, $F(1, 6) = 4.22$, $p = .09$, $\eta^2 = .41$, but this was nearing significance.

Overall bias was broken down into bias of increasing disparity to evaluate whether bias on perceived unified trials differs as disparity between the visual and auditory target increases. There were two disparities in azimuth, 8.78° and 17.56°, and two disparities in depth, 14 cm and 28 cm. Bias at the smallest disparity in depth (M = 78.04%, SD = 35.63) decreased as the space between the two signals doubled (M = 39.58%, SD = 37.16), $F(1,6) = 6.08$, $p = .05$, $\eta^2 = .45$. The effect between the small disparity (M = 86.84%, SD = 21.86) and large disparity (M = 69.61%, SD = 35.45) did not quite reach significance in azimuth, $F(1,6) = 4.53$, $p = .08$, $\eta^2 = .43$ (see Figures 10 & 11).

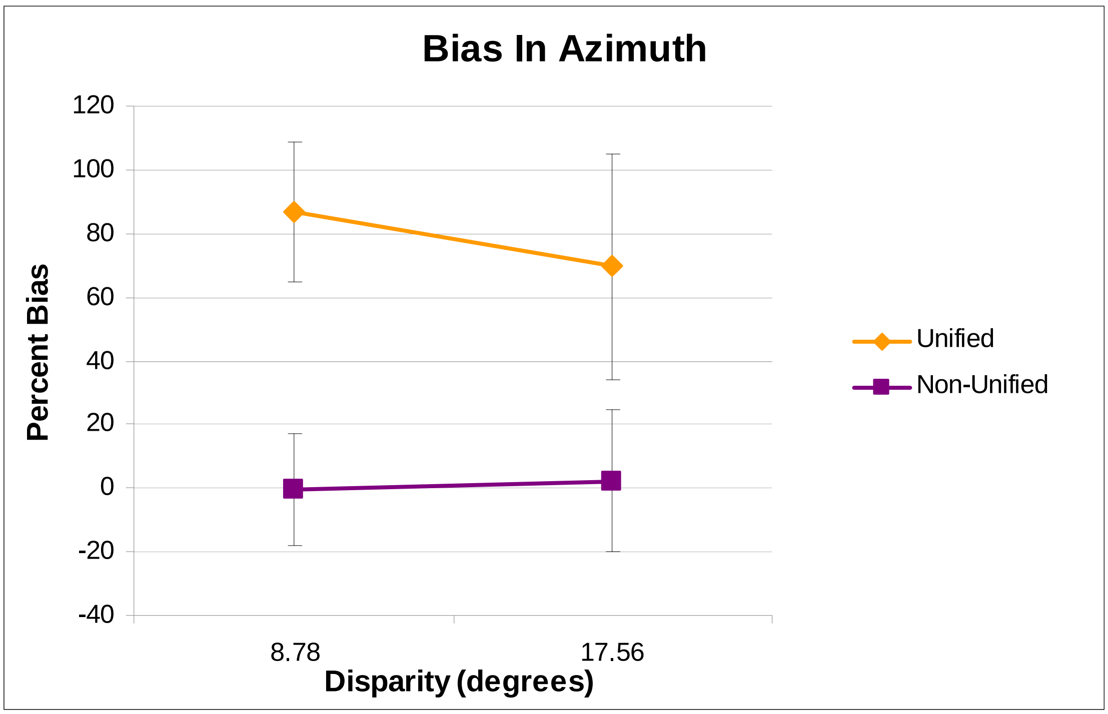

Figure 10. Bias In Azimuth: Bias is plotted on the y-axis and disparity between the visual and auditory target is plotted on the x-axis. Error bars are standard deviations.

Of final interest was how visual accuracy and the ratio of unimodal auditory to visual variance of participants' localizations were related to bias on perceived unified trials. Unfortunately, across participants, bias on unified trials did not positively correlate with the ratio of auditory to visual variance for either unimodal session in azimuth or depth, at anything less than $p = .38$. However, visual error in the first unimodal session

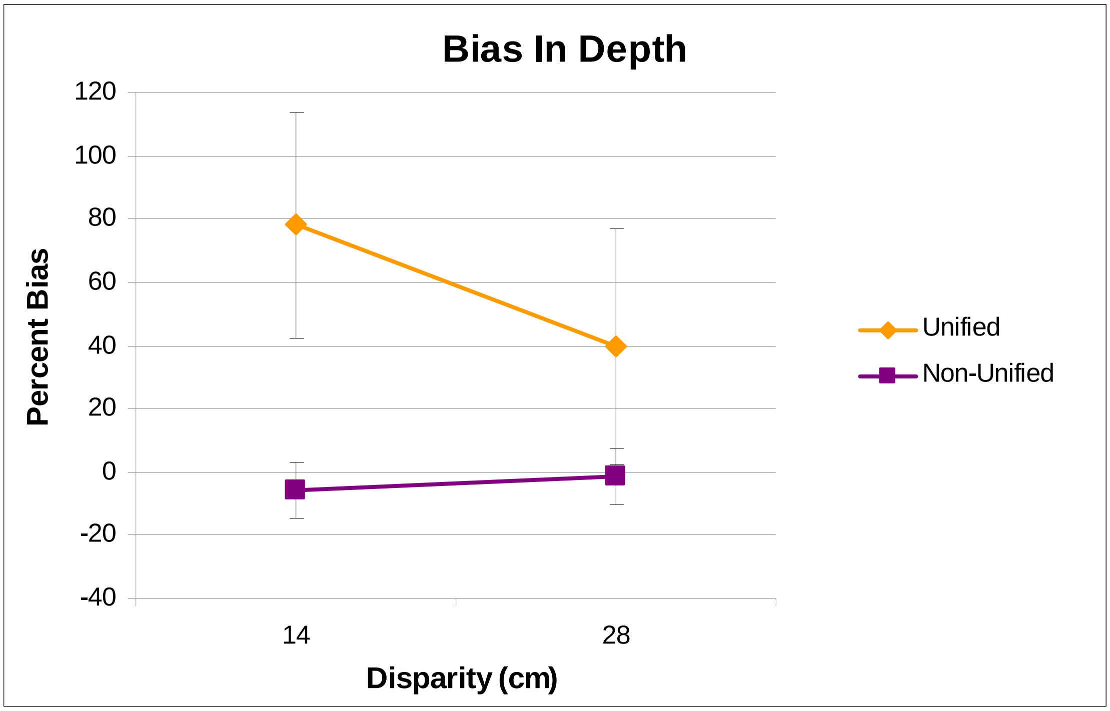

Figure 11. Bias In Depth: Bias is plotted on the y-axis and disparity between the visual and auditory target is plotted on the x-axis. Error bars are standard deviations.

did negatively predict bias on unified trials in both azimuth, $r = -.78$, $n = 7$, $p = .04$, and depth, $r = -.86$, $n = 7$, $p = .01$. This effect disappeared when examining the relation between the visual error from the second unimodal session and bias in azimuth, $r = -.45$, $n = 7$, $p = .32$, and depth, $r = -.61$, $n = 7$, $p = .14$ (see Figures 12 & 13).

After scanning the visual localization data from the first to the second unimodal session, it was evident that some participants' error increased while others decreased. It seems that one explanation for the increase in visual localization error may be due to carlessness from quick responding in the second unimodal session. Therefore, I correlated visual accuracy in unimodal trials with response time. I predicted that unimodal visual accuracy in session two should be negatively correlated with response time. However, these correlations were not significant. Visual accuracy in unimodal session two was not negatively related to unimodal session two response time in azimuth

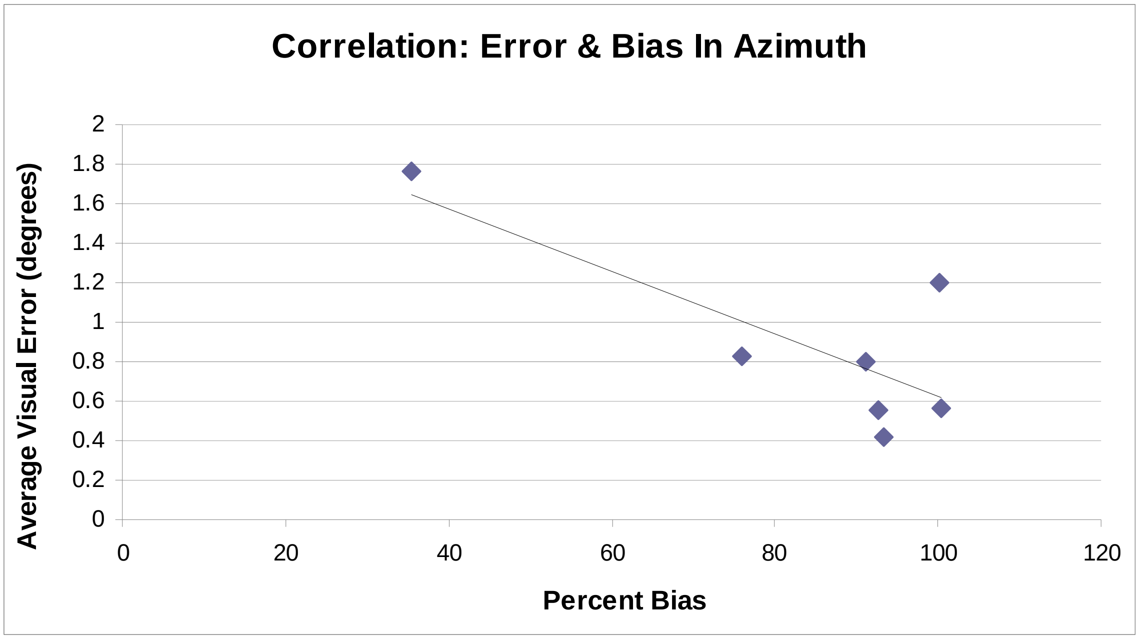

Figure 12. Correlation: Error & Bias In Azimuth: Average azimuth visual error from unimodal session one in degrees is plotted on the y-axis and bias in azimuth is plotted on the y-axis.

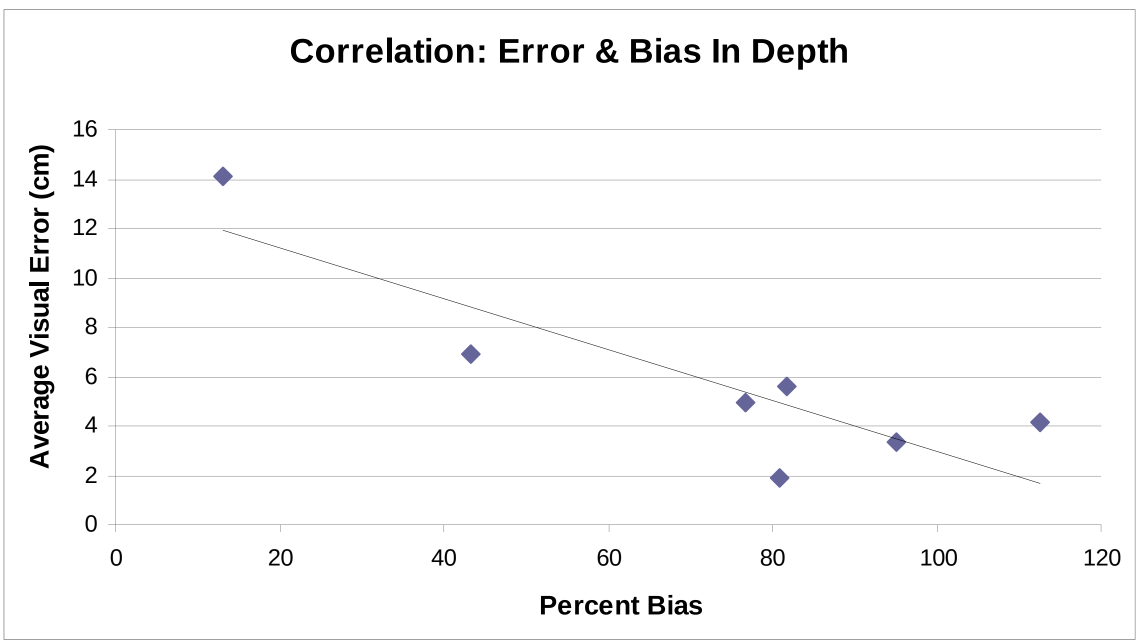

Figure 13. Correlation: Error & Bias In Depth: Average depth visual error from unimodal session one in cm is plotted on the y-axis and bias in depth is plotted on the y-axis.

at r = -.59, n = 7, p = .16, and in depth at r = -.67, n = 7, p = .10. However, these relations were larger than the correlation between unimodal session one visual accuracy and unimodal session one response time in azimuth, $r = -.25$, $n = 7$, $p = .59$, and in depth at $r = -.42$, $n = 7$, $p = .35$. Additionally, the visual average response time for the two unimodal sessions ($M = 2619.66$, $SD = 246.94$) was significantly and negatively correlated with the average response time for the average of the accuracy scores of the two unimodal sessions in azimuth ($M = .89$, $SD = .42$), $r = -.77$, $n = 7$, $p = .05$, and in depth ($M = 5.93$ $SD = 3.94$), $r = -.87$, $n = 7$, $p = .01$. This suggests that overall, response time is negatively correlated to the unimodal sessions, but only for visual localizations (see Figures 14 & 15), and this is more true for the second unimodal session, after practice has taken place.

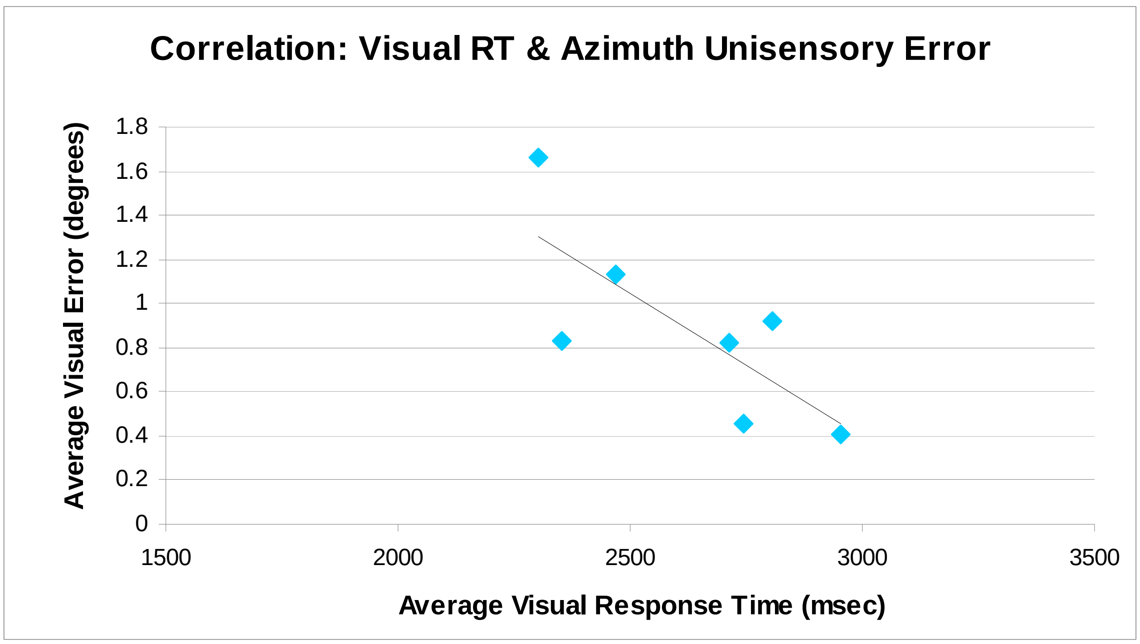

Figure 14. Correlation: Visual RT & Azimuth Unisensory Error: Average azimuth visual error from both unimodal sessions in degrees is plotted on the y-axis and average response time to visual signals from both unimodal sessions in msec is plotted on the x-axis.

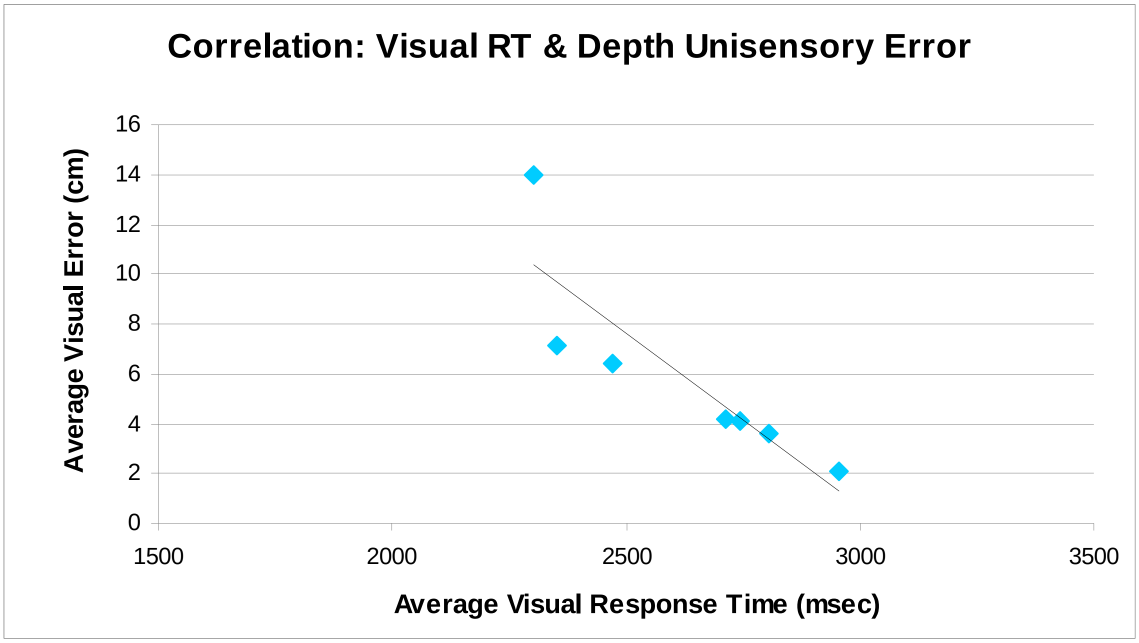

Figure15. Correlation: Visual RT & Depth Unisensory Error: Average depth visual error from both unimodal sessions in cm is plotted on the y-axis and average response time to visual signals from both unimodal sessions in msec is plotted on the x-axis.

DISCUSSION

The current experiment has expanded upon the field of multisensory integration as many goals in the current study were accomplished. While prior studies have shown that the discriminability of visual and, particularly, auditory signals in depth is poor (Brungart, Durlach, & Rabinowitz, 1999; Mershon & Bowers, 1979; Viguier, Clement, & Trotter, 2001; Zahorik, 2002a), the current experimental setup uniquely allowed for the examination of visual bias in depth. Another gain from the design of this project was knowledge of how to create visual and auditory signals that are distinguishable not only in azimuth, but also in depth. Finally, the negative relationship between visual unisensory error and bias on perceived unified trials in the current study presents an insight into understanding a potential mechanism behind visual bias. A summary and explanation of the findings as well as the limitations and future directions of this research project follows.

Unisensory Localization

Participants completed two unimodal sessions, one at the start of the experimental trials and a second after the eight multimodal sessions. This provides an evaluation of the possible existence of a practice effect in terms of localization performance and what this may mean for visual bias. Response time, which was not expected to differ between visual and auditory localizations, did differ at unimodal session two. Visual responses were almost significantly faster than auditory responses by session two, and this may have been driven by the finding that localization response time to visual targets decreased significantly from the first to second session. Interestingly, only auditory localization became more accurate and less variable from session one to session two, and this was most evident for depth localization. This practice effect may be due to auditory localization, particularly in depth, having the most room for improvement throughout the duration of the experiment. Even in the azimuth dimension, auditory localization was not without a high degree of error; mean error of auditory localizations in azimuth during unimodal session one was around the actual disparity between azimuth targets (i.e., around 9°). Because visual signals were already localized with a large degree of accuracy, participants did not become “better” at responding to them, but they did become faster, from session one to session two.

Due to the practice effect found from unimodal session one to session two, it was of interest to examine the frequency of the visual and auditory localization responses made at each location in azimuth in depth. The thought was that participants may be learning the existence of the nine separate locations from unimodal session one to session two and may begin to localize in categories; when participants miss one of the locations, they may localize at another location that they have learned, rather than localizing near the actual location. If this were true, then one would expect that localization responses would not follow a normal distribution at each location by unimodal session two. For unimodal session one in azimuth, participants localized with the highest frequency around the actual location and their localization responses clustered around the actual location, following a normal distribution. This was true for auditory responses as well as visual responses. However, in the first unimodal session in depth, the localization frequency did not always seem to be normally distributed around each of the nine locations, but this was only true for auditory localizations. In particular, individuals seemed to pull auditory signals closer to them than they actually were (i.e., around 20 cm), especially when the sounds were coming from the front depth arc closest to the participant. The visual localizations in depth were normally distributed, much like the localizations made in azimuth.

By unimodal session two, response localizations did appear to become less normally distributed, but only for auditory localizations. From unimodal session one to session two, auditory localizations in azimuth had a tendency to be displaced to other locations in azimuth, particularly towards the center, at 0° when the sound originated from approximately -8° to the left or 8° to the right. Auditory localization in depth did not follow a normal distribution in unimodal session one, and this did not seem to change much by session two, other than participants did not seem to pull the auditory signal as close to them as was shown in session one. The visual localizations remained fairly normally distributed from session one to session two. These findings seem to indicate that participants are learning the locations of the lights and sounds through practice. In doing so, even though participants become more accurate, they had a tendency to categorize their auditory localizations such that, when they did not hit one location, they tended to hit another location.

An explanation for these findings may be suggestive of a number of perceptual phenomena. It seems to argue most for the fallibility of auditory localization, particularly in depth. Participants are so poor at determining the exact location of auditory signals that they must rely on something other then their perception to make localizations, such as the memory of where signals, likely visual ones, have originated from in previous cross-modal sessions. In the depth dimension, participants seemed very sloppy in their auditory localizations in session one, and while their localizations became more accurate by session two, they began categorizing their responses. This is further supported by the fact that participants' localizations of auditory signals in the azimuth dimension became more flat and less normally distributed from session one to session two, indicating that participants are relying upon these learned positions. In line with this explanation is that the practice effect is due to a learning mechanism, i.e., reinforcement. On cross-modal trials that participants perceive as being unified, they are mostly localizing the auditory stimulus at the position of the visual stimulus. Because the two stimuli are occurring together, either in actuality or in the participants' perception, the participants are being reinforced by localizing the signal with what they likely feel is a high degree of accuracy (i.e., around the visual signal).

One of the most important findings of the current study was the degree of accuracy and variability of the visual signals in the depth dimension. Unlike a previous study (Bowen, 2009), visual localization accuracy and variability was significantly superior to the accuracy of auditory localization in depth, and all participants seemed to follow this pattern. This finding verified that the measures that were taken to increase the saliency of the visual signals were successful. Specifically, by creating multiple fixation planes, using signals with blurred, rather than sharp, edges, and by increasing the duration of the signal to 1000 ms, visual signals were able to be localized with a good deal of accuracy in depth. Additionally, using a longer duration auditory signal and a speaker that equally represented all frequencies likely aided in auditory localization in depth as well. While these measures were likely necessary to produce visual bias in depth, it is also interesting in and of itself to find a fair degree of localization accuracy with low variability in the depth dimension. Prior studies have shown how poor localization performance is in this dimension (Brungart, Durlach, & Rabinowitz, 1999; Mershon & Bowers, 1979; Viguier, Clement, & Trotter, 2001; Zahorik, 2002a), and the current study demonstrates several ways to lower the discrimination threshold for signals in depth, particularly in the visual modality.

Multisensory Localization

Response time followed an expected pattern; participants were faster at responding to cross-modal trials than to unimodal session one trials, and this was true for both perceived unified and nonunified trials in comparison to auditory signals and for unified trials in comparison to visual signals. This finding dovetails nicely with the finding from the unimodal response time evaluation; unimodal stimuli, particularly visual ones, elicited responses with shorter reaction times from session one to session two. Participants did not make faster localization responses on perceived unified versus nonunified trials, indicating that any time participants are presented with two signals, regardless of the spatial coincidence of the signals, they are able to make localization responses more quickly than if they are given only one signal. However, this finding seems to disappear with localization practice.

Localization performance in azimuth and depth was more accurate for perceived unified trials than for perceived nonunified trials, and this effect seems to be driven by localizations made on trials that are both perceived as unified and actually spatially coincident. A contrast test comparing the perceived unified and coincident condition to the three remaining multimodal conditions revealed this to be the case; participants made the most accurate localizations in azimuth and depth on trials they perceived to be unified and that were indeed coincident. The variability of the error scores of the participants followed a similar pattern; participants were less variable in their localizations of perceived unified versus nonunified trials, but this was only true for the azimuth dimension. Again, an overall analysis revealed that localization variability in all four conditions of the cross-modal sessions was different in azimuth, but this did not quite reach significance in depth. However, in both depth and azimuth, variability of the perceived unified and coincident trial was lower than the variability of all other conditions. These findings, particularly the evaluation of performance error, suggest that participants' localization of auditory signals in the presence of visual distracters follow a similar pattern in both azimuth and depth. This is further evidence that localization accuracy on unimodal trials, particularly visual trials, may be important for the perceptual process of visual bias to take effect.

Of particular importance in the evaluation of the unisensory accuracy and variability was its comparison to multisensory accuracy and variability. According to a prediction by the MLE model, bimodal variance is typically smaller than the variance of the best unisensory response. For this reason, it was predicted that the error and variability of the localizations made on trials perceived as unified and spatially coincident would be lower than those in visual trials. However, this turned out not to be the case. Moreover, the mean accuracy and variability of the cross-modal and unimodal conditions were not in the expected direction. The fact that visual-only accuracy was so tight, with the mean error of the two sessions at approximately 6 cm in depth and approximately 0.9° in azimuth, could have accounted for this finding. Perhaps the saliency of the visual signal was increased to such a large extent that, during multimodal sessions, there was not an opportunity for the auditory stimulus to influence localization performance.

However, without using measures to increase the saliency of the visual and auditory signals, visual bias may not have occurred in the depth dimension.

Visual Bias

Visual bias was examined in terms of perceived unification. In order to evaluate the results by comparing bias on perceived unified to nonunified trials, it was of interest to examine the frequency histograms of unification judgments among the differing disparities in azimuth and depth, as well as among participants. In both azimuth and depth, participants seemed to be more likely to indicate “nonunified” when the targets were spatially distal rather than proximal, i.e., when targets were 17° versus 8° apart in azimuth or 14 cm versus 28 cm apart in depth. When the disparity was small, i.e., 8° or 14 cm apart, it seemed as if participants seemed to be indicating “unified” and “nonunified” fairly equally, or perhaps “unified” slightly more than “nonunified.” Interestingly, there seemed to be a consistent pattern among unification judgments in azimuth in that when the sound was coming from the left of the light, and when the distance between the signals was 8°, participants seemed more likely to indicate “unified”. However, I have no explanation for this unique perceptual finding.

Additionally, as expected, participants seemed more likely to indicate unified when the light and sound were actually spatially coincident and nonunified when they were not spatially coincident, according to the frequency histograms. By examining the frequency of responses of unification judgments on a participant level, it seemed that participants were, overall, more likely to respond “unified” than “nonunified” or to indicate each relatively equally. This falls in-line with the finding that, overall, the frequency of “unified” responses seems higher than “nonunified”.

One of the main goals of the current study was to enable participants to discriminate between signals in depth in order to be able to respond to stimuli at three different depths. Indeed, bias of auditory targets toward visual stimuli at different depth disparities reached about 72% on perceived unified trials and was slightly negative (e.g., -3%) on perceived nonunified trials. Interestingly, bias on perceived unified trials did not differ significantly from 100%, i.e., perfect bias, and perceived nonunified bias did not differ from no bias, or 0% bias. A similar pattern was found in azimuth, with bias on perceived unified at 84% and nonunified at 1%, which replicates many findings on visual localization in azimuth (Bertelson & Aschersleben, 1998; Bertelson & Radeau, 1981; Bowen, 2009; Hairston et al., 2003; Wallace et al., 2004; Warren, 1979; Warren, Welch, and McCarthy, 1981). However, it did not replicate the finding by Wallace et al. (2004) that participants negatively bias on trials they perceive as nonunified. Because visual bias was so high, and did not differ from 100% bias on perceived unified trials, these findings may be a result of visual capture or domination. When participants indicate unity, it may be that they are localizing the light and not localizing the auditory signal in the presence of the light. Additionally, it is unknown why visual bias was so high in the depth dimension. It may be that visual accuracy in both dimensions is so incredibly precise that it dominates the auditory signal, however, it is not necessary for the visual error to be as small as it is in azimuth for visual capture to occur. Overall, visual bias can be extended from the azimuth dimension to the depth dimension, demonstrating a unique finding in this study in comparison to the previous literature on multisensory depth localization.

Another facet of the current study was to examine the difference between bias when the auditory and visual signals were spatially proximal or distal. Previous studies have found that visual bias is prominent when signals are proximal. However, beyond an upper threshold disparity, participants no longer unify them, and, thus, do not bias. Given the current experimental setup, only two disparities were present in azimuth, 8.78° and 17.56°, and depth, 14 cm and 28 cm. A comparison between bias at 8.78° and at 17.56° on perceived unified trials in azimuth did not reveal a significant reduction in bias, but did approach significance. However, participants did bias more on trials when the signals were 14 cm versus 28 cm apart in depth. If participants were given greater stimulus disparities, these differences, both of which were small, would be expected to increase.

A final goal of the current study was to provide evidence of the MLE model's predictions regarding visual bias. MLE suggests that individuals rely on, or weight, the visual and auditory signals in multimodal localizations to a predetermined extent based on the variance of the unimodal localizations, which is what was found by Alais & Burr, 2004, Battaglia, Jacobs, and Aslin (2003), and, using haptic signals as opposed to auditory signals, by Ernst and Banks (2002). Therefore, one conclusion based on these assumptions is that the ratio of auditory to visual variance in unimodal localizations may be related to visual bias on perceived unified trials. The larger the ratio, i.e., the more variance in the auditory localizations as compared to the visual localizations, the more participants would rely on the visual signals, causing more bias. However, there was not a strong relationship between the ratio of auditory to visual variance and bias on unified trials in either azimuth or depth. It seems it was not the precision of the unimodal responses that led to the weight participants gave to the auditory and visual signals on perceived unified trials, or at least there was no reliable relationship among all participants to this effect.

Another assumption regarding visual bias is that the accuracy of visual responses is what causes participants to perceive unification between the visual and auditory signals, and thus bias to the visual stimuli. Indeed, visual bias was negatively related to visual error in both azimuth and depth from the first unimodal session, but not from the second unimodal session. Visual localization overall was not shown to be different from unimodal sesson one to session two, so the finding that bias did not correlate with visual localization on session two must have been related to the pattern of visual localization on an individual level from session one to two. A scan of the visual localization data from the first to second unimodal sessions in azimuth and depth revealed that some participants' accuracy scores increased while others decreased. Therefore, it seemed that at least accuracy in the first unimodal session may be predictive of visual bias. In order to examine whether the lack of correlation between visual bias and visual accuracy in the second unimodal session was possibly due to fast response time, and, thus, sloppy responding, I examined the relationship between visual accuracy and response time. Response time for unimodal session two was more highly and negatively correlated to visual accuracy in that session than response time from session one was related to visual accuracy in session one, but the correlations were not significant. However, average response time and unimodal visual accuracy across both unimodal sessions was significantly and negatively correlated, suggesting a trade-off between response time and accuracy.

Overall, it is possible that the increased saliency in both the depth and azimuth dimensions enhanced the discriminability of the unimodal stimuli, which caused participants to bias to visual signals in depth. Qualification for this conjecture is the finding that unimodal visual accuracy was predictive of visual bias. Additionally, because visual bias on perceived unified trials did not statistically differ from 100% bias, it seems that the visual signal was strong enough (i.e., able to be accurately localized) to dominate the auditory target. Perhaps when visual accuracy exceeds a certain amount, such as the accuracy of localization in cross-modal trials that are both spatially coincident and perceived as unified, localization does not follow a distribution that would be predicted by unimodal variance, such as the MLE model would predict. Perhaps the distribution of localization responses in the current data can be better explained by using a Bayesian model, which could take into account unimodal variance but also account for the prior information that the visual system is, regardless of all things being equal, inherently relied upon to a greater degree than the auditory system. Strong visual accuracy may have caused the visual system to be perceptually favored in such a way that it no longer matters whether the auditory to visual variance is low or high because the participants are not giving any weight to the auditory signal in multimodal trials, and, thus, demonstrating visual capture.

Conclusions, Limitations, & Future Directions

In conclusion, the unimodal accuracy of the visual signal correlates with the amount of visual bias in perceived unified trials. Because visual error and variability was significantly smaller than that for the auditory modality, participants are placing a greater reliance on the visual signal when making their multimodal localizations, despite being told to localize only the auditory signal. Increasing the saliency of the visual and auditory signals enabled the current experimental design to replicate visual bias findings in azimuth, as well as produce the finding in depth. Not only was this effect demonstrated in depth, a unique finding in itself, but it was large enough to be statistically equal to 100% bias.

The two hypotheses stemming from MLE predictions, that localization on perceived unified and coincident trials would be superior to visual-only localization and that the higher the auditory to visual variance was, the more participants would bias on perceived unified trials, were not supported. This may be due to a general bias of the visual system; participants may rely heavily on the visual signal due to the small error with which it is localized in unimodal trials. In other words, the data may better fit a distribution of responses predicted by a Bayesian theory, which could allow for the incorporation of prior knowledge of statistical information as well as unimodal localization performance.

The exceedingly accurate localization of visual signals, particularly in azimuth, implies that the results of the current study should be interpreted with caution. Because visual bias depends on a number of factors, including spatial and temporal proximity, the perception of unification, as well as the ability for signals from both sensory modalities to influence each other, the data in the current study seem to follow a pattern suggested by visual capture. The increased saliency of the visual signal may suggest that participants are not integrating sensory information from multiple modalities, but instead are disregarding the auditory stimulus and only localizing the visual stimulus on perceived unified trials. As such, it is possible to view the increased saliency of the visual signals in azimuth to be somewhat of a limitation in the current study. Perhaps by degrading the visual stimulus in a future study by shortening the stimuli duration, or using a modulated auditory stimulus to increase its saliency, would allow for multisensory integration rather than visual dominance. Additionally, if the practice effect is not entirely due to reinforcement gained on perceived unified multisensory localizations, then giving participants many unimodal practice sessions to increase their auditory localization accuracy before beginning cross-modal sessions may solve the problem posed by this limitation. On a related note, another limitation of the current study was the fact that, while participants were always instructed to remain fixated on the noius points, there was no way to ensure that participants did not move their eyes to the signals during their presentation. If participants did move their eyes to the stronger stimulus during presentation, i.e., the visual signal, this could explain why participants did not integrate the visual and auditory stimuli and only seemed to localize the visual signal.

A possible problem with the current study was the existence of a participant who was an outlier among all of the data sets. Because only 7 individuals participated in the current study, the analyses for this thesis were conducted with all participants. However, response time, accuracy, variability, and visual bias were reanalyzed after removing the outlying subject's data. Analyses of the response time, accuracy, and variability of the unimodal and cross-modal trials followed the same pattern after the outlier was removed from the data set. However, while visual bias results remained similar, the difference between bias in perceived unified trials in depth between the small (i.e., 14 cm) disparity and large (28 cm) disparity only approached significance, $F(1,5) = 5.52$, $p = .07$, and was no longer seemed to be following a trend in azimuth, $F(1,5) = 2.95$, $p = .15$. More importantly, the ratio of auditory to visual variance in the second unimodal data set, while not significantly related to bias in depth in the original data set, approached significance in the new data set, $r = .77$, n = 6, $p = .07$. Also, unfortunately, visual error in the first unimodal session no longer negatively predicted bias on perceived unified trials in azimuth, $r = .02$, n = 6, $p = .98$, or in depth, $r = -.59$, n = 6, $p = .22$. Additionally, the correlation between visual-only accuracy and response time followed a similar pattern, but response time across the two unimodal sessions no longer significantly predicted the error of the average visual accuracy across the two unimodal sessions in azimuth, $r = -.63$, n = 6, $p = .18$, as it still did in depth, $r = -.98$, n = 6, $p < .001$.

The results after the removal of the outlying participant's data suggest that one should be careful when interpreting the original correlation between visual-only accuracy and bias on perceived unified trials. The large drop in the correlation between the two in azimuth when the outlying subject was removed suggests that this finding may not demonstrate a potential mechanism for visual bias, but may have been driven by the outlying subject. In order to demonstrate that visual bias is, or is not, predicted by visual-only accuracy, more subjects are needed to allow for a wider representation of the population from which we were sampling. Additionally, the correlation between visual error and bias in depth, as well as between the ratio of auditory to visual variance and bias in depth both approached significance after removal of the outlier. This indicates that perhaps when the visual signal is not overwhelming (as it appears to be in azimuth), visual bias may follow a pattern of integration suggested by the MLE model in that they may be weighting the signals in the cross-modal trials based on their unimodal variance and accuracy. This could be better examined in a future experiment that does not involve such a strong and overwhelming visual stimulus.

One possible direction for future examination of visual bias in azimuth and depth would be to replicate the current study in a larger experimental chamber in order to allow for more signal disparity than the 8° to 17° in azimuth and the 14 to 28 cm in depth. In this way, other aspects of visual bias in azimuth could be examined in, and possibly extended to, depth. If the disparity between the visual and auditory signals were large, perhaps individuals would demonstrate a higher reliance on the auditory and visual variances when determining the influence that each signal has on multimodal response localization in perceived unified conditions. Because the visual signal is so strong, participants may not be weighting it to the extent that the auditory to visual variance would predict, according to the MLE model. Additionally, increasing the distance between the visual and auditory signal may allow participants to demonstrate negative visual bias, i.e., the pulling of the auditory signal in the opposite direction of the visual signal. Supposedly, if participants perceive the auditory and visual signal as too far apart, and, thus, as nonunified, they may pull the auditory target even further from the visual signal than it already is, as was found in azimuth by Wallace et al. (2004).

An additional future direction could be to ask participants to make not only auditory localizations in the presence of visual distracters, but to also make visual localizations in the presence of auditory distracters. If participants are biasing approximately 80% of the target disparity in azimuth and approximately 70% of the disparity in depth on perceived unified trials, then it would be interesting to see how the auditory signal may be influencing the localization as well. For instance, participants are always first presented with a multimodal auditory and a visual stimulus, separated by some distance, but they are asked on half of the trials to localize the auditory stimulus, and on the other half to localize the visual stimulus. Of the trials perceived as unified, if participants bias towards the visual signal when localizing the auditory signal 80% of the distance, then they should bias toward the auditory signal when localizing the visual signal 20% of the distance in the opposite direction. A similar experimental design was set up by Warren (1979) and was studied only in azimuth. While Warren (1979) did not demonstrate this sum to 100% bias on trials perceived as unified by the participants, the strong saliency of the visual and auditory stimuli in our experimental design may make the procedure worth re-examining in azimuth and examining for the first time in depth.

Overall, this thesis has shown that when participants perceptually fuse two stimuli that are spatially disparate, they will bias about 80% of the disparity towards visual stimuli in azimuth and at about 70% in depth. This study is unique in that it directly measures and results in visual bias in the depth dimension. Additionally, by increasing the saliency of visual and auditory signals, individuals were able to discriminate between differing visual and auditory signals in depth. Finally, findings from this study have lead to the conclusion that visual accuracy in unimodal trials may be predictive of visual bias in perceived unified multimodal trials, such that the smaller the amount of error visual stimuli are localized with, the more participants will bias auditory targets toward visual signals in both azimuth and depth. This bias seems to begin decreasing as signal disparities increase, particularly in depth, although this finding is not very strong, likely due to the confines of the experimental parameters under which the current study took place.

REFERENCES

Alais, D. ,& Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14,* 257-262.

Anastasio, T., Patton, P., & Belkacem-Boussaid, K. (2000). Using Bayes' rule to model multisensory enhancement in the superior colliculus. *Neural Computation, 12,* 1165-1187.

Ashmead, D., LeRoy, D., & Odom, R. (1990). Perception of the relative distances of nearby sound sources. *Perception & Psychophysics, 47,* 326-331.

Barutchu, A., Danaher, J., Crewther, S., Innes-Brown, H., Shivdasani, M., & Paolini, A. (2010). Audiovisual integration in noise by children and adults. *Journal of Experimental Child Psychology, 105,* 38-50.

Battaglia, P., Jacobs, R., & Aslin, R. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of Optical Society of America, 20(7),* 1391-1396.

Bertelson, P., & Aschersleben, G. (1998) Automatic visual bias of perceived auditory location. *Psychonomic Bulletin & Review, 5,* 482-489.

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics, 29,* 578-584.

Blake, R., & Sekuler, R. (2006). *Perception.* McGraw Hill: New York.

Blohm, G., Khan, A., Ren, L., Schreiber, K., & Crawford, J. (2008). Depth estimation from retinal disparity requires eye and head orientation signals. *Journal of Vision, 8,* 1-23.

Bowen, A. (2009). *Multisensory integration in azimuth and depth.* Unpublished manuscript, Department of Psychology, Wake Forest University, Winston-Salem, North Carolina.

Brungart, D. (1999). Auditory localization of nearby sources. III. Stimulus effects. *Journal of the Acoustical Society of America, 106,* 3589-3602.

Brungart, D., Durlach, N., & Rabinowitz, W. (1999). Auditory localization of nearby sources. II. Localization of a broadband source. *Journal of the Acoustical Society of America, 106,* 1956-1968.

Brungart, D., & Rabinowitz, W. (1999). Auditory localization of nearby sources. Head-related transfer functions. *Journal of the Acoustical Society of America, 106,* 1465-1479.

Brungart, D., Rabinowitz, W., & Durlach, N. (2000). Evaluation of response methods for the localization of nearby objects. *Perception & Psychophysics, 62,* 48-65.

Butler, R., Levy, E., & Neff, W. (1980). Apparent distance of sounds recorded in echoic and anechoic chambers. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 745-750.

Choe, C., Welch, R., Gilford, R., & Juola, J. (1975). The "ventriloquist effect": Visual dominance or response bias? *Perception & Psychophysics, 18,* 55-60.

Coleman, P. (1968). Dual role of frequency spectrum in determination of auditory distance. *The Journal of the Acoustical Society of America, 44,* 631-632.

Ernst, M., & Banks, M. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429-433.

Gardner, M. (1968). Proximity image effect in sound localization. *Journal of the Acoustic Society of America, 43,* 163.

Godfroy, M., Welch, R., Sandor, P., & Roumes, C. (2008). Accuracy and precision of auditory-visual localization in the two-dimensional frontal field: A test of the modality precision model. *Experimental Brain Research.* Manuscript submitted for publication.

Hairston, W., Wallace, M., Vaughan, J., Stein, B., Norris, J., & Schirillo, J. (2003). Visual localization ability influences cross-modal bias. *Journal of Cognitive Neuroscience, 15(1),* 20-29.

Hartmann, W. (1999). How we localize sound. *Physics Today,52,* 24-30.

Jacobs, R. (2008). *Bayesian estimation*. Retrieved March 1, 2010, from University of Rochester, Brain and Cognitive Sciences Computational Cognition Cheat Sheets Web site: http://www.bcs.rochester.edu/people/robbie/jacobslab/cheat_sheets.html

Körding, K., Beierholm, U., Ma, W., Quartz, S., Tenenbaum, J., & Shams, L. (2007). Causal inference in multisensory perception. *PLos ONE, 2(9), e943.doi:10.1371/journal.pone.0000943.*

Levy, E., & Butler, R. (1978). Stimulus factors which influence the perceived externalization of sound presented through headphones. *The Journal of Auditory Research, 18,* 41-50.

McKee, S., Levi, D., & Bowne, S. (1990). The imprecision of stereopsis. *Vision Research, 30,* 1763-1779.

Mershon, D., & Bowers, J. (1979). Absolute and relative cues for the auditory perception of egocentric distance. *Perception, 8,* 311-322.

Mershon, D., Desaulniers, D., Amerson, T., & Kiefer, S. (1980). Visual capture in auditory distance perception: proximity image effect reconsidered. *Journal of Auditory Research, 20,* 129-136.

Plenge, G. (1974). On the differences between localization and lateralization. *Journal of the Acoustic Society of America, 56,* 944-951.

Savioja, L., Huopaniemi, J., Lokki, T., & Väänänen, R. (1990). Creating interactive virtual acoustic environments. *Journal of the Audio Engineering Society, 47(9)*, 675-705.

Schiffman, H. (2000). *Sensation & perception (5th ed.).* John Wiley & Sons, Inc.: New York, NY.

Stein, B., & Meredith, A. (1993). *The merging of the senses*. MIT Press: Cambridge, MA.

Vesa, S., & Lokki, T. (2006, September). Detection of room reflections. *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)*, Montreal, Canada.

Viguier A., Clement, G., & Trotter, Y. (2001). Distance perception within near visual space. *Perception, 30,* 115-124.

Wallace, M., Roberson, G., Hairston, W., Stein, B., Vaughan, J., & Schirillo, J. (2004). Unifying multisensory signals across time and space. *Experimental Brain Research, 158,* 252-258.

Warren, D. (1979). Spatial localization under conflict conditions: Is there a single explanation? *Perception, 8,* 323-337.

Warren, D., Welch, R., & McCarthy, T. (1981). The role of visual-auditory "compellingness" in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception & Psychophysics, 30(5)*, 557-564.

Witten, I., & Knudsen, E. (2005). Why seeing is believing: Merging auditory and visual worlds. *Neuron, 48,* 489-496.

Wightman, F., & Kistler, D. (1989a). Headphone simulation of free-field listening. I: Stimulus synthesis. *Journal of the Acoustic Society of America, 85(2),* 858-867.

Wightman, F., & Kistler, D. (1989b). Headphone simulation of free-field listening. II: Psychophysical validation. *Journal of the Acoustic Society of America, 85(2),* 868-878.

Zahorik, P. (2001). Estimating sound source distance with and without vision. *Optometry and Vision Science, 78,* 270-275.

Zahorik, P. (2002a). Assessing auditory distance perception using virtual acoustics. *Journal of the Acoustical Society of America, 111,* 1832-1846.

Zahorik, P. (2002b). Direct-to-reverberant energy ratio sensitivity. *Journal of Acoustical Society of America, 112,* 2110-2117.

CURRICULUM VITA

Amanda L. Bowen

Current Address
530 Haven Ridge Dr. Apt 11
Winston-Salem, NC 27104

Permanent Address
3680 Labrador Lane
Suffolk, VA 23434

EDUCATION

Wake Forest University, Winston Salem, NC

M.A. in general psychology (2010)
Thesis: "Visual Localization Accuracy Determines the Bias of Auditory Targets in Azimuth and Depth"
Advisor: Dr. Jim Schirillo
GPA: 3.50

Christopher Newport University, Newport News, VA

B.S. in psychology; minor in women and gender studies (2008)
Major GPA: 4.00, Cumulative GPA: 3.96

RESEARCH EXPERIENCE

Research assistant, Wake Forest University, Winston Salem, NC, Aug 2008-Apr 2010

- Collected, analyzed, and presented data for a research project entitled "Multisensory Integration in Two Dimensional Space".
- Developed research project idea, collected data, prepared manuscript for, and defended a thesis project entitled "Visual Localization Precision Determines the Bias of Auditory Targets in Azimuth and Depth".
- Presented research projects at conferences.
- Established proficiency in multivariate statistics and in using the Statistical Package for Social Sciences (SPSS).
- Faculty Advisor: Dr. Jim Schirillo, Ph.D., Professor of Psychology, WFU

Research assistantship, Wake Forest University, Winston Salem, NC, Aug 2008-Apr 2010

- Held position as a research assistant for the WFU psychology department.
- Oversaw mass testing sessions in which questionnaires were administered to over 200 students
- Proctored and graded exams for classes of up to 50 students.
- Reanalyzed data from a developmental psychology thesis project
- Faculty Supervisors: Dr. Julie Baker, Ph.D., Lecturer, WFU; Dr. Christy Buchanan, Ph.D., Professor of Psychology, WFU.

Research assistant, Christopher Newport University, Newport News, VA, Feb 2006-May 2008

- Designed research project and developed questionnaires under supervision of faculty advisor.
- Collected and presented data through presentations and posters at conferences.
- Mentored new research assistants.
- Faculty Advisor: Dr. Jeffrey Gibbons, Ph.D., Associate Professor of Psychology, CNU

TEACHING EXPERIENCE

Teaching assistant, Wake Forest University, Winston Salem, NC, Aug 2009-Apr 2010

- Teaching assistant for PSY 320/620 (Physiological Psychology), Jan 2010-Apr 2010
    - Assisted physiological psychology class of approximately 18 students with laboratory activities and classroom and reading materials.
    - Faculty Supervisor: Dr. Jim Schirillo, Ph.D., Professor of Psychology, WFU
- Teaching assistant for PSY 341 (Research in Developmental Psychology), Aug 2009-Dec 2009
    - Assisted developmental psychology class of approximately 15 students by grading exams, editing students' writing and giving feedback on student presentations.
    - Faculty Supervisor: Dr. Andrew Irwin-Smiler, Ph.D., Visiting Assistant Professor, WFU

Writing associate, Christopher Newport University, Newport News, VA, Aug 2006-May 2007

- Attended training sessions for writing tutors, revised and edited papers for experimental psychology undergraduates, held meetings with students to discuss technical writing, instructed students on American Psychological Association (APA) guidelines for writing research reports
- Faculty Supervisor: Dr. Jeffrey Gibbons, Ph.D., Associate Professor of Psychology, CNU

CONFERENCE PRESENTATIONS

Bowen, A., Muday, J., & Schirillo, J. (2010, February). *Integrating perceptual information in azimuth and depth.* Poster session presented at the annual meeting of North Carolina Cognition Group Meeting, Winston Salem, NC.

Gibbons, J., Rothwell, C., Bowen, A., & Mack, N. (2008, March). *The role of the fading affect bias in addiction*. Poster session presented at the annual meeting of Southeastern Psychological Association, Charlotte, NC.

Bowen, A., & Gibbons, J. (2007, April). *Storytelling, depression, and the fading affect bias.* Paper session presented at the annual meeting of the Paideia Student Research Conference, Newport News, VA.

Bowen, A., & Gibbons, J. (2006, November). *Storytelling, depression, and the fading affect bias.* Poster session presented at the annual meeting of the Sigma Xi Student Research Poster Session, Newport News, VA.

Bowen, A., Alan, K., Gerth, M., & Gibbons, J. (2006, April). *Storytelling behaviors and the fading affect bias.* Paper session presented at the annual meeting of the Paideia Student Research Conference, Newport News, VA.

Bowen, A., Warme, R., & Gibbons, J. (2006, March). *Storytelling behaviors are related to the fading affect bias*. Poster session presented at the annual meeting of the Eastern Psychological Association, Baltimore, MD.