

ARMY RESEARCH LABORATORY



**Factors Affecting Auditory Localization and Situational
Awareness in the Urban Battlefield**

by Angélique A. Scharine and Tomasz R. Letowski

ARL-TR-3474

April 2005

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

DESTRUCTION NOTICE—Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Aberdeen Proving Ground, MD 21005-5425

ARL-TR-3474

April 2005

Factors Affecting Auditory Localization and Situational Awareness in the Urban Battlefield

Angélique A. Scharine and Tomasz R. Letowski
Human Research and Engineering Directorate, ARL

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> <i>OMB No. 0704-0188</i>	
<p>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>				
1. REPORT DATE (DD-MM-YYYY) April 2005		2. REPORT TYPE Final		3. DATES COVERED (From - To) September 2003 to September 2004
4. TITLE AND SUBTITLE Factors Affecting Auditory Localization and Situational Awareness in the Urban Battlefield			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Angélique A. Scharine and Tomasz R. Letowski (both of ARL)			5d. PROJECT NUMBER 6102A74A	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory Human Research and Engineering Directorate Aberdeen Proving Ground, MD 21005-5425			8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-3474	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT Soldiers conducting military operations in an urban terrain require heightened auditory situational awareness because of the complexity of the terrain and highly limited field and range of view. In such situations, people tend to rely more heavily on the sounds they hear and vibrations they feel through the sense of touch. However, the complexity of the urban terrain affects not only vision but also hearing and most notably, the perception of the direction of incoming sound. This report presents a summary of the literature that outlines the acoustic factors affecting a human's ability to localize sound sources in the urban environment. These factors include the acoustic environment of the urban terrain, elements of battleground activities, and limits of human localization capabilities. In addition, the report identifies the areas of research that would clarify localization issues and allow for improvements in training and equipment.				
15. SUBJECT TERMS acoustic measurements; auditory testing facilities				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT SAR	18. NUMBER OF PAGES 60
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified		
			19b. TELEPHONE NUMBER (Include area code) 410-278-5957	

Contents

List of Figures	v
1. Introduction	1
2. Sound Localization Basics	2
2.1 Azimuth.....	3
2.2 Elevation.....	6
2.3 Distance.....	9
2.4 Auditory Localization Capabilities and Limits	10
3. Acoustics of the Urban Environment	11
3.1 Walls and Buildings: Physical Properties of the Environment	14
3.1.1 Reflection and Reverberation.....	14
3.1.2 Sound Path Barriers.....	17
3.1.3 Vibration.....	18
3.2 Battlefield Conditions: Noise-Induced Chaos.....	18
3.2.1 Noise.....	18
3.2.2 Multiple Sound Sources: Acoustic Distractors	19
3.3 Other Factors	20
3.3.1 The Effect of Vision on Auditory Localization.....	20
3.3.2 Moving Sound and Moving Listener.....	22
3.3.3 Localizability of Target Sound Sources	25
4. Research Questions	27
4.1 Localizability of Typical Battle Sounds.....	27
4.2 Effect of Reverberation on Localization	27
4.3 Effect of Echoes and Flutter Echoes on Localization	28
4.4 Localizing Multiple Sounds	28
4.5 Moving Sound and Moving Listeners.....	29
4.6 The Interaction of Auditory Localization With Vision	29
4.7 Auditory Training.....	29
5. Conclusions	30
6. References	31

Appendix A. Localization Accuracy	41
Appendix B. Minimum Audible Angle (MAA)	43
Appendix C. Minimum Audible Movement Angle (MAMA)	45
Appendix D. Signal-to-Noise Ratio Needed for Localization	49
Distribution List	51

List of Figures

Figure 1. Visual example of azimuth and elevation	3
Figure 2. Cone of confusion.....	7
Figure 3. Effect of frequency band on localization in the median plane.	8
Figure 4. Echo effect.....	14

INTENTIONALLY LEFT BLANK

1. Introduction

Military operations in an urban terrain (MOUT) are very difficult to conduct because of the complex terrain features and low reliability of sensory information. Narrow streets, smoke obscuring views, reflected and reverberating sounds, overwhelming burning smells, sudden gusting winds, and flying debris create a very confusing environment. When conducting reconnaissance missions or making movement decisions, Soldiers rely primarily on visual information. However, during MOUT, visual cues are frequently obscured or are completely lacking. In such situations, audition becomes the first source of information about the presence of an enemy and the direction of incoming weapon fire. Even if visual cues are available, audition plays a critical role in human behavior because it is the only directional tele-receptor that operates throughout the full 360-degree range. However, veterans of urban warfare and Soldiers in training report that it is quite difficult to identify the locations of sound sources in an urban environment. For example, during urban fights, Soldiers may hear tanks moving but do not know where they actually are at a given moment. Gunfire sounds reflected multiple times from various walls provide no clues about the directions of incoming fire. This is a serious problem for the attacking and defending forces, especially in modern times when MOUT is increasingly common. Defensive forces have the advantage of concealment; the offensive force must determine the locations of enemy resources, and this requires entry into unknown buildings and territories. However, the defending forces risk being isolated and imprisoned in the same buildings that protect them. Therefore, both attacking and defending Soldiers must maintain situational awareness (SA) at all times.

Since World War II, many systems and devices have been developed with the intent to provide aid to Soldiers conducting urban reconnaissance. Most of these systems are designed with the goal of giving the Soldier knowledge about whether buildings and rooms are occupied before he or she enters them. However, all these systems have a limited range of uses and they are difficult to use during movement. In addition, they augment the cognitive and sensory load, and Soldiers report a preference for natural sensory information. Even with the improved supporting systems, there are numerous situations when the Soldiers are forced to rely solely on their own perceptual skills.

This report discusses the effects of the urban environment on one specific element of auditory perception: auditory localization. Numerous studies demonstrate that the auditory system's ability to localize a sound source is vulnerable to distortion by other factors. During difficult listening conditions created by noise and reverberation, we may still be able to detect or even identify a sound source, but we may not be able to determine its location. Thus, the objective of this report is to describe the acoustical characteristics of the urban environment and examine their possible detrimental effects on auditory localization. This analysis is based on an

examination of a large body of research describing human localization behavior in various laboratory contexts to outline the possible sources of and the severity of error. However, there is an operational gap between laboratory conditions and the very noisy, highly reverberant, and constantly changing urban battlefield environment. Such environments and the human behavior in such environments are the ultimate object of interest in this analysis. Therefore, an integral part of this report is also the discussion of potential research questions, technological advances, and training paradigms that have been identified through literature analysis and contacts with Soldiers. It is hoped that analysis and the subsequent research efforts will improve understanding of human auditory abilities and provide guidelines for improved survivability and effectiveness of fighters conducting operations in the urban setting.

2. Sound Localization Basics

Numerous acoustic cues have been shown to be used for auditory orientation in space. The importance of specific cues depends on the type of environment and the sound sources operating in this environment. Moreover, the listener's auditory capabilities and listening experience affect the degree to which individual cues are used. A clear understanding of human capabilities and the mechanisms by which acoustic signals are altered by an environment is important for prediction of the character and the extent of potential localization errors. Thus, in order to understand the capabilities and limitations of auditory spatial orientation in a specific environment, it is necessary to review the primary auditory cues and the elements of the acoustic environment that affect these cues.

Auditory orientation in space involves estimates of and information about four elements of the acoustic environment:

1. The azimuth at which the specific sound source is situated in the horizontal plane and the angular spread of the sound sources of interest in the horizontal plane (horizontal spread or panorama) (see figure 1),
2. The zenith (elevation) at which the specific sound source is situated in the vertical plane and the angular spread of the sound sources of interest in the vertical plane (vertical spread) (see figure 1),
3. The distance to the specific sound source or the difference in distance between two sound sources situated in the same direction (depth), and
4. The size and the shape of the acoustic environment in which the observer is situated (spaciousness, volume).

The first three elements are the polar coordinates of the sound source in Cartesian space with origin of the space anchored at the listener's location. The fourth element is a global measure of

the extent of space that affects the listener. All together, they provide cues regarding a dynamic relationship between the space, the sound source, and the listener.

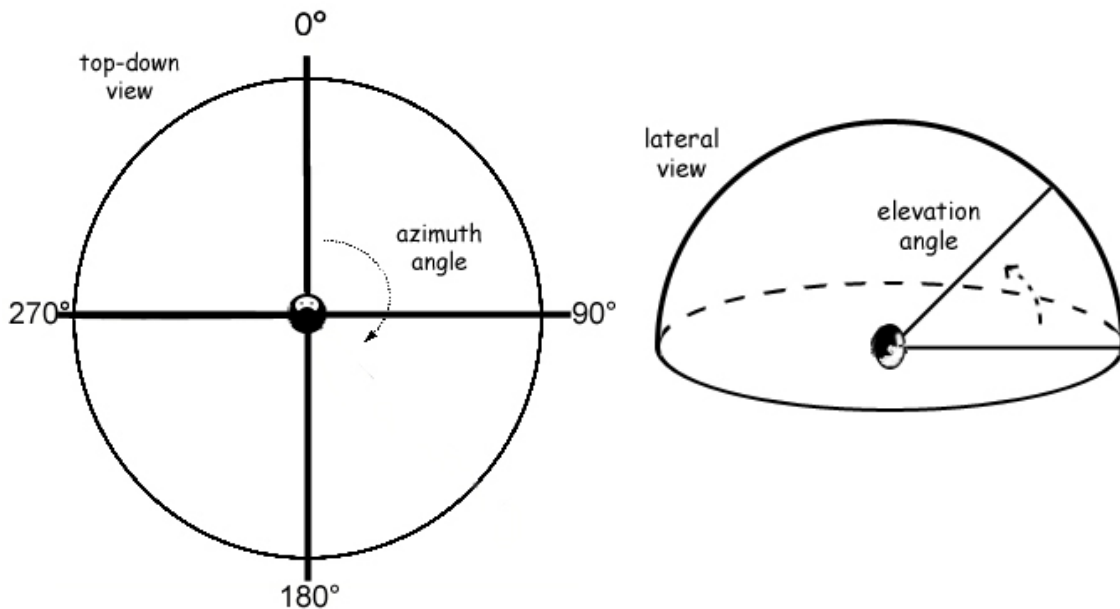


Figure 1. Visual example of azimuth and elevation.

A listener's auditory spatial orientation is based on the differences between sounds entering two ears of the listener (binaural cues), reflections of sounds from the listener's pinnae, head and shoulders (monaural cues), the listener's familiarity with the sound sources and the environment, and dynamic behavior of the sound sources and the listener. The following sections provide information about specific acoustic cues that are used to locate sound sources in azimuth, elevation, and distance. Cues about the size of the acoustic space are not directly related to localization of sound sources but rather to an understanding of the relationship between the environment and the listener when visual cues are not available. They are discussed later in the context of the urban environment. However, it needs to be stressed that the perceived size of the acoustic environment has a direct effect on estimation of the distance from the listener to the sound source when the listener is provided with a frame of reference (distance calibration).

2.1 Azimuth

Sound source localization in the horizontal plane (azimuth) uses binaural (two ears) and monaural (one ear) cues (Blauert, 1999). There are two binaural cues: (a) interaural level differences (ILD) referred to also as interaural intensity differences (IID), and (b) interaural time differences (ITD) or interaural phase difference (IPD). The terms ILD and IID have the same connotation and can be used interchangeably, but there is a slight difference in meaning between ITD and IPD. This difference is described later.

Sound arriving at the two ears of the listener from a sound source situated at a specific azimuth is more intense in the proximal ear than in the distal ear because of the “baffling effect” of the head casting an “acoustic shadow” on the distal ear. At low frequencies, the dimensions of the human head are small in comparison to the wavelength of the sound wave; the difference in sound intensity between two ears is small because of sound diffraction around the head. At high frequencies, the intensity differences caused by the dimensions of the human head are sufficient to provide clear localization cues. Higher frequencies and a larger head size cause a larger baffling effect and a larger interaural intensity difference (IID or ILD). When the sound source is situated in front of one ear of the listener, the IID reaches its highest value for a specific frequency and can be as large as 8 dB at 1 kHz and 30 dB at 10 kHz (Steinberg & Snow, 1934). Thus, IID is a powerful binaural localization cue at high frequencies but fails at low frequencies. Please note that the complex sound arriving at the proximal ear is not only more intense but is also richer in high frequencies (brighter) than the sound arriving at the distal ear. These spectral differences may provide the listener with an additional cue for resolving spatial locations of several simultaneous sound sources such as various musical instruments playing together or two or more vehicles moving at various directions.

At low frequencies, sound localization in the horizontal plane depends predominantly on temporal binaural cues (ITD and IPD). Sound arriving at the two ears of the listener from a sound source situated at a specific azimuth strikes the proximal ear earlier than the distal ear. Assuming that the human head can be approximated by a sphere, the resulting time difference can be calculated with the equation

$$\Delta t = (r/c)(\alpha + \sin \alpha),$$

in which r is the radius of the sphere (human head) in meters, c is the speed of sound, and α is the angle (azimuth) of incoming sound in radians.

The maximum possible time difference between sounds from the same sound source entering the two ears of the listener is about 0.8 ms ($r = 0.1$ m and $c = 340$ m/s) and depends on the size of the human head and the distance of the sound source from the listener. This maximum ITD occurs when the sound source is situated next to one of the listener’s ears. Smaller ITDs indicate a less lateral sound source location. The minimum perceived difference in azimuth occurs when the sound is arriving from 0 degrees (defined as directly in front of the listener) and is equal to about 2 to 3 degrees and corresponds to an interaural time delay of 0.020 to 0.030 ms.

The ITD is used to calculate the difference in arrival time for clicks, onset transients, and non-periodic sounds. Thus, ITD cues can be used for low and high frequency sounds that differ in their amplitude envelopes (onset transients) if the information about the onset transient is available (Leakey, Sayers, & Cherry, 1958; Henning, 1974). For continuous periodic sounds, the time delay of the sound arriving at the farther ear is equivalent to a phase shift between sounds arriving at both ears of the listener. Therefore, in the case of continuous periodic sounds, the term IPD is commonly used to describe the difference in times of arrival. This phase difference

(phase shift) uniquely describes the azimuth of the sound source if the time difference between both arrivals is less than the duration of a half-cycle of the waveform (180 degrees) in air. In the frequency domain, it means that a unique relation between the phase shift and the direction of incoming sound is maintained through low frequencies to approximately 500 to 750 Hz when the half-period of the wavelength becomes greater than the time delay between the two ears. At this frequency, a sound source situated at one ear of the listener produces waveforms at the two ears, which are out of phase, and the IPD cue becomes ambiguous. The listener does not know whether the phase shift of 180 degrees is a result of the waveform in the right ear being a half-cycle behind or a half-cycle ahead of the waveform in the left ear. This means that identical IPD cues are generated by the sound source at the right ear and the left ear of the listener. Small head movements may resolve this ambiguity so there is no well-defined frequency limit in effectiveness of the IPD cues. However, it is generally assumed that phase differences provide useful localization cues for frequencies of approximately 1.0 to 1.5 kHz. In this frequency range, small head movements are sufficient to differentiate between potential sound source locations on the left or the right side of the listener. Above this frequency, the number of potential sound source locations is larger than two and the IPD cue is no longer effective. The IPD cues are the strongest for frequencies between about 500 and 750 Hz and are less effective for higher (ambiguity) and lower (small change in phase) frequencies.

The two mechanisms just described are the foundation of the “duplex theory” of sound localization (Rayleigh, 1907). According to this theory, sound source location in space is defined by the IPD mechanism at low frequencies and the IID mechanism at high frequencies. Because the frequency ranges in which these two binaural cues operate poorly overlap, localization errors in the horizontal plane are the largest for sound sources emitting signals in the 1000-Hz to 3000-Hz range. Moreover, people are very sensitive to sounds in this frequency range, and any reflections can be very detrimental to spatial orientation. In addition, Sandel, Teas, Feddersen, and Jeffress (1955) reported that the listeners have a natural tendency (bias) to underestimate the deviation of the sound source from the median plane for tones in the 1000- to 5000-Hz range. All these effects together make middle frequency sounds very difficult to localize. Recall also that simpler (more tonal) signals cause poorer localization accuracy. Last but not least, binaural cues provide reliable information about position on the left-right axis; however, they are very ineffective for estimation of sound location in the vertical plane (elevation) or along the front-back axis. Human ability to localize sounds along these dimensions is based primarily on the monaural cues described in section 3.2.

One additional binaural mechanism that plays an important role in sound source localization is the “precedence effect” (Wallach, Newman, & Rosenzweig, 1949). The precedence effect, also known as “the law of the first wavefront” (Gardner, 1968; Blauert, 1999) or “Haas effect” (Haas, 1972), is an inhibitory effect that allows one to localize sounds, based on the signal that reaches the ear first (the direct signal) and inhibits the effects of reflections and reverberation. It applies to inter-stimulus delays larger than those predicted from the finite dimensions of the human head but shorter than ~50 ms. If the interval between two sounds is very small (less than 0.8 ms), the

precedence effect does not operate and the sound image is heard in a spatial position defined by the ITD. However, if the time difference between two brief sounds exceeds 0.8 ms and is shorter than 5 ms for single clicks and 30 to 50 ms for complex sounds, both sounds are still heard as a single sound. The location of this fused sound image is determined largely by the location of the first sound. This is true even if the lagging sound is as much as 10 dB higher than the first sound (Wallach, et al., 1949). However, at higher intensities of reflections, the shift in an apparent position of the sound source attributable to the presence of an interaural time delay can be compensated by the interaural intensity difference inducing the shift in the opposite direction. If the time delay exceeds 30 to 50 ms, both sounds are not fused and are heard separately as a direct sound and an echo (see section 4). The precedence effect operates primarily in the horizontal plane, but it can also be observed in the median plane (Rakerd & Hartmann, 1992, 1994).

The effect of the delayed sound on the spatial position of the fused event depends on the interval between the lead and lag. The lagging sound tends to “pull” the perceived sound location away from that of the lead. It is noteworthy that if the primary sound and the secondary sound differ greatly in their spectral (timbral) characteristics, the precedence effect may not occur. This means that the sound reflection from the wall, which is highly dissimilar from the original sound, may be heard separately from the original sound even if the time delay is less than 30 to 50 ms (Divenyi & Blauert, 1987). The precedence effect does not completely eliminate the effect of the delayed sound even if its level is relatively low. It makes the delayed sounds part of a single fused event and it reduces the effect of directional information carried by the delayed sounds. However, the changes in the pattern of reflections can still be detected and they can affect the perceived size of the sound source, its loudness, and its timbre (Blauert, 1999).

2.2 Elevation

Sound source elevation and sound source position along the front-back axis are determined primarily by the monaural cues. Despite the general success of binaural cues and the “duplex theory” in explaining localization of sound sources in space, they still leave an unresolved region known as the “cone of confusion,” i.e., a cone extending outward from each ear and centered on the lateral axis connecting the two ears of the listener. All locations on this cone have the same binaural differentials (see figure 2) and cannot be resolved by binaural cues¹ (Oldfield & Parker, 1986). Therefore, other perceptual mechanisms are needed to specify the location of the sound source on the cone. This is the domain of the monaural cues. Monaural cues are directionally dependent spectral changes that occur when sound is reflected from the folds of the pinnae and the shoulders of the listener. Passive filtering of sound caused by the concave surfaces and ridges of the pinna is the dominant monaural cue used in sound localization. The filtering effect of shoulders is weaker but it is also important since it operates in slightly different frequency range. The resulting spectral transformation of sound traveling from the sound source to the ear

¹This is not strictly true. The “cone of confusion” model assumes a spherical head. However, auditory localization error patterns generally support the belief that this model approximates human behavior well.

canal (and reflected from the body and pinnae) is direction dependent. This directional function is called the head-related transfer function and is often referred to as the HRTF². The resulting spectral changes are largest in the frequency ranges above approximately 4 kHz and can be best interpreted in reference to the spectral content of the original sound. The richer the sound is, the more useful the monaural information will be.

People can localize sound sources in the horizontal plane with one ear but localization error is much greater (~30 to 40 degrees) than that resulting from the use of binaural cues (~3 to 4 degrees). Lack of clear horizontal information affects listener self-confidence and makes monaural cues and related head movements less effective in the judgment of sound source elevation or front-back position. Similarly, elimination of monaural cues affects localization effectiveness of binaural cues in the horizontal plane. Thus, monaural and binaural cues cannot be treated as linearly related and they enhance each other.

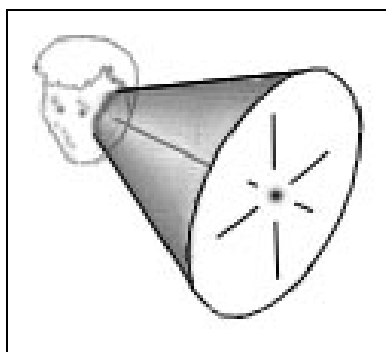


Figure 2. Cone of confusion.

It needs to be stressed that monaural spectral changes occur relative to the original sound source, and therefore, their interpretation requires some familiarity with the original sound source. For example, Plenge and Brunschen (1971) reported that short, unfamiliar sounds were consistently localized by their subjects at the rear of their actual location (front-back error). After a short familiarization session, the number of such errors greatly decreased. In addition, small physiological (unintentional) movements of the head aid in sound localization by providing the listener with information about the spectral characteristics for different head positions (Noble, 1987). However, head movements are only beneficial for sounds of durations greater than approximately 400 to 500 ms. If the sound is very short, it disappears before the head movement is initialized or before the head makes a sufficient rotation (when the head was already moving). Moreover, some sounds have a tendency to be localized low or high, independent of the actual position of the sound source. For example, people have a tendency to localize 8-kHz signals as coming directly from above. Figure 3 presents a graph from Blauert (1999), which shows the effect of frequency band on perceived location in the median plane. The vertical axis gives the

²The monaural filtering effect of each pinna is measured for each ear separately. However, because the HRTF consists of these two filters together, binaural cues are present also.

percentage of judgments placing the sound behind, above, or in front of the listener as a function of the frequency of the stimulus. These data support the notion that humans are not normally as adept at localizing elevation and front-back position of a sound source as they are at localizing the horizontal position of a sound source along the left-right axis. This makes estimates of elevation and front-back position especially susceptible to non-specific factors such as expectations, eye position and sound loudness (Davis & Stephens, 1974; Getzmann, 2002; Hartmann & Rakerd, 1993; Hofman & Opstal, 1998).

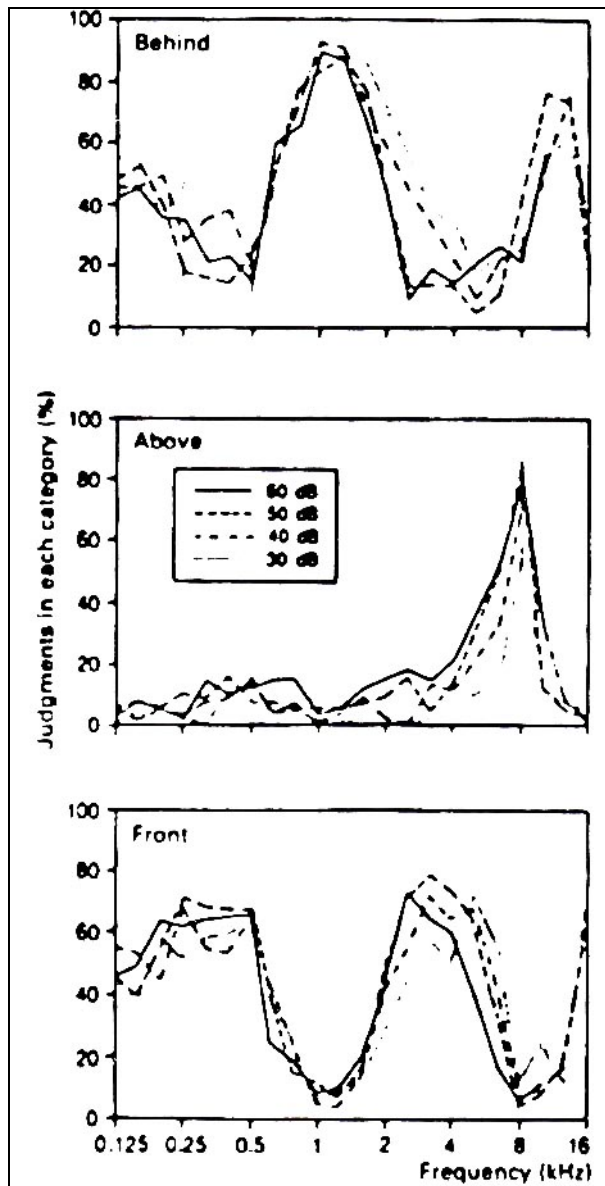


Figure 3. Effect of frequency band on localization in the median plane.

2.3 Distance

Auditory distance estimation is primarily affected by sound loudness (intensity), sound spectrum, and temporal offset (decay). All these cues require some knowledge about the original sound source and the acoustical characteristics of the environment. Their effect depends also on the expectations of the listener and other sensory information. Because of the complexity of conditions affecting auditory distance judgments, these judgments are quite inaccurate and result in about 20% error or more (Moore, 1989). In addition, many people cannot translate perceived distance into numerical judgments, and people differ greatly in the assumed frame of reference when judging the distance. These difficulties create a real problem with the reliability and validity of reported data and need to be addressed.

The most natural auditory distance estimation cue seems to be sound intensity (Mershon & King, 1975). According to the inverse square law of sound propagation in open space (see section 4.1.1), sound intensity decreases by 6 dB per doubling of the distance from the receiver. Therefore, a comparison of the currently perceived intensity to the expected intensity of the original sound source at a specific distance can provide one cue for estimating the distance to the sound source in an open environment. However, this cue requires some familiarity with the specific source of the sound or at least with the specific class of sound sources. In addition, the listener's movement toward or away from the operating source may provide a needed frame of reference (Ashmead, LeRoy, & Odom, 1990).

In rooms and other closed spaces, the decrease of sound intensity may initially follow a 6-dB rule but soon becomes less because of room reflections from nearby surfaces (e.g., the floor). This decrease continues as long as the energy of the direct sound exceeds that of the reflected sounds and a direct sound field becomes a reverberant field. The distance from a sound source where both sound energies are equal is called the *critical distance*. Inside the critical distance, sound localization is practically not affected by sound reflections from space boundaries because of the precedence effect. The precedence effect, however, may not operate at larger distances and higher intensities of reflected sounds. Therefore, the closer the listener is to the sound source and the farther both of them are from the space's boundaries, the less effect the environment has on the localization accuracy.

Another cue for distance estimation is the changes in sound spectrum caused by the frequency-dependent absorption of sound energy by the air. Sounds arriving at the listener from larger distances may sound as if they were low-pass filtered when compared to the original sounds. Humidity has a similar effect on attenuation of high frequencies. If one has knowledge of the original sound source as well as knowledge of the weather conditions and intervening environment (e.g., walls, objects), the spectral changes attributable to air absorption provide useful information about the distance to the sound source (Brungart & Scott, 2001; McGregor, Horn, & Todd, 1985; Mershon & King, 1975). However, without the listener's familiarity with the sound source, the changes in sound spectrum provide only relative but not absolute information about the distance to the sound source (Little, Mershon, & Cox, 1992).

Sounds reflected (reverberated) from the ground, walls, and other objects last longer and decay more slowly than the original sound. The more reverberant the environment and the larger the distance between the sound source and the listener, the more extended in time the sound is perceived to be by the listener. Therefore, reverberation constitutes a very effective if not the main cue for distance estimation in most environments (both indoors and outdoors). As distance between the sound source and the listener increases, the amount of direct sound energy arriving at the listener's ears decreases and the amount of reverberant (reflected) energy increases (Mershon, Ballenger, Little, McMurtry, & Buchanan, 1989; Nielsen, 1993). However, the specific ratio of these two energies depends also on the directivity of the sound source and the listener's hearing, the size of the space, and the position of the sound source relative to the walls and the listener (Mershon & King, 1975). Furthermore, small and highly reflective spaces may create the same perceptual effects as larger and more damped spaces. Thus, reverberation information coming from unknown and unseen spaces (such as adjacent rooms or buildings) is unlikely to provide usable distance information until the listener becomes familiar with the space. It is also important to recall that the distance judgments are complicated by the difficulty most persons have expressing distance in numeric units. This ability, however, can be developed with experience and by specialized training.

2.4 Auditory Localization Capabilities and Limits

Sound localization requires the integration of binaural information in the brain stem. ITD and IID information is computed in the lateral superior olive (SO) and then later mapped into the inferior colliculi (IC) (Gelfand, 1998). Because neural output from IC is processed by specific (the auditory cortex) and non-specific centers, auditory sensory information is combined with visual sensory information and cognitive expectations, all affecting the perceptual orientation of a person in space. Thus, the elements affecting sound localization in space can be divided into physical elements (i.e., sound, source, and environment related) and psychological elements such as attention and memory.

Precision of sound source localization depends primarily on the type of sound source, the listener's familiarity with the source, and the type of acoustic environment. It is also affected by the sound duration, relative movements of the sound source and listener, and presence of other sounds in the space. A listener's expectations and other sensory information can also affect his or her judgments.

Three types of precision measures are used in localization studies: localization accuracy (LA), minimum audible angle (MAA), and minimum audible movement angle (MAMA). Appendices A, B, and C provide results from selected studies of LA, MAA, and MAMA measures, respectively. Localization accuracy (LA) is defined as an absolute precision in reporting the direction of incoming sound. Average LA error for horizontal localization of a sound source ranged from 1 to 15 degrees, depending on several factors such as the observation region (Oldfield & Parker, 1984) and the frequency content (Butler, 1986) of the signal. Reported errors frequently did not include

the front-back errors. Elevation errors were slightly higher (4 to 20 degrees) than horizontal errors (Oldfield & Parker, 1984; Carlile, Leong, & Hyams, 1997). Accuracy varies with the method used to “point” or estimate the location of the sound source.

MAA refers to the smallest angular separation of two sound sources that can be discriminated. Listeners may be asked to indicate if the second of a pair of sounds comes from the right or the left of the first reference sound. Data from selected studies are given in appendix B. In general, listeners are able to distinguish differences in azimuth as small as 1 degree (Mills, 1958). The MAA increases slightly when the sounds are situated near 90 degrees, and this finding has been replicated in a number of studies. However, the ability to discriminate differences in elevation is much worse ranging from 6 to 20 degrees. Some listeners were unable to localize sounds with precision better than 20 degrees (Grantham, Hornsby, & Erpenbeck, 2003). Some factors that affect MAA precision are the frequency content of the stimuli, the time delay between the onsets of the presented stimuli, and the amount of stimulus overlap. It is believed that inter-stimulus onset delays of at least 150 to 200 ms are required to discriminate the MAA because such time is required for the auditory system to process the frequency content of a signal (the monaural information).

MAMA refers to the minimum movement of a sound across a given axis required to detect a sound as moving. The ability to detect and localize moving sounds is discussed in section 4.3.2. Appendix C provides a sample of the data from several selected studies. Generally, people require 4 to 20 degrees of horizontal movement (more for movement in elevation) to detect that movement has occurred (Perrott & Musicant, 1977; Chandler & Grantham, 1992).

3. Acoustics of the Urban Environment

When gathering data about the environment and making decisions about movements, people rely predominantly on visual observations and visual memory. In urban environments, many visual cues are missing or obscured and acoustic information becomes an important factor that affects SA. Even when visual information is available, the importance of audition cannot be overstated since the ears are the only directional tele-receptor that operate in the full 360-degree sphere. People respond to sound by turning their heads toward incoming sound and use both hearing and vision for more accurate localization of the potential source of sound. Therefore, awareness of the specific acoustic environment surrounding the Soldier in an urban battlefield is critical for a Soldier’s effectiveness and safety.

The acoustic environment can be defined as a sound field created by all sound sources and other physical objects surrounding the listener. This sound field is a combination of direct sound waves radiated by acoustic sources and numerous sound reflections created when the sound waves bounce back from objects in the space and the space boundaries. The acoustic environment is also affected by a number of other acoustic phenomena. These include diffusion

(scattering), diffraction (bending around the edges), refraction (bending during transmission to other media), acoustic shadow, interference (e.g., acoustic beats), standing waves, amplification (resonating), and attenuation (damping). Additionally, the acoustic environment is affected by the presence of a background noise and the relative movements of sound sources and the listener within the environment. Background noise is a spatially uniform sound created by external sound sources through vibrations of space boundaries and by internal sound sources through multiple reflections of sounds from space boundaries and other objects within the space. Background noise can also include the higher order reflections from the target sound of interest. Therefore, some parts of the background noise may be correlated with the sound of interest and others are independent.

These phenomena affect human ability to identify the exact position of a sound source as well as other aspects of auditory awareness such as sound detection and identification. They can be called *acoustic signal processing phenomena* or *sound modifiers* because they affect all spatio-spectro-temporal characteristics of the sounds arriving at the listener.

The urban environment differs from rural or open environments in that sounds are bounced back and forth with relatively small loss in sound energy from a large number of closely spaced reflective surfaces. These include hard walls with and without openings, parallel walls, hard ceilings and floors, and numerous stationary and moving objects. These strong multiple reflections together create a high level of correlated background noise and provide false or ambiguous sound localization cues that reveal more about the environment topography than about the actual position of the sound source within the environment. Sound reflections as well as the other acoustic factors discussed previously are not necessarily unique to the urban environment, but they become especially important in the physically complex urban battlefield because of their number and strength as well as the lack of visual support in object localization. Last but not least, multi-story buildings with windows, balconies, a variety of roofs, and highly reflective streets and parking lots create a three-dimensional acoustic environment in which sounds must be localized in azimuth as well as in elevation and depth.

Previous discussion (sections 2.1, 2.2, and 2.3) indicated that human ability to localize a sound source is affected by the kind of information that is available in the sound itself and by the degree to which this information becomes a part of background noise in the environment. Recall that monaural localization cues require prior knowledge of the sound source and the acoustic context in which the source operates. These cues provide little help to the Soldier who is ignorant of the identity of the sound source or has never been in the environment. As a result, the ambiguous localization cues and unfamiliar listening conditions, together with scarcity of visual information, make the “visual-capture effect” (see section 3.3.1) a dominant source of localization errors in the urban environment.

All sounds reflected from nearby and distal objects can be divided into three overlapping classes: early reflections, late reflections, and echoes. When the reflected sound wave reaches the ear within approximately 50 ms of the direct sound, both sounds are combined perceptually into one

prolonged sonic event with the perceived sound source location dictated by the *precedence effect*. Such reflections are called *early reflections*. They increase overall sound intensity (loudness) without changing the perceived incoming direction and duration of the signal. They also increase spatiality (perceived size) of the sound source and cause a perceived change in the sound spectrum (timbre) referred commonly to as *sound coloration*. However, there is an intensity limit within which the precedence effect operates. If the intensity of the reflected sound is sufficiently high in comparison to that of the direct sound, it may cause a shift in perceived sound source location toward the direction of reflected sound (see section 3.1). Even at lower intensities, reflected sounds can cause a perceived change in the sound spectrum, referred commonly to as *sound coloration*, which may provide false cues regarding the sound source location.

Late reflections are the reflections that arrive 50 ms or more after the direct sound. In most rooms, late reflections are very dense and cannot be differentiated from one another. They also become weaker with time and with the number of walls from which the sound was reflected. They extend the decay of the sound and increase the likelihood of overlap with subsequent sounds, thereby causing masking and smearing effects.

The gradual decay of sound in a space (room) is called *space (room) reverberation*. Reverberation is a product of all sound reflections arriving at a given point in space. Keep in mind, however, that early reflections contribute mainly to perceived loudness of sound, whereas late reflections contribute to perceived size of the space and related rate of sound decay. Therefore, for all the practical purposes, sound *reverberation* can be defined as a sequence of dense and spatially diffuse reflections from space boundaries, which cannot be resolved by the human ear and are perceived as a gradual decay of the sound in the space. Reverberation is characterized by reverberation time (RT_{60}) that is defined as the time needed for a sound level at a given point in space to decrease by 60 dB from the moment of sound source offset. Reverberation time is proportional to the volume of the space, reflectivity of space boundaries, and frequency of the sound. This relationship is most frequently expressed by the *Norris-Eyring formula*:

$$RT_{60} = 0.161 \frac{V}{\sum S_i \ln(1 - \alpha_i)},$$

in which V is the volume of the space (m^3) and S_i and α_i are an average coefficient of absorption and the area of the i element of the space boundaries, respectively. In reflective environments (where $\alpha < 0.3^3$), $\ln(1 - \alpha) \approx -\alpha$ (with a maximum error of 5.7%) and the previous equation can be simplified to the form

$$RT_{60} = 0.161 \frac{V}{\sum S_i \alpha_i}.$$

³This criterion is met by many of the laboratories at ARL (Scharine, Tran, & Binseel, 2004).

Echoes are late reflections that are distinguishable as separate acoustic events from the direct signal. They can be heard when the signal is not masked by other reflections and other simultaneous sounds. In order for an echo to appear, the distance between the paths traveled by the direct and reflected sounds needs to exceed 17 meters (assuming that the speed of sound equals 340 m/s at 20 °C) (figure 4).

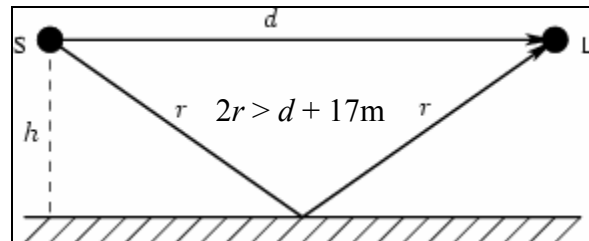


Figure 4. Echo effect.

When a sound is repeatedly reflected between two parallel flat surfaces, the resulting product is a sequence of echoes called flutter echo. Flutter echo sound is a sequence of noise pulses. If the surfaces are less than 30 feet apart, individual echoes blend together into a single periodic event with fundamental frequency defined by the distance between the walls. Such flutter echo becomes a zing-sounding (buzzing, ringing) flutter tone that is easy to detect but very annoying. Flutter sounds only originate when the reflected surfaces are parallel to each other and will not appear if the walls are skewed by as little as 5 degrees.

3.1 Walls and Buildings: Physical Properties of the Environment

Reflective surfaces of walls, buildings, and rooms modify the distribution of sound energy in the space and alter direction and spectro-temporal properties of sounds arriving at the listener's ears. The properties of these sounds depend on the shape and relative positions of individual surfaces, structural support and construction material, and spatial arrangement of these surfaces in reference to the position of the sound source in the space. The closer the sound source to a reflective surface, the stronger the reflection. The farther the sound source is from the reflective surface, the more the reflection is delayed, increasing the probability of our hearing an echo. The listener's task is to predict the location of the sound source, based on the sounds arriving at the ears and the listener's knowledge about the sound source and environment. For example, if the listener knows that the terrain behind the building directly in front of him or her is empty and grassy, it cannot be a location of a tank moving with a rambling high pitch sound, even if the localization cues indicate that direction. If the sound coming from that direction is heard as a rambling high pitch sound, it must be a reflection of a sound coming from another direction and the listener's task is to identify this direction.

3.1.1 Reflection and Reverberation

Sound arriving at the listener's ears is composed of direct and reverberant (reflected) energy. These reflections can impede localization in both the horizontal and vertical planes. Since the

reflected sounds can be quite strong and last beyond the end of the direct sound, they can attract the listener's attention toward the direction of the reflection rather than the direction of the original sound source. In an open (free) field, the direct sound energy produced by an omnidirectional sound source decreases gradually with increasing distance at a rate of 6 dB per doubling of the distance as described by the inverse square law formula (Howard & Angus, 1998):

$$I_d = \frac{Q_s W_s}{4\pi r^2},$$

in which I_d is the intensity of direct sound at a given point in space (W/m^2), Q_s is the directivity⁴ of the sound source (compared to sphere), r is the distance from the sound source (m), and W_s is the acoustic power of the sound source (W). Please note that I_d , Q_s , and W_s are frequency dependent.

In closed or semi-closed spaces, the attenuation of direct sound energy can be less than in an open field because reflective surfaces are present near the sound source. This is greatly affected by the directivity coefficient and spatial orientation of the sound source. At large distances, sound pressure becomes dominated by reverberated energy and becomes independent of the distance to the sound source. During the sound presentation, reverberant energy in the space is directly dependent on the energy of the sound source, the size of space, and acoustic properties of space boundaries and can be roughly estimated via the following equation:

$$W_r = W_s 4 \left(\frac{1-\alpha}{S\alpha} \right) = W_s \frac{4}{R},$$

in which W_r and W_s are reverberant sound power and sound source power, respectively; α is an average coefficient of the absorption of space boundaries; S is the total area of the space boundaries (m^2), and R is the room constant (m^2). The equation assumes an omnidirectional sound source, steady state sound, and acoustic symmetry of the space. For the points in space far away from the sound source, the energy of the reflected sounds dominates the sound field and creates a spatially diffuse field with sound pressure level changing in space and time according to a normal distribution with a standard deviation equal to (Lubman, 1968)

$$\sigma = \frac{4.34}{\sqrt{\left(1 + \left(\frac{BT}{6.9}\right)\right)}} dB,$$

⁴Directivity is a measure of the directional characteristic of a sound source. It can be quantified as a directivity index in decibels or as a dimensionless value of Q. Sound from a point source would send sound in all directions equally, and this would represent a Q value of 1. Sound radiating in a hemispherical pattern would have a Q value of approximately 2 (Beranek, 1960).

in which T is the reverberation time in seconds and B is the signal bandwidth in Hz. The longer the reverberation time and the more wide band the signal, the smaller variability of reflected sound energy in space (Lubman, 1968).

The shape and material of reflective surfaces and their geometrical relation to each other affect the distribution of sound energy in space and the temporal envelope of the sound signal reaching the listener. In general, the effects of reverberant energy on sound source localization depend on whether the energy is from early reflections, from non-directional late reflections creating a noise floor correlated with the direct sound (reverberation), from strong directional reflections, or from echoes that are perceived as distinct sound events. Early reflections are fused perceptually with the direct sound and have two possible effects on auditory orientation. If the localization cues produced by the early reflection are congruent with those of the direct sound, then the reflected energy can be beneficial, increasing signal detectability and localizability (Rakerd & Hartmann, 1985). This is true especially if one is primarily interested in horizontal localization because the reflected sounds from the ground (floor) and (when indoors) a ceiling contain the same directional cues as the direct sound and therefore increase the strength of localization cues. For example, Hartmann (1983) found that by lowering ceilings and thus causing the early reflection to occur earlier, horizontal localization performance was improved. However, if the reflected energy arrives from directions that are incongruent with the direction of arrival of the direct signal, the perceived image of the sound source may become less defined (larger) or even drawn toward the direction of the reflected sound (Rakerd & Hartmann, 1985). These effects are especially noticeable in situations when the *precedence effect* is compromised or fails to operate. In the case of elevation, even reflections with congruent horizontal cues can be detrimental to accurate vertical sound source localization. Guski (1990) found that a single reflective surface above the head of the listener (a ceiling) disrupted localization in elevation more than if it were located below (a floor)⁵. This can be explained by the atypical nature of this acoustical configuration. Humans are accustomed to encountering floors without ceilings in the outdoors; however, it is rare to encounter a ceiling with no floor.

Reverberation effects lasting beyond 50 ms after the end of the sound (late reflections) impair localization. Hartmann (1983) asked listeners to perform a localization task in a chamber where the wall panels could be adjusted to vary their absorption coefficient and the ceiling could be raised or lowered. He found that the ability to localize broadband (square wave) sounds was better for the less reflective room. Reverberation changes localization cues in several ways (Kopčo & Shinn-Cunningham, 2002). First, by the introduction of variability into the spectral information, monaural information is reduced. Second, by the addition of noise into the signal, binaural interaural level differences are reduced. Finally, reflections may create a second energy peak (a false onset cue) that is temporally implausible, adding false ITDs to the real ones. All these effects worsen as source distance increases and the ratio of reverberant to direct energy increases.

⁵An anechoic chamber was used so that the only reflective surface was the ceiling.

To examine how sensitive listeners are to the configuration of walls in a room, Kopčo and Shinn-Cunningham (2002) measured the localization performance of six listeners placed in several positions relative to the walls of a small room (5 by 9 meters). The sound sources were placed at three distances (0.15, 0.40, and 1.00 meter) from the listening positions. Performance was affected by reverberation; there was evidence of bias caused by the fusion of early reflections with the direct sound when a wall was situated opposite an ear. Further, localization performance was worse at increased distances, which suggested that the reduced direct-to-reverberant energy ratio had a negative impact. Similar data for larger distances were also reported by Henry and Letowski (2004). However, Kopčo and Shinn-Cunningham (2002) also reported that contrary to predictions, localization performance was only modestly affected by the listener's position within the room. This latter finding suggests that listeners are able to adapt somewhat to the acoustical properties of a room and discount those features in their estimates. This was supported by measurements of the output of medial superior olive (MSO) neurons which suggest that although instantaneous temporal information is obscured by reverberation, cross-correlation over time may allow room information and sound information to be segregated (Shinn-Cunningham & Kawakyu, 2003).

If a listener can adapt to a particular acoustic environment, are they aware of these acoustic features? It has been shown that listeners are sensitive to sound reflections that might indicate changes in the physical environment such as the movement of a wall (Clifton, Freyman, Litovsky, & McCall, 1994). However, because humans are unlikely to encounter such implausible changes in the real world, this information is easily misused. If a particular localization cue is not possible or if it signals improbable circumstances (e.g., the walls are moving or changing absorptiveness), then listeners will weigh that information less and rely on other information to localize the sound (Rakerd & Hartmann, 1985). Thus, it appears that the auditory system can detect acoustic features sufficiently to ignore improbable cues. However, it is unlikely that the auditory system can interpret this information further to give information about the size and or position of walls. Shinn-Cunningham and Ram (2003) simulated the presentation of nine sound sources (white noise presented from nine locations) for four listener locations within a "virtual⁶" room. The listener's task was to indicate the perceived position of the sound sources within this room. Listeners were unable to do this accurately. For those who were able to determine their position, the perception of location seemed to be most dependent on the difference between the amounts of direct sound energy perceived by the two ears. Listeners identified themselves as being near a wall on the side where direct energy was strongest.

3.1.2 Sound Path Barriers

In MOUT environments, sounds may come from behind fences and barriers, around walls, from adjacent rooms, or from nearby buildings. All these structures can occlude the original sound source, forcing sound to travel around them. Sound traveling around barriers has different

⁶Sounds were recorded with an acoustic manikin and presented to the listener via headphones.

spectral characteristics than unimpeded sound; specifically, the longer wavelengths of the low frequency components are less vulnerable to acoustic shadow and more likely to travel around the obstruction. Farag, Blauert, and Alim (2003) found that the effects on localization ability could be predicted if they assumed that localization was based on the resulting redirected pathway. They found that if the sound's pathway was occluded by a wooden panel, the perceived location of the auditory event shifted in a manner consistent with that predicted by the precedence effect. In this case, the sound was free to travel around both sides of the panel and the percept was shifted toward the side of the panel for which the sound path was shortest. It is, as of yet, unclear whether the listener can perceive that the sound has been diverted or if this perception depends on the absorptive properties of the occluder.

3.1.3 Vibration

Sounds travel through other media than air. Pipes, building construction elements, and underground infrastructure of an urban area will propagate sound waves faster and farther (with less loss of energy) than air. Such structures emit the sounds through their large surfaces and outlets (pipes), behaving as waveguides and distributed sound sources. These structure-borne sounds add to the auditory confusion of the urban terrain because the real source of the sound can be far away from the sound-emitting element. In addition, waves traveling through the structures can be repeatedly reflected by two parallel surfaces, creating a phenomenon of standing waves, which are a source of mechanical vibrations. These mechanical vibrations become the secondary sources of sound that are not necessarily spatially congruent with the location of the forces that created the sounds. Therefore, it is quite frequently impossible to determine the location of the primary source of vibrations in the absence of reliable airborne sound localization cues.

3.2 Battlefield Conditions: Noise-Induced Chaos

3.2.1 Noise

Noise is an important psychological weapon. The U.S. Army field manual for urban offensive operations (U.S. Department of the Army, 2003) states that surprise, concentration, tempo, and audacity are especially characteristic of urban maneuvers. Soldiers report that noise is an essential element in offensive urban operations. It can be used to surprise and startle the opposition and to convey speed and authority. For example, intense sounds (music, noise, messages) played from loudspeakers mounted on low-flying helicopters or on moving vehicles may annoy and disorient the enemy as well as mask other sounds that we want to make undetectable by the enemy. However, the use of such noise sources in close combat urban environment can also mask important auditory localization cues, making the urban battleground even more ambiguous and dangerous.

Rakerd and Hartmann (1986) note the importance of the temporal cues provided by onset⁷ transients for sound source localization. In their experiments, localization was always worst for the conditions where the onsets of the signals were essentially removed by introduction of the sound very gradually. Because the binaural cue of interaural time difference relies on the lag between the onset of the sound reaching each ear, an important source of localization information is lost when the onsets are removed. In the battlefield environment, onsets can be effectively “removed” by the masking effects of ambient noises and long sound decays (the effect of a preceding sound).

In order to localize a single target sound in noisy background, the signal-to-noise-ratio (SNR) or sensation level (SL)⁸ of the target sound must be high enough not only for the sound to be heard but also to be interpretable. Appendix D presents data from four studies, two that investigated the SNR and two that investigated the SL needed for accurate localization in the horizontal plane. An SNR of at least -7 to -4 dB was needed to achieve 50% accuracy (to within ± 15 degrees) for listeners with normal hearing. The SL needed to be at least 9 dB in order for listeners with normal hearing to achieve similar performance.

3.2.2 Multiple Sound Sources: Acoustic Distracters

Most localization research has focused on the ability to localize a single sound source, either in quiet or in noise. However, in a natural environment, there are usually multiple sound sources, any one of which may require attention. If two sounds occur simultaneously, it may be difficult to attend to one sound sufficiently to localize it. In general, the closer in space two sound sources are, the greater the difficulty of localizing either of them properly (Smith-Abouchacra, 1993; Zurek, Reyman, & Balakrishnan, 2004). Smith-Abouchacra (1993) presented listeners in an anechoic room with a target and a distracter at various target-to-distracter intensity level ratios. The relative angular separation between target and distracter was varied from 0 to 315 degrees, and eight horizontal positions encompassing the entire perimeter were used. The presence of a distracter had several detrimental effects on localization. First, detection of the target decreased with its horizontal distance from the distracter. Second, presence of a directional noise (one situated at a single location) caused listeners’ localization percepts to be biased either toward or away from the distracter. The direction of the bias depended on the positions of the target and distracter. Perceptions of targets situated frontally and in the same hemisphere as the distracter tended to be shifted toward the distracter if the distracter was more laterally located. Otherwise, they were shifted in the opposite direction. Localization estimates of targets in the rear hemisphere were susceptible to front-reverse errors and tended to be shifted away from the distracter.

⁷The onset is the beginning portion of a sound: the “attack”.

⁸Sensation level refers to the number of decibels by which a sound exceeds a person’s hearing threshold.

Braasch and Hartung (2002) conducted a similar experiment where targets were presented from each of 13 positions in the frontal hemisphere coincidental with a distracter in one of three locations (0, 30, or 90 degrees azimuth). Here, the target-to-distracter sound level ratios ranged from 0 to -15 dB and testing was done in both anechoic and reverberant conditions. Similar to Smith-Abouchacra, they found that detection was difficult when the target and distracter were positioned close together. When the target was presented at the same sound level as the distracter, listeners exhibited a bias to localize the target in the direction away from the distracter. Furthermore, when the target was at lower sound intensity levels, spatial resolution decreased, that is, localization judgments were clustered into three primary locations: left, center, and right. Reverberation exacerbated these difficulties.

The fact that reverberation masks the spatial cues that aid the listener in isolating individual sounds means that detection and identification of target sounds are also impaired. The reflections of sounds produced by the distracting sound sources increased the masking effect of distracters and raised the detection threshold of a target sound even further (Zurek et al., 2004). Accurate sound source localization in such environments is additionally complicated by the presence of front-back confusions that result from not only false physical cues but also the listener's potential lack of familiarity with the specific sound source.

If sounds occur in proximity to and earlier in time than the target, they can disrupt localization even if the distracter serves to draw attention to the same spatial region in which the target is to occur. Kopčo and Shinn-Cunningham (2002) presented target sounds with an auditory cue that signaled the particular region in which the target sound was to occur. For cue-target delays as long as 300 ms, the cue interfered with localization of the target, biasing localization toward the cue.

3.3 Other Factors

3.3.1 The Effect of Vision on Auditory Localization

In urban terrain, events happen around corners and behind walls where visual information is neither available nor relevant. Given degraded visual information, sound provides an important source of additional information. However, as shown previously, auditory localization information is often ambiguous or difficult to interpret. This raises an important question: how do vision and audition interact when both types of information are available?

A number of studies have measured the effect of an auditory stimulus on perceived intensity, the discrimination threshold, or the detection threshold of a visual stimulus (Gilbert, 1941). Many found facilitative effects of sound on visual orientation (Chason & Mockovak, 1970; Chason & Berry, 1971; Hartmann, 1933; Kravkov, 1934, 1939a, & 1948; Symons, 1963). Similar effects have been found for color (Allen & Schwartz, 1940; Costa & Bertoldi, 1936; Jakovlev, 1940; Kravkov, 1936 & 1939b). However, a number of other studies found negative effects, no effects, or large individual differences (Burnham, 1942; Ince, 1968; Kravkov, 1939a; Loveless, Brebner,

& Hamilton, 1970; Warner & Heimstra, 1971, 1972). These effects depend on a number of factors such as the frequency content of either the sound (Maruyama, 1959) or the visual object (Costa & Bertoldi, 1936; Jakovlev, 1940; Kravkov, 1939a), the temporal relationship of the multimodal information (Kravkov, 1934), the degree of adaptation to the noise (Burnham, 1942), the task characteristics, and the individual characteristics of the person being tested (Ince, 1968). The value of this literature is in pointing out that in some circumstances, sound can affect visual sensations and that visual information, if barely available, may become clearer because directional congruent sound is present. However, in an urban environment, the Soldier seldom encounters such congruent situations and in most cases has to confront acoustic reflections, visual reflection (windows, metal walls, etc.), or both that provide contradictory cues.

It is more useful then, to consider the complex environment where some form of scene analysis, either visual or auditory, is needed to interpret events in the scene. Normally, because of their transient nature, sounds alert the person to the presence and approximate location of the sound source, but vision supplements and refines this information. The information relied upon when vision and audition are incomplete or misinterpreted depends on the information most likely to be correct (Wada, Kitagawa, & Noguchi, 2003).

Vision is superior to audition for acuity of spatial information (Perrott, Costantino, & Ball, 1993). As a result, visual location information is weighted more heavily than auditory localization cues. One example of this is known as the “ventriloquism effect” (Thomas, 1941). As the name implies, this phenomenon is commonly associated with the perception that the ventriloquist’s “dummy” is producing the voice rather than the ventriloquist. A more general term is “visual capture,” which occurs when a visual object causes an auditory stimulus to be mislocalized to its location (Bertelson & Radeau, 1981; Driver, 1996; Jack & Thurlow, 1973; Kitajima & Yamashita, 1999; Mateeff, Hohnsbein, & Noack, 1985; Shimojo, Miyauchi, & Hikosaka, 1997; Spence & Driver, 2000).

It seems that we are quite willing to trust visual location information over auditory cues. For example, it rarely concerns people who attend movies that the loudspeakers are placed on the walls to the side and the rear of the audience. Visual capture is made more probable if the visual and auditory events are proximal in location (Bermant & Welch, 1976) or synchronous (Radeau & Bertelson, 1987). The more compelling the visual and auditory objects are, the more likely they are to be fused or grouped together as a single event (Warren, Welch, & McCarthy, 1981; Radeau & Bertelson, 1977). Cognitive expectations also affect the strength of the capture, meaning that sounds will be localized in part according to where the sound source is expected to be (Weerts & Thurlow, 1971). The potential result is that a convenient but innocuous visual object is presumed to be the source of an alarming sound, or an innocuous sound is judged to be threatening because it comes from a visual object that is deemed to be dangerous.

On the other hand, audition is superior to vision for the detection of temporal changes. A number of studies demonstrate that a visual flicker paired with an auditory flutter will appear to

synchronize together (Wada et al., 2003; Welch & Warren, 1980). For example, a single flash accompanied by two auditory beeps will be perceived as two flashes (Shams, Kamitani, Thompson, & Shimojo, 2002). In general, the temporal onset of visual objects is drawn toward auditory signals that occur in the same temporal and spatial region (Aschersleben & Bertelson, 2003; Bertelson & Aschersleben, 2003).

The ability to monitor a visual scene is limited by the inability to monitor the entire scene for changes. This insensitivity to temporal changes is part of a larger visual phenomenon known as “change blindness”. For example, viewers are sometimes even unable to detect changes that are in the center of focus. Levin and Simons (1997, 2000) presented observers with videos containing a number of scene changes. With each scene change, an object was exchanged, added, or removed. Even though such a change would truly be remarkable if it really happened, it often went unnoticed. This effect is not an artifact of using video; it was replicated in a real-world interaction where a live conversation partner was switched during a small interruption (Levin, Simons, Angelone, & Chabris, 2002).

Auditory information allows one to perceive more events simultaneously, thus allowing for more parallel processing of information. Unlike vision, one can detect auditory events that are outside one’s central focus. The importance of expanded parallel processing, even when visual information is not ambiguous, can be found in the following statement made by a doctoral student, Jason Corey (1998).

While (I was) creating and editing a sound track for an animated film, it became apparent that sounds occurring synchronously in time with visual events on the screen had an effect on how I perceived the visuals. For one particular scene, there happened to be a great deal of activity happening on the screen. Without a sound track, there were many events that were not perceived until a sound effect was synchronized with the particular visual events.

It seems that by having sound accompany a visual, many more details of the visual are perceived...

This anecdote and the previous research findings underscore the complexity of the scene analysis tasks required in urban terrain. Because the tasks that are vulnerable to visual or auditory capture depend on whether the information needed is temporal or spatial, further analysis of the informational needs of Soldiers in urban terrain is needed. When vulnerabilities are discovered, tools and strategies can be developed as aids to avoid misalignment of degraded cues.

3.3.2 Moving Sound and Moving Listener

Movement can both aid and hinder auditory localization. The effects of movement depend on whether it is the sound source or the listener that is moving and whether the localization activity is concurrent with the movement. A moving object emitting a sound may cease to move but continue to sound, or it may continue to move but cease to sound or it may sound briefly and then cease both movement and sounding. A moving listener may be moving during the sound

event or moving after the event has ended. Furthermore, a listener may move only his or her head, change orientation, or move his or her entire body. Any one of these factors will affect the precision of auditory localization, and any combination of dynamic events is probable in an urban environment.

Although sound sources may travel along an infinite number of pathways, the moving sound studied in many laboratory experiments rotates around the listener. Usually, this is accomplished by a loudspeaker mounted on a rotating boom. Occasionally, apparent movement is created by multiple loudspeakers positioned in an arc. Many of these experiments are based on the MAMA paradigm described in section 3.4. (See appendix C for some examples of data from MAMA studies.) The consequence of this is that the localization cues used to detect motion and in this case, angular velocity and acceleration, are used to estimate horizontal and vertical position⁹. It is probable that information derived from movement in depth is equally limited as that used to estimate depth.

Humans process moving sounds differently than stationary ones (Clarke, Adriani, & Bellmann, 1998; Griffiths, Bench, & Frackowiak, 1994; Griffiths et al., 1996; Hall & Moore, 2003). This sensitivity provides an adaptive advantage, allowing us to detect changes in the environment that signal potential danger and opportunities (Neuhoff, 2001). However, we are not as precise at localizing a sound while in motion. For example, Perrott and Musicant (1977), using a MAMA paradigm, found that listener estimates of the starting location of the angular sweep were consistently shifted in the direction of movement. Estimates of the end points were also mis-localized, but errors depended on the duration of the signal, which suggests that listeners might not be able to detect velocity. Other studies show that listener estimates of velocity are proportional to the actual velocity and that listeners can discriminate acceleration and deceleration (Perrott, Buck, Waugh, & Strybel, 1979; Perrott, Costantino, & Cisneros, 1993; Waugh, Strybel, & Perrott, 1979). Unfortunately, listeners seem to be unable to use the velocity, the duration, or the localization information to estimate the location of the beginning and end points of the sound source with the same degree of accuracy as achieved when the sound is stationary (Grantham, 1986).

It might seem incongruent that we are sensitive to but not accurate in localizing moving sounds. However, consider how one might interact with a moving sound. Unless the movement is directly toward the listener, interception requires some form of tracking. As long as the sound is moving, the ongoing location is changing and precise localization is probably irrelevant. If movement stops and sound continues, the listener has been alerted and can now locate the stationary signal. The difficulty arises when the sound has stopped and movement either ceases or continues. Unless the listener has also been able to locate the target and see it, he or she must

⁹This statement is made in spite of the fact that there is an ongoing debate about whether the perceptual mechanisms used for auditory motion perception are the same as those used for stationary perception. It is justified by evidence that suggests that the perception of the location of a moving sound at time t is not significantly different than one based the estimation of the end points and proportion of the total duration (Grantham, 1986).

rely on memory of where the sound appeared to be when it ended. Memory of the last location of a sound is subject to auditory representational momentum, a bias in which memory for the last location of a sound is shifted in the direction of movement (Getzmann, Lewald, & Guski, 2004; Hubbard, 1995; Nagai, Kazai, & Yagi, 2002).

Humans are fairly adept at detecting movement and the direction of movement (Perrott et al., 1993; Strybel, Manligas, & Perrott, 1992). Perceived velocity is proportional to actual velocity (Perrott et al., 1979), and tracking is improved if visual information is available (Somers, Das, Dell'Osso, & Leigh, 2000; Stream, Whitson, & Honrubia, 1980). Movements of the head while the body is otherwise stationary can reinforce binaural cues (Wightman & Kistler, 1999), making auditory localization more accurate, especially if the sound is continuous and the sound source is stationary (Fisher & Freedman, 1968; Handzel & Krishnaprasad, 2002). Similarly, tilting of the head creates binaural differences that strengthen the weaker monaural cues (Noble, 1987; Perrett & Noble, 1997). However, this presumes that the sound source is not moving and that the listener is not changing body position or spatial location relative to the rest of the environment.

It is assumed that Soldiers are moving at least some part of their heads or bodies nearly all the time. However, relatively little research has been conducted to date about the human ability to localize a sound while the whole body orientation or spatial location is being changed. Although not necessarily attributable to the movement, there seems to be a small but significant effect of posture on localization accuracy. Lewald, Dörrscheidt, and Ehrenstein (2000) found that listeners consistently under-rotated when orienting toward a sound or a visual target, which suggests that proprioceptive calibration of head position is subject to error. Visual feedback reduced these errors significantly, but even left-right position relative to the median plane of the head is shifted when the head is rotated on the torso. Lackner (1973) found similar localization errors that were consistent with erroneous proprioception. These findings imply that a small amount of localization error is introduced by the normal variability of body positions that would be expected in non-laboratory conditions. It is likely that localization that occurs while body positions are being changed would contain errors attributable to the sound source's changing position relative to the ears and to mis-estimation of the frame of reference. Obviously, this does not apply to sounds that last during the whole process of movement. In such case, the changes in body position (as with the changes in head orientation) may actually aid in the location of the true position of the sound source.

Movement that occurs after the sound event stops also introduces frame-of-reference errors into the localization estimates. If the remembered position of the head or body during presentation of the sound source relative to head position is incorrect, this error will affect the current estimate of the sound's location (Kopinska & Harris, 2003).

Research on auditory localization by a listener moving during or after the sound has not been tested directly but has been tested indirectly. In research on whether we are able to determine "time-to-direct-contact" information (acoustic *tau*) from distance cues, Ashmead, Davis, and

Northington (1995) found that if listeners began walking toward a brief sound while it was still sounding, they were more accurate than if they waited until the sound ceased. Studies of blind navigation suggest that blind walkers can use the change in the accumulation of sound reflections from a wall to detect the wall's presence (Rosenblum, Gordon, & Jarquin, 2000) or to maintain a constant distance from a wall when the walkers are walking parallel to it (Ashmead, LeRoy, & Odom, 1990; Ashmead & Wall, 1999; Ashmead et al., 1998).

Research on spatial navigation has used the ability to “learn” an environment through the use of auditory targets to investigate whether spatial coding is best for sensory-based spatial modalities (vision and audition) or with verbal labels (Klatsky, Lippa, Loomis, & Golledge, 2002, 2003). A listener presented with multiple targets will perform better on a pointing task when cues are in a spatial (visual or auditory) modality than when only verbal descriptions of angle and distance are provided. However, if a person is asked to move to a new way point after training and then point to the targets, the estimates of remembered locations are more accurate if the original target was presented visually than for the auditory or the verbal target cues. The observed limitations in following verbal descriptions can be related to general human difficulties in translating perceptual sensations into numbers and vice versa and may be alleviated to a degree by a specialized training.

3.3.3 Localizability of Target Sound Sources

A sound can only be localized if it contains sufficient localization cues. Strong onset information is the best source of the binaural cues necessary for horizontal localization of unfamiliar sounds (Rakerd & Hartmann, 1985, 1986). However, in part because binaural cues are ambiguous, monaural cues are also important. Therefore, the richer the spectral content, the more easily localized the sound. Hartmann found that it was difficult to localize tonal stimuli in the presences of reverberation or noise (Hartmann, 1983, 1989). This is partly because not enough spectral information remained after the binaural information was lost¹⁰. In addition, Tran, Letowski, and Abouchacra (2000) reported that localizability¹¹ of target sounds depends on the high frequency content of the target and is improved when the target is stationary rather when it oscillates slightly around its central position. Another finding of this study was that measured accuracy of locating the target in space correlated very well with the listeners' impression of target localizability.

¹⁰Hartmann and Rakerd posited that improved localization of complex signals was attributable to the fact that a broadband signal consists of a series of impulses that provide additional temporal information beyond that available from the onset transient (precedence). They only tested horizontal localization across a small arc for which binaural cues were probably most important. It is likely that complex spectral content aids by providing more monaural and binaural cues.

¹¹Localizability is defined here as the presence of information within the sound reaching the listener's ears, which allows a human listener to identify the spatial location of the sound's source. Localization ability refers to a listener's ability to use this information. Although localization ability often depends on the localizability of a sound, it can also be affected by cognitive factors and individual ability.

Monaural cues require the presence of broadband spectral content. However, the spectral content can be absent from the sound or be uninterpretable by the human ear if the sound is too short. Duration of the sound is important because the ear is not capable of integrating the spectral information of extremely short sounds (< 100 ms) (Vliegen & Optstal, 2004). Therefore, localization in elevation of very brief sounds is poor (Hartmann & Rakerd, 1993; Hofman & Opstal, 1998; MacPherson, 2000). Longer sounds (> 500 ms) are easier to localize because they allow listeners more time to move their heads in relation to the sound and gather more information about the position of the sound source to determine its location (Fisher & Freedman, 1968).

Recent publications by Abouchacra and Letowski (2001) and Abouchacra, Emanuel, Blood, and Letowski (1998) show no effect of sound intensity on localization accuracy in the horizontal plane as long as the signal is clearly audible and not accompanied by sound reflections. However, the intensity of the sound seems to affect localization accuracy in the vertical plane (Davis & Stephens, 1974; Hartmann & Rakerd, 1993; MacPherson, 2000). The effect is especially strong for short sounds. This may be attributable to nonlinear compression by the cochlea or to spreading of activation of hair cells on the cochlea (loss of information about spectral differences because of saturation of the neural response). This explains why one of the most difficult sounds to localize is sniper fire. The firing sound is very short, loud, and elevated. Other sounds, such as the disturbance of air along the bullet's path and the impact of the bullet on a surface, can disrupt or bias this information.

Recall that monaural cues are the result of direction-dependent changes in the sound spectrum. Therefore, the listener needs to be familiar with a sound in order to be able to localize a sound in elevation or distance (Philbeck & Mershon, 2002). Determination of the elevation of and distance to unfamiliar sounds is more difficult than judgment of the horizontal position of the sound source because the monaural cues are the only cues available. However, familiarity with the sound source (auditory memory) is also important for sound localization in the horizontal plane, especially for resolving front-back confusions. People usually do not have difficulty turning their heads in a proper direction when called by a familiar voice but seem to be less precise when called by a stranger. Familiarity with various sound sources and the environment itself also provides the listener with more information about which sounds to attend to in an environment. This may not improve localization accuracy specifically, but it will reduce the cognitive load by allowing one to ignore irrelevant information and in effect, accelerate the localization process. For example, when a person spends a night in a new home, numerous sounds may be abnormally alerting. After a few nights in the same space, however, the person will become more accustomed to the sounds, and only sounds that are out of place will be noticed. A Soldier or squad of Soldiers conducting reconnaissance in an urban area is not likely to be familiar with the normal sounds or the acoustics of the environment. This will reduce SA in the best of circumstances. During the course of operations, the emotional and cognitive load will reduce the interpretability of cues even further. Therefore, some aural training in common

regional sounds may be very beneficial for the better use of monaural cues. Such specialized training may be built on the basis of an auditory skill development program FAME (familiarization, acquisition, monitoring, and evaluation), which is undergoing development at the Human Research and Engineering Directorate of the U.S. Army Research Laboratory. It is also feasible, although much more difficult, to develop a training program in which reflective characteristics of an actual space are implemented in auditory virtual reality scenarios.

4. Research Questions

The objective of this document was to list potential sources of sound localization problems in an urban environment in order to identify areas where further research would be useful. This section outlines seven potential areas for future research. This list is not meant to be all inclusive, nor is it possible or feasible to conduct simultaneous research in all these areas. However, these areas identify key issues and are intended to clarify the frame for future studies in this area.

4.1 Localizability of Typical Battle Sounds

Considerable research has been conducted to identify cues that are used to localize pure tones, complex tones, and noise bursts. Investigation of the frequency ranges used to determine elevation and reinforce binaural information has highlighted the importance of high frequency information. Very loud short sounds are difficult to localize in elevation, and these sounds are likely to be present in the urban context.

However, for several reasons, it is unclear how this information translates to the urban battlefield. One primary factor is that there is no uniform typical battlefield context. Moreover, even if there were, it would not be an ideal place for basic research because of its continuous variability and mortal danger. It would not be clear what sounds were present at any given moment and which sound information presents problems. Yet, to make basic research military relevant, some effort is needed to identify main sources and types of sounds present in specific battlefield environments as well as to identify the sounds that are ambiguous and difficult to interpret and localize.

Once these sounds are identified and isolated, further research can clarify the information present in the sound source, the disruptions present in the environment, the methods to make such sounds more apparent to the listener, and the extent to which humans are capable of localizing these sounds in given contexts.

4.2 Effect of Reverberation on Localization

Reverberation has been shown to confound the localization perception. However, reverberation in an environment is one of the physical characteristics of that environment. This means that

reverberation contains information about specific acoustic properties of the space and the main sources of reflected sounds. To what degree this information is available to and interpretable by the listener has yet to be determined. Can, for example, a listener discriminate between types of reverberation (e.g., glass, wood, brick)? Can the listener, while moving in a smoke-filled room determine whether he or she is moving toward or away from the entrance to a cave, a room, or another semi-open space?

4.3 Effect of Echoes and Flutter Echoes on Localization

Rakerd and Hartmann (1985) and Guski (1990) have studied the effect of a single reflective surface on localization ability. A systematic study of the effects of two or more reflective surfaces on auditory localization has yet to be completed. In the natural environment, the floor is almost always present as a reflective surface. When we pair this with a single wall, the effect of two surfaces can be studied. There are two three-wall configurations: the corridor (the street) and the corner, which are of great interest to MOUT tactics. Note that a listener can be positioned inside and outside the hallway (street) or the corner. Three walls and a floor form four reflective surfaces; four walls and a floor form five surfaces, and the ceiling forms a sixth surface. All these configurations affect in their specific way human perception of acoustic environment. In addition, the pattern of reflections varies with the listener's position within the room, and thus, localization performance should be measured at several locations within each of these configurations. Various coefficients of reflectivity of wall materials also need to be studied. Can, for example, a listener discriminate between sources of reflected sounds? Given a choice between several visible or non-visible surfaces, can the listener identify which surface is the source of the reflected energy? Is the visual appearance of the wall affecting auditory judgment (e.g., steel wall disguised as wood)?

Another typical acoustical circumstance is that sounds may not reach the ear via a direct pathway because of occlusions. While the localization information is expected to be disrupted by these obstacles, it is not clear whether humans are able to discriminate reflected sounds from direct sounds. In other words, can a listener recognize that the sound source location information is inaccurate?

4.4 Localizing Multiple Sounds

The battlefield is likely to contain multiple sounds, more than one of which is likely to be relevant. Most research to date has concentrated on how well a person is able to localize a single sound source in the presence of directional or non-directional noise. Measuring the ability of people to attend to multiple sounds may highlight limits attributable to attentional resources, memory, and ability to respond accurately. However, if some of the existing paradigms such as dual tasks and post-trial cueing were used, this type of question could be investigated and provide useful information.

4.5 Moving Sound and Moving Listeners

The listener and the sound source can both move. The listener may only move the head in the left-right dimension or he or she may tilt the head in a number of directions. The listener may be walking and therefore changing the whole body's position relative to the sound. Time becomes a factor because motion indicates change in position over time. The sound also changes with time. The sound may or may not be present when sound source localization is attempted. Thus, memory may become a factor for the sound or for the spatial environment and one's position in it.

Some of these facets have been examined in previous research. Although it is clear that movement of the head allows for some resolution of ambiguous information, head movements are only beneficial with sounds of longer durations. It is also clear that precise localization of a moving sound source is more difficult than that of a stationary source. What has not been tested is whether intentional whole body movement of a listener can aid in sound localization. If so, in which direction relative to the sound source should movement occur? If there is an advantage for the listener to move, does this advantage remain if both the listener and the sound source are moving? It would be beneficial to determine auditory localization accuracy during a variety of movement trajectories for the sound and the listener.

4.6 The Interaction of Auditory Localization With Vision

The interaction of visual and auditory information is very dependent on the kinds of information available and the information requirements of the task to be performed. It is quite likely that in an environment with limited visual information, auditory cues will become more important. However, given the limited resolution of auditory spatial information, a Soldier may rely on whatever visual information is available, even if such information does not correspond to the sound source. This can have serious consequences if a benign object is judged to be a threat or vice versa.

Observation of training exercises might suggest specific circumstances for investigation. Relevant questions to consider include what visual factors cause auditory localization cues to be ignored? Does this mis-attribution or "visual capture" occur as a function of cognitive load or emotional stress?

Perhaps an even more important question is whether there are circumstances in which auditory cues can capture vision and increase SA. Observation may identify sounds that signal situations requiring attention and strategies that increase SA.

4.7 Auditory Training

As presented here, the analysis of sound localization cues and the ways in which the urban terrain makes these cues ambiguous point toward familiarization with sound sources and the operational environment as the most important ways to improve MOUT effectiveness. This familiarization can be obtained through various forms of training in real and virtual environments. It is advan-

tageous to Soldier survivability and effectiveness to be familiar with a specific sound source when the source is situated in various noisy and distracting environments, at various distances from the listener, or outside and inside a building. For example, a sound source in front of a wall or behind the wall sounds very different. It is also important for Soldiers to be auditorially sensitive to the differences in gun shot sounds, vehicle signatures, and other sounds likely to be encountered in a MOUT environment. Similarly, early recognition of the sound of footsteps on a specific surface may aid in the “friend or foe” decision and in Soldiers taking appropriate action. All these abilities are subject to training. Even limited focus training in one or two aspects of auditory orientation in MOUT seems to warrant research experimentation.

5. Conclusions

Given the high cognitive load of the war fighter and the need to adapt to a changing environment, auditory training may increase SA by resolving ambiguous information and by highlighting flawed information. It is not likely that one will be able to sufficiently adapt to a new acoustic environment without occupying it for some time. However, one might be able to use “rules of thumb” for identifying sound source locations and may be aware of potential misleading cues. Knowing how to strategically position one’s head may minimize errors. Strategic walking may increase available information.

Further, by identifying the sources and limitations of localization ability in reverberant contexts, a Soldier might be able to minimize information available to enemy forces by using a strategically placed sound source to hide or mask his own noises. An important issue seems to be the Soldier’s ability to determine the urban-related specifics of a sound. Is it a direct or reflected sound? Is this sound coming from outside or from inside a building? Is this sound reflected from a glass and metal surface, a wooden surface, or the asphalt surface of a street?

Unfortunately, the currently available information is not sufficient to prescribe or test all these strategies. It is therefore important that a complete database of human auditory localization capabilities be developed. By identifying human capabilities and limits, it may be possible to identify key sound sources and develop gadgets to aid in the localization of such sounds. An example of such an aid is the helmet-mounted microphone to help Soldiers identify the direction of incoming sound (e.g., sniper detection). The need for other devices such as stethoscopes or adaptive listening aids may become apparent with further research.

6. References

- Abouchacra, K.; Letowski, T. *Localization accuracy of speech in the presence of non-directional and directional noise maskers*. Paper presented at the 17th International Congress on Acoustics, Rome, Italy 2001, 2-7 September 2001.
- Abouchacra, K. S.; Emanuel, D. C.; Blood, I. M.; Letowski, T. R. Spatial perception of speech in various signal to noise ratios. *Ear & Hearing* **1998**, *19* (4), 298-309.
- Allen, F.; Schwartz, M. The effect of stimulation of the sense of vision, hearing, taste, and smell upon the sensibility of the organs of vision. *Journal of General Physiology* **1940**, *24*, 105-121.
- Aschersleben, G.; Bertelson, P. Temporal ventriloquism: Crossmodal interaction on the time dimension: 2. Evidence from sensorimotor synchronization. *International Journal of Psychophysiology* **2003**, *50*, 157-163.
- Ashmead, D. H.; Davis, D. L.; Northington, A. Contribution of listeners' approaching motion to auditory distance perception. *Journal of Experimental Psychology-Human Perception and Performance* **1995**, *21*, 239-256.
- Ashmead, D. H.; LeRoy, D.; Odom, R. D. Perception of the relative distances of nearby sound sources. *Perception and Psychophysics* **1990**, *47* (4), 326-331.
- Ashmead, D. H.; Wall, R. S. Auditory perception of walls via spectral variations in the ambient sound field. *Journal of Rehabilitation Research and Development* **1999**, *36*, 313-322.
- Ashmead, D. H.; Wall, R. S.; B., E. S.; Ebinger, K. A.; Snook-Hill, M. M.; Guth, D. A.; Yang, X. F. Echolocation reconsidered: Using spatial variations in the ambient sound field to guide locomotion. *Journal of Visual Impairment & Blindness* **1998**, *92*, 615-632.
- Beranek, L.L. *Noise Reduction*. McGraw-Hill Book Company: New York, NY, 1960.
- Bermant, R.; Welch, R. Effect of degree of separation of visual-auditory stimulus and eye position upon spatial interaction of vision and audition. *Perceptual and Motor Skills* **1976**, *43*, 487-493.
- Bertelson, P.; Aschersleben, G. Temporal ventriloquism: Crossmodal interaction on the time dimension: 1. Evidence from auditory-visual temporal order judgment. *Journal of Psychophysiology* **2003**, *50*, 147-155.
- Bertelson, P.; Radeau, M. Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics* **1981**, *29*, 578-584.

- Blauert, J. *Spatial hearing: The psychophysics of human sound localization*. (John S. Allen, Trans.). (Revised ed.). Cambridge, MA: MIT Press, 1999.
- Braasch, J.; Hartung, K. Localization in the presence of a distracter and reverberation in the frontal horizontal plane. I. Psychoacoustical data. *Acta Acustica United with Acustica* **2002**, *88*, 942-955.
- Brungart, D. S.; Scott, K. R. The effects of production and presentation level on the auditory distance perception of speech. *Journal of the Acoustical Society of America* **2001**, *110*, 425-440.
- Burnham, R. W. A study of auditory 'brightness'. *Journal of Experimental Psychology* **1942**, *30*, 490-494.
- Butler, R. A. The bandwidth effect on monaural and binaural localization. *Hearing Research* **1986**, *21*, 67-73.
- Carlile, S.; Leong, P.; Hyams, S. The nature and distribution of errors in sound localization by human listeners. *Hearing Research* **1997**, *114*, 179-196.
- Chandler, D. W.; Grantham, D. W. Auditory spatial resolution in the horizontal plane as a function of reference angle: Microstructure of the azimuth function. *Journal of the Acoustical Society of America* **1992**, *93*, 2350-2351.
- Chason, L.; Mockovak, W. P. *The effects of sound on color intensity perception*; U.S. Air Force Academy Technical Report; Dec, 70-76, 1970.
- Chason, L. R.; Berry, G. A. *Cross-sensory effects of white noise on color perception*; USAF Academy Technical Report; May, 71-76, 1971.
- Clarke, S.; Adriani, N.; Bellmann, A. Distinct short-term memory systems for sound content and sound localization. *Neuroreport* **1998**, *9*, 3433-3437.
- Clifton, R. K.; Freyman, R. L.; Litovsky, R. Y.; McCall, D. Listeners' expectations about echoes can raise or lower echo threshold. *Journal of the Acoustical Society of America* **1994**, *95*, 1525-1533.
- Corey, J. *Correlation of vision and audition*, [message posted to discussion board by Al Bregman]. Available: www.auditory.org [1998, February 9, 1998], 1998.
- Costa, A.; Bertoldi, M. Variations of the visual field accompanying muscular exertion. *Variazioni del campo visivo da sforzo muscolare. Archivio Italiano di Psicologia* **1936**, *14*, 240-248.
- Davis, R. J.; Stephens, S.D.G. The effect of intensity on the localization of different acoustical stimuli in the vertical plane. *Journal of Sound and Vibration* **1974**, *35*, 223-229.

- Divenyi, P.; Blauert, J. On creating a precedent for binaural patterns: When is an echo an echo? In W. A. Yost & C. Watson (Eds.), *Auditory Processing of Complex Sounds*. New Jersey: Erlbaum, 1987.
- Driver, J. Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature* **1996**, *381*, 66-68.
- Farag, H.; Blauert, J.; Alim, O. A. Psychoacoustic investigations on sound-source occlusion. *Journal of the Audio Engineering Society* **2003**, *51*, 635-646.
- Fisher, H. G.; Freedman, S. J. The role of the pinna in auditory localization. In S. J. Freedman (Ed.), *The Neuropsychology of Spatially Oriented Behavior* (pp. 135-152). Homewood, IL: Dorsey Press, 1968.
- Gardner, M. B. Historical background of Haas and/or precedence effect. *Journal of the Acoustical Society of America* **1968**, *43* (6), 1243-&.
- Gelfand, S. A. Hearing: An introduction to psychological and physiological acoustics (3rd ed.). New York, NY: Marcel Dekker, 1998.
- Getzmann, S. The effect of eye position and background noise on vertical sound localization. *Hearing Research* **2002**, *169*, 130-139.
- Getzmann, S.; Lewald, J.; Guski, R. Representational momentum in spatial hearing. *Perception* **2004**, *22* (5), 591-599.
- Gilbert, G. M. Inter-sensory facilitation and inhibition. *Journal of General Psychology* **1941**, *24*, 381-407.
- Grantham. Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *Journal of the Acoustical Society of America* **1986**, *79*, 1939-1949.
- Grantham, D. W.; Hornsby, B.W.Y.; Erpenbeck, E. A. Auditory spatial resolution in horizontal, vertical, and diagonal planes. *Journal of the Acoustical Society of America* **2003**, *114*, 1009-1022.
- Griffiths, T. D.; Bench, C. J.; Frackowiak, R.S.J. Human cortical areas selectively activated by apparent sound movement. *Current Biology* **1994**, *4* (10), 892-895.
- Griffiths, T. D.; Rees, A.; Witton, C.; Shakir, R. A.; Henning, G. B.; Green, G.G.R. Evidence for a sound movement area in the human cerebral cortex. *Nature* **1996**, *383* (6599), 425-427.
- Guski, R. Auditory localization - Effects of reflecting surfaces. *Perception* **1990**, *19*, 819-830.
- Haas, H. Influence of a single echo on audibility of speech. *Journal of the Audio Engineering Society* **1972**, *20* (2), 146-159.

- Hall, D. A.; Moore, D. R. Auditory neuroscience: The salience of looming sounds. *Current Biology* **2003**, *13* (3), R91-R93.
- Handzel, A. A.; Krishnaprasad, P. S. Biomimetic sound-source localization. *IEEE Sensors Journal* **2002**, *2*, 607-616.
- Hartmann, W. M. Localization of sound in rooms IV: The Franssen effect. *Journal of the Acoustical Society of America* **1989**, *86*, 1366-1373.
- Hartmann, G. W. II. Changes in visual acuity through simultaneous stimulation of other sense organs. *Journal of Experimental Psychology* **1933**, *16*, 393-407.
- Hartmann, W. M. Localization of sound in rooms. *Journal of the Acoustical Society of America* **1983**, *74*, 1380-1391.
- Hartmann, W. M.; Rakerd, B. Auditory spectral discrimination and the localization of clicks in the sagittal plane. *Journal of the Acoustical Society of America* **1993**, *94*, 2083-2092.
- Henning, G. B. Lateralization and binaural masking-level difference. *Journal of the Acoustical Society of America* **1974**, *55* (6), 1259-1262.
- Henry, P. P.; Letowski, T. R. *Auditory localization across distance*. Paper presented at the American Auditory Society, Scottsdale, AZ, 2004.
- Hofman, P. M.; Opstal, J. V. Spectro-temporal factors in two dimensional human sound localization. *Journal of the Acoustical Society of America* **1998**, *103*, 2634-2648.
- Howard, D.; Angus, J. *Acoustics and Psychoacoustics (2nd ed.)*. Boston, MA: Focal Press, 1998.
- Hubbard, T. L. Auditory representational momentum - Surface form, direction, and velocity effects. *American Journal of Psychology* **1995**, *108* (2), 255-274.
- Ince, L. P. Effects of low-intensity acoustical stimulation on visual thresholds. *Perceptual and Motor Skills* **1968**, *26*, 115-121.
- Jack, C. E.; Thurlow, W. R. Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual & Motor Skills* **1973**, *37*, 967-979.
- Jakovlev, P. A. The effect of sound on the limits of visual color fields. *ViestniK Oftal'Mologii* **1940**, *17*, 459.
- Kitajima, N.; Yamashita, Y. Dynamic capture of sound motion by light stimuli moving in three-dimensional space. *Perceptual and Motor Skills* **1999**, *89*, 1139-1158.
- Klatsky, R. L.; Lippa, Y.; Loomis, J. M.; Golledge, R. G. Learning directions of objects specified by vision, spatial audition or auditory spatial language. *Learning and Memory* **2002**, *9*, 364-367.

- Klatzky, R. L.; Lipka, Y.; Loomis, J. M.; Golledge, R. G. Encoding, learning, and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision. *Experimental Brain Research* **2003**, *149*, 48-61.
- Kopčo, N.; Shinn-Cunningham, B. Effects of cuing on perceived location of auditory sources. In P. Sincak, J. Vascak, V. Kvasnicka, & J. Pospichal (Eds.), *Intelligent Technologies: Theory and Applications* (pp. 201-206): IOS Press, 2002.
- Kopinska, A.; Harris, L. R. Spatial representation in body coordinates: Evidence from errors in remembering positions of visual and auditory targets after active eye, head, and body movements. *Canadian Journal of Experimental Psychology* **2003**, *57*, 23-37.
- Kravkov, S. V. Changes of visual acuity in one eye under the influence of the illumination of the other or of acoustic stimuli. *Journal of Experimental Psychology* **1934**, *17*, 805-812.
- Kravkov, S. V. The influence of sound upon the light and color sensibility of the eye. *Acta Ophthalmologica. KjoBenhavn.* **1936**, *14*, 348.
- Kravkov, S. V. The effect of the intensity of the auditory stimulus on the color sensitivity of the eye. *Viestnik Oftal'Mologii* **1939a**, *15*, 100.
- Kravkov, S. V. The influence of the loudness of the indirect sound stimulus on the color sensitivity of the eye. *Acta Ophthalmologica. KjoBenhavn.* **1939b**, *17*, 324-331.
- Kravkov, S. V. Interaction of the sense organs *Vzaimodeistvie organov chuvstv*, 1948.
- Lackner, J. R. The role of posture in sound localization. *Quarterly Journal of Experimental Psychology* **1973**, *26*, 235-251.
- Leakey, D. M.; Sayers, B. M.; Cherry, C. Binaural fusion of low-frequency and high-frequency sounds. *Journal of the Acoustical Society of America* **1958**, *30* (3), 222-223.
- Letowski, T.; Mermagen, T.; Abouchacra, K. S. *Directional detection and localization of a bolt click sound in jungle- and pink-noise*. Paper presented at the Noise-Con 2004, Baltimore, Maryland 2004, July 12-14, 2004.
- Levin, D. T.; Simons, D. J. Failure to detect changes to attended objects in motion pictures, 1997.
- Levin, D. T.; Simons, D. J. Perceiving stability in a changing world: Combining shots and integrating views in motion pictures and the real world 2000.
- Levin, D. T.; Simons, D. J.; Angelone, B. L.; Chabris, C. F. Memory for centrally attended changing objects in an incidental real-world change detection paradigm. *British Journal of Psychology* **2002**, *93*, 289-302.

- Lewald, J.; Dörrscheidt, G. J.; Ehrenstein, W. H. Sound localization with eccentric head position. *Behavioral Brain Research* **2000**, *108*, 105-125.
- Little, A. D.; Mershon, D. H.; Cox, P. H. Spectral content as a cue to perceived auditory distance. *Perception* **1992**, *21* (3), 405-416.
- Loveless, N. E.; Brebner, J.; Hamilton, P. Bisensory presentation of information. *Psychological Bulletin* **1970**, *73*, 161-199.
- Lubman D. Fluctuations of sound with position in a reverberant room. *Journal of the Acoustical Society of America* **1968**, *44*, 1491-1502.
- MacPherson, E.A.; Middlebrooks, J.C. Localization of brief sounds: Effects of level and background noise. *Journal of the Acoustical Society of America* **2000**, *108*, 1834-1849.
- Makous, J. C.; Middlebrooks, J. C. Two-dimensional sound localization by human listeners. *Journal of the Acoustical Society of America* **1990**, *87*, 2188-2200.
- Maruyama, K. The effect of intersensory tone stimulation on absolute light threshold. *Tohoku Psychologica Folia* **1959**, *17*, 51-81.
- Mateeff, S.; Hohnsbein, J.; Noack, T. Dynamic visual capture: apparent auditory motion induced by a moving visual target. *Perception* **1985**, *14*, 721-727.
- McGregor, P.; Horn, A. G.; Todd, M. A. Are familiar sounds ranged more accurately? *Perceptual and Motor Skills* **1985**, *61*, 1082.
- Mershon, D. H.; Ballenger, W. L.; Little, A. D.; McMurtry, P. L.; Buchanan, J. L. Effects of room reflectance and background-noise on perceived auditory distance. *Perception* **1989**, *18*, 403-416.
- Mershon, D. H.; King, L. E. Intensity and reverberation as factors in auditory-perception of egocentric distance. *Perception & Psychophysics* **1975**, *18*, 409-415.
- Mills, A. W. On the minimum audible angle. *Journal of the Acoustical Society of America* **1958**, *30*, 237-246.
- Moore, B.C.J. An introduction to the psychology of hearing (3rd ed.), 1989.
- Nagai, M.; Kazai, K.; Yagi, A. Larger forward memory displacement in the direction of gravity. *Visual Cognition* **2002**, *9* (1-2), 28-40.
- Nielsen, S. H. Auditory distance perception in different rooms. *Journal of the Audio Engineering Society* **1993**, *41*, 755-770.
- Neuhoff, J. G. An adaptive bias in the perception of looming auditory motion. *Ecological Psychology* **2001**, *13* (2), 87-110.

- Noble, W. Auditory localization in the vertical plane: Accuracy and constraint on bodily movement. *Journal of the Acoustical Society of America* **1987**, *82*, 1631-1636.
- Oldfield, S. R.; Parker, S.P.A. Acuity of sound localisation: a topography of auditory space. I. Normal hearing conditions. *Perception* **1984a**, *13*, 581-600.
- Oldfield, S. R.; Parker, S.P.A. Acuity of sound localisation: a topography of auditory space. II. Pinna cues absent. *Perception* **1984b**, *13*, 601-617.
- Oldfield, S. R.; Parker, S.P.A. Acuity of sound localisation: a topography of auditory space. III. Monaural hearing conditions. *Perception* **1986**, *15*, 67-81.
- Perrett, S.; Noble, W. The effect of head rotations on vertical plane sound localization. *Journal of the Acoustical Society of America* **1997**, *102*, 2325-2332.
- Perrott, D. R.; Buck, V.; Waugh, W.; Strybel, T. Z. Dynamic auditory localization: Systematic replication of the auditory velocity function. *Journal of Auditory Research* **1979**, *19*, 277-285.
- Perrott, D. R.; Costantino, B.; Ball, J. Discrimination of moving events which accelerate or decelerate over the listening interval. *Journal of the Acoustical Society of America* **1993**, *93*, 1053-1057.
- Perrott, D. R.; Costantino, B.; Cisneros, J. Auditory and visual localization performance in a sequential discrimination task. *Journal of the Acoustical Society of America* **1993**, *93*, 2134-2138.
- Perrott, D. R.; Musicant, A. D. Minimum auditory movement angle - Binaural localization of moving sound sources. *Journal of the Acoustical Society of America* **1977**, *62*, 1463-1466.
- Perrott, D. R.; Pacheco, S. Minimum audible angle thresholds for broadband noise as a function of the delay between the onset of the lead and lag signals. *Journal of the Acoustical Society of America* **1989**, *85*, 2669-2672.
- Perrott, D. R.; Tucker, J. Minimum audible movement angle as a function of signal frequency and the velocity of the source. *Journal of the Acoustical Society of America* **1988**, *83*, 1522-1527.
- Philbeck, J. W.; Mershon, D. H. Knowledge about typical source output influences perceived auditory distance (L). *Journal of the Acoustical Society of America* **2002**, *111*, 1980-1983.
- Plenge, G.; Brunschen, G. Signalkenntnis und richtungsbestimmung in der medianebene bei sprach [A priori knowledge of the signal when determining the direction of speech in the median plane]. *Proceedings of the 7th International Congress on Acoustics 19-H-10*, 1971.
- Radeau, M.; Bertelson, P. Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations. *Perception and Psychophysics* **1977**, *22*, 137-146.

- Radeau, M.; Bertelson, P. Auditory-visual interaction and the timing of inputs. *Psychological Research* **1987**, *49*, 17-22.
- Rakerd, B.; Hartmann, W. M. Localization of sound in rooms, II: The effects of a single reflecting surface. *Journal of the Acoustical Society of America* **1985**, *76*, 524-533.
- Rakerd, B.; Hartmann, W. M. Localization of sound in rooms, III: Onset and duration effects. *Journal of the Acoustical Society of America* **1986**, *80*, 1695-1706.
- Rakerd B.; Hartmann W. M. Precedence effect with and without interaural differences - Sound localization in three planes. *Journal of the Acoustical Society of America* **1992**, *92*, 2296.
- Rakerd B.; Hartmann W. M. More on the precedence effect in three planes. *Journal of the Acoustical Society of America* **1994**, *95*, 2917.
- Rayleigh, Lord (J. W. Strutt, 3rd Baron of Rayleigh). On our perception of sound direction. *Philosophical Magazine* **1907**, *13*, 214-232.
- Rosenblum, L. D.; Gordon, M. S.; Jarquin, L. Echolocating distance by moving and stationary listeners. *Ecological Psychology* **2000**, *12*, 181-206.
- Saberi, K.; Perrott, D. R. Minimum audible movement angles as a function of sound source trajectory. *Journal of the Acoustical Society of America* **1990**, *88*, 2639-2644.
- Sandel, T. T.; Teas, D. C.; Feddersen, W. E.; Jeffress, L. A. Localization of sound from single and paired sources. *Journal of the Acoustical Society of America* **1955**, *27*, 842-852.
- Scharine, A. A.; Tran, P.; Binseel, M. S. *ARL Acoustic Measurements in Buildings 518 and 520 at APG*; Memorandum Report 0180; U.S. Army Research Laboratory: Aberdeen Proving Ground, 2004.
- Shams, L.; Kamitani, Y.; Thompson, S.; Shimojo, S. Visual illusion induced by sound. *Cognitive Brain Research* **2002**, *14*, 147-152.
- Shimojo, S.; Miyauchi, S.; Hikosaka, O. Visual motion sensation yielded by non-visually driven attention. *Vision Research* **1997**, *37*, 1575-1580.
- Shinn-Cunningham, B.; Kawakyu, K. *Neural representation of source direction in reverberant space*. Paper presented at the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, 2003.
- Shinn-Cunningham, B.; Ram, S. *Identifying where you are in a room: Sensitivity to room acoustics*. Paper presented at the International Conference on Auditory Display, Boston, MA 2003, 6-9 July 2003.

- Smith-Abouchacra, K. M. *Detection and localization of a target signal in multiple-source environments*. Unpublished Dissertation, Pennsylvania State University, University Park, PA, 1993.
- Somers, J. T.; Das, V. E.; Dell'Osso, L. F.; Leigh, R. J. Saccades to sounds: Effects of tracking illusory visual stimuli. *Journal of Neurophysiology* **2000**, *84*, 96-101.
- Spence, C.; Driver, J. Attracting attention to the illusory location of a sound: reflexive crossmodal orienting and ventriloquism. *Neuroreport: For Rapid Communication of Neuroscience Research* **2000**, *11*, 2057-2061.
- Stream, R. W.; Whitson, E. T.; Honrubia, V. Visual tracking of auditory stimuli. *Journal of Auditory Research* **1980**, *20*, 233-243.
- Steinberg; Snow. Physical Factors. *Bell Systems Technical Journal* **1934**, *132*, 245-258.
- Strybel, T. Z.; Manligas, C. L.; Perrott, D. R. Minimum audible movement angle as a function of the azimuth and elevation of the source. *Human Factors* **1992**, *34*, 267-275.
- Symons, J. R. The effect of various heteromodal stimuli on visual sensitivity. *Quarterly Journal of Experimental Psychology* **1963**, *15*, 243-251.
- Thomas, G. J. Experimental study of the influence of vision on sound localization. *Journal of Experimental Psychology* **1941**, *28*, 163-177.
- Tran, T. V.; Letowski, T.; Abouchacra, K. S. Evaluation of acoustic beacon characteristics for navigation tasks. *Ergonomics* **2000**, *43* (6), 807-827.
- U.S. Department of the Army. Urban operations, *Field Manual 3-06*. Washington, DC: Department of Defense, 2003.
- Vliegen, J.; Optstal, A.J.V. The influence of duration and level on human sound localization. *Journal of the Acoustical Society of America* **2004**, *115*, 1705-1713.
- Wada, Y.; Kitagawa, N.; Noguchi, K. Audio-visual integration in temporal perception. *International Journal of Psychophysiology* **2003**, *40*, 117-124.
- Wallach, H.; Newman, E. B.; Rosenzweig, M. R. The precedence effect in sound localization. *The American Journal of Psychology* **1949**, *62*, 315-336.
- Warner, H. D.; Heimstra, N. W. Effects of intermittent noise on visual search tasks of varying complexity. *Perceptual and Motor Skills* **1971**, *32*, 219-226.
- Warner, H. D.; Heimstra, N. W. Effects of noise intensity on visual target-detection performance. *Human Factors* **1972**, *14*, 181-185.

- Warren, D. H.; Welch, R. B.; McCarthy, T. J. The role of visual-auditory “compellingness” in the ventriloquism effect: Implications for transivity among the spatial senses. *Perception & Psychophysics* **1981**, *30* (6), 557-564.
- Waugh, W.; Strybel, T. Z.; Perrott, D. R. Perception of moving sounds: Velocity discrimination. *Journal of Auditory Research* **1979**, *19*, 103-110.
- Weerts, T. C.; Thurlow, W. R. The effects of eye position and expectation on sound localization. *Perception and Psychophysics* **1971**, *9*, 35-39.
- Welch, R. B.; Warren, D. H. Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin* **1980**, *88*, 638-667.
- Wightman, F. L.; Kistler, D. J. Resolution of front-back ambiguity in spatial hearing by listener and source movement. *Journal of the Acoustical Society of America* **1999**, *105*, 2841-2853.
- Zurek, P. M.; Reyman, R. L.; Balakrishnan, U. Auditory target detection in reverberation. *Journal of the Acoustical Society of America* **2004**, *115*, 1609-1620.

Appendix A. Localization Accuracy

(Oldfield & Parker, 1984)

	Horizontal Error	Vertical Error
0	4.0	6.0
10	5.0	6.5
20	5.0	7.0
30	5.0	7.0
40	5.0	6.0
50	6.5	6.5
60	6.0	6.5
70	7.0	6.0
80	6.5	7.0
90	9.0	6.0
100	8.5	6.0
110	13.0	7.0
120	14.0	7.0
130	14.0	9.5
140	18.0	10.0
150	19.0	10.5
160	18.0	13.5
170	10.5	15.0
180	9.0	10.5

	Horizontal Error	Vertical Error
40	12.0	9.5
30	11.5	9.5
20	9.5	9.5
10	8.0	9.0
0	9.0	7.5
-10	8.0	8.5
-20	9.5	8.5
-30	8.0	8.0
-40	9.0	9.0

The graph on the left gives the average absolute horizontal and vertical error for each indicated horizontal position (collapsed over elevation). The graph on the right gives the average absolute horizontal and vertical error for each indicated vertical position collapsed across all the horizontal positions.

(Butler, 1986)

Average Absolute Error (°)							
Monaural			Binaural (with/without reversals)				
Horizontal Region							
	Front-Rear	Middle (52.5-127.5°)	Total	Front-Rear	Middle (52.5-127.5°)	Total	
Bandwidth*	2 kHz	51	35	44	42/10	29/16	36
	4 kHz	46	17	31	23/9	22/13	22
	6 kHz	44	17	30	9/8	17/12	13
	8 kHz	36	15	26	10/9	13/11	12

*noise centered at 8 kHz

This graph gives monaural and binaural average absolute localization error as a function of frequency bandwidth.

(Makous & Middlebrooks, 1990)

Average Absolute Error (Horizontal°/Vertical°)										
Horizontal Position										
	0°	20°	40°	60°	80°	100°	120°	140°	160°	
Vertical Position	45°	3.5/10.2	6.0/8.4	8.7/7.8					16.3/11.4	
	25°	2.4/6.0	5.0/5.4	7.4/7.0	8.1/6.8		8.1/5.9	15.9/11.3	14.5/19.1	
	5°	2.2/4.1	3.8/4.6	6.1/3.9	9.7/4.2	5.2/3.2	6.2/6.5	10.3/7.3	12.0/7.3	15.6/9.2
	-5°	1.5/3.5	3.7/4.1	5.5/3.4	9.4/4.0	5.5/4.6	7.4/7.2	10.4/7.8	13.0/8.0	12.7/6.9
	-25°	2.5/6.5	5.2/8.0	5.2/5.9	7.1/5.4			9.6/7.4	10.7/7.9	12.8/7.9
	-45°	3.6/7.5	6.5/7.6						9.4/10.5	11.6/15.7

This table gives the average absolute error for a number of horizontal and vertical positions. These values do not include error attributable to front-back confusions (6% of trials). The average horizontal standard deviation was 3.62 degrees and the average vertical standard deviation was 3.17 degrees.

(Carlile, Leong, & Hyams, 1997)

Subjects were trained to localize by pointing their noses at the sound source. Average localization errors were on the order of 3 degrees in azimuth and 4 degrees in elevation. These numbers do not include error attributable to front-back confusions (3.2% of cases). They found no up-down confusions and the majority of front-back errors were between 60 and 120 degrees' azimuth.

Appendix B. Minimum Audible Angle (MAA)

(Mills, 1958)

		Reference Azimuth			
MAA		0°	30°	60°	75°
Frequencies (Hz)	250	1.0	1.5	3.5	6.5
	500	1.0	1.5	3.5	7.5
	750	1.0	1.5	3.5	8.0
	1000	1.0	1.5	3.5	8.5
	1250	1.5	3.0	5.0	n/a
	1500	3.0	5.0	n/a*	n/a
	1750	3.1	6.0	n/a	n/a
	2000	3.0	6.5	n/a	n/a
	3000	2.0	6.5	5.0	6.5
	4000	2.0	5.0	10.5	6.5
	6000	2.0	5.0	n/a	n/a
	8000	4.0	12.5	n/a	n/a
	10000	2.5	3.8	n/a	n/a

*Unable to measure MAA

This table gives the MAA for the indicated pure tone signals at each of several reference azimuth positions. Listeners were asked to judge whether the second sound came from left or right of reference sound.

(Perrott & Pacheco, 1989)

		MAA
Interstimulus Delay (ms)	0	4.3
	1	4.7
	5	4.3
	10	3.6
	20	3.5
	30	2.9
	40	3.4
	50	2.4
	75	2.2
	100	1.5
	150	0.9
	200	1.2

The MAA was measured as a function of inter-stimulus delay. The azimuthal reference position was 0 degrees. An inter-stimulus delay of 0 ms indicates that the two signals were simultaneous. Data show that 100 to 150 ms need to resolve spatial information. Very short delays fuse together as a single moving perception.

(Saberri & Perrott, 1990)

MAA	
0°	0.97
10°	0.78
20°	0.9
30°	0.9
40°	0.9
50°	1.06
60°	1.06
70°	1.24
80°	1.8
90°	3.65

The MAA was measured for sounds varying in both azimuth *and* elevation. The experimental setup included 30 loudspeakers on boom that were rotated in 10-degree increments. The speaker array angle indicates the angle of the tilt of the boom. The azimuthal reference point was 0 degrees.

(Grantham, Hornsby, & Erpenbeck, 2003)

MAA		Stimulus Bandwidth		
		wideband	highpass	lowpass
Speaker Array Angle	Horizontal	1.6	1.6	1.5
	Diagonal	2.8	4.1	3.4
	Vertical*	6.5	11.6	19.0

*Only 6 of 20 were able to do the vertical task.

The task in this experiment differed from the one used in Saberri and Perrott in that stimulus items were recorded by a KEMAR manikin placed in three positions relative to the loudspeaker array. These KEMAR recordings were then presented to the listener, and Saberri and Perrott made estimates from these recordings. The azimuthal reference point was 0 degrees.

Appendix C. Minimum Audible Movement Angle (MAMA)

(Perrott & Musicant, 1977)

Velocity (°/s)	MAMA (degrees)	Threshold duration (ms)
90	8.3	92.2
180	12.9	71.7
360	21.2	58.9

Listeners were asked to localize the onset and offset of a 500-Hz sine wave with different durations. The sound was moved to the right or the left, always beginning at 0 degrees' azimuth. The perceived onsets were shifted in direction of movement. The perceived offsets also tended to be shifted, but the size of the shift depended on the duration of the signal. Therefore, the average signal duration required to accurately direct direction of movement is also provided as a function of velocity.

(Perrott & Tucker, 1988)

MAMA	Velocity (°/s)		
	8-16	32-64	128
500	3.7	6.0	7.4
730	2.6	3.8	7.0
950	3.1	4.2	6.0
1170	5.0	5.5	7.2
1500	6.5	7.6	8.5
1800	6.1	8.6	11.0
2250	5.7	7.5	8.4
3000	6.6	8.5	9.3
3700	5.2	6.6	8.7

This table gives the MAMA measured as a function of velocity and frequency. Because the difference in the MAMA for 8 and 16 (°/s) or 32 and 64 (°/s) were not significantly different, average values are given here.

(Chandler & Grantham, 1992)

		Velocity (deg/s)						Minimum integration time (ms)
		0 (MAA)	10	20	45	90	180	
MAMA								
		@ 0° azimuth						
Frequency (Hz)	500	2.8	5.9	6.7	9	11.7	15.9	292
	1000	2.4	8.2	10.4	10.3	13.0	18.9	294
	3000	8.4	11.7	17.1	15.6	19.4	25.1	335
	5000	5.0	9.5	12.7	15.8	18.1	20.9	295
	1/3 octave*	2.3	8.6	10.2	11.7	17.2	22.9	420
	1 octave*	2.3	6.7	7.8	11.3	14.2	19.0	352
	wideband	1.2	5.2	5.7	8.2	12.0	17.3	362
		@ 60° azimuth						
	3000 Hz	11.3	12.7	23.0	19.2	34.3	52.1	1082
	1/3 octave*	8.7	12.6	20.0	32.5	41.9	65.0	1455
	1 octave*	11.0	13.5	19.7	24.5	32.5	43.1	812
	wideband	1.5	8.3	11.0	16.6	20.3	28.4	551

(centered at 3 kHz)

This experiment measured the MAMA of listeners from two reference azimuth positions (0 and 60 degrees) for pure tones and several bandwidths. The task differed in that the signal was moving to the right or was stationary and the task was to indicate if the sound had ended in the “same” or “different” location.

(Strybel, Manligas, & Perrott, 1992)

MAMA (degrees)		
Horizontal Position	-80	3.4
	-40	2.2
	-20	1.8
	-10	1.3
	0	1.1
	10	1.7
	20	1.3
	40	1.8
	80	3.1

This table gives the MAMA as measured from azimuthal reference points from -80 to 80 degrees.

(Grantham, Hornsby, & Erpenbeck, 2003)

		MAMA (degrees)		
		wideband	highpass	lowpass
Array Orientation*	horizontal	4.2	5.4	4.5
	diagonal	7.3	11.4	7.2
	vertical	15.3	22.6	...

*Orientation with respect to the KEMAR head

The task in this experiment differed from other experiments in that stimulus items were recorded by a KEMAR manikin placed in three positions relative to the loudspeaker array. These KEMAR recordings were then presented to the listener and Grantham et al. made estimates from these recordings. The azimuthal reference point was 0 degrees.

...No MAMA could be obtained in this condition.

INTENTIONALLY LEFT BLANK

Appendix D. Signal-to-Noise Ratio Needed for Localization

Percent correct localization of a target in the presence of non-directional noise as a function of SNR.

(Abouchacra, Emanuel, Blood, & Letowski, 1998)

Type of Noise		Signal to Noise Ratio											Quiet	
		Angle	-18	-15	-12	-9	-6	-3	0	3	6	9		12
Target Sound "Northwest" (60 dB)	diffuse white noise	45	9	20	36	43	59	67	66	75	76	88	88	97
		90	46	57	69	70	82	82	89	89	89	96	96	100
		135	15	34	37	45	45	45	45	60	58	64	77	85
		180	2	12	36	76	96	96	96	96	96	96	96	100

For 50% correct or better in all directions (errors within ± 22.5 degrees) requires an SNR of -9 dB.

(Letowski, Mermagen, & Abouchacra, 2004)

Target Sound		Direction of Background Noise*							
		Angle	-18	-15	-12	-9	-6	-3	0
Target Sound bolt-click	Quiet (30 dB)	overall	4	14	24	39	58	66	68
	Pink (70 dB)	overall	0	3	5	22	41	56	71
	Jungle (70 dB)	overall	6	8	13	33	53	66	75

For 50% correct or better (errors within ± 15 degrees) requires an SNR of at least -4 to -7 dB.

Percent correct localization of a target in the presence of noise as a function of sensation level.

(Smith-Abouchacra, 1993)

Direction of Background Noise*		Target Sensation level				
		Angle	0	6	12	18
Target Sound "Northwest"	Nondirectional (65 dB A-wtd)	0	25	37	90	90
		45	25	25	55	60
		90	28	53	80	88
		135	8	15	26	34
		180	15	29	50	57

(Abouchacra & Letowski, 2001)

		Angle				
		0	6	12	18	
Target Sound "Northwest"	Nondirectional	overall	25	39	90	90
	0°	overall	38	50	94	96
	45°	overall	25	33	96	99
	90°	overall	29	50	86	98
	135°	overall	22	27	74	92
	180°	overall	21	36	71	91

*For both studies, the background noise was speech shaped and presented at 65 dB A-wtd.

In both studies, a 9 dB SL was required for 50% correct or better (errors within ± 15 degrees).

<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>	<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>
*	ADMINISTRATOR DEFENSE TECHNICAL INFO CTR ATTN DTIC OCA 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218 *pdf version only	1	AFRL/HECB ATTN DR DAVID DARKOW 2610 7TH ST WRIGHT PATTERSON AFB OH 45433-7901
1	DIRECTOR US ARMY RSCH LABORATORY ATTN IMNE ALC IMS MAIL & REC MGMT 2800 POWDER MILL RD ADELPHI MD 20783-1197	1	MAJ CYNTHIA EADES 3MDG/SGOSLA 10435 KUTER AVE ELMENDORF AFB AK 99506
1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL CI OK TL TECH LIB 2800 POWDER MILL RD ADELPHI MD 20783-1197	1	BRIAN WIXOM 30725 DEL REY ROAD TEMECULA CA 92591
1	BEN MOZO COMMUNICATIONS & EAR PROTECTION PO BOX 311174 3700 SALEM ROAD ENTERPRISE AL 36331-1174	2	TEMCO COMMUNICATIONS INC ATTN SIDRIC NAKAJO BRIAN MURAKAMI 13 CHIPPING CAMPDEN DR SOUTH BARRINGTON IL 60010
1	US ARMY AEROMEDICAL RSCH LAB DIRECTOR AIRCREW PROTECTION DIV ATTN WM A AHROON 6901 ANDREWS AVENUE PO BOX 620577 FORT RUCKER AL 36362-0577	4	US ARMY NATICK SOLDIER CENTER ATTN AMSSB RTP D KRASNECKY (3 CYS) MATT CORREA BLDG 5 RM 116 15 KANSAS STREET NATICK MA 01760-5018
1	MINE SAFETY APPLIANCES COMPANY ATTN RAY CLARKE 121 GAMMA DRIVE PITTSBURGH PA 15238	1	AIR FORCE RSCH LABORATORY LEAD JSF VIBROACOUSTICS ATTN RICH MCKINLEY WRIGHT PATTERSON AFB OH 45433-7901
3	KAISER ELECTRO-OPTICS ATTN JIM MELZER 2752 LOKER AVE WEST CARLSBAD CA 92008-6603	1	USAMRMC C STRATEGIC PARTNERSHIP OFC ATTN COL N L VAUSE BLDG 810 RM B36 FT DETRICK MD 21701
1	DEFENCE RSCH AND DEVELOPMT CANADA HUMAN-COMPUTER INTERACTION GROUP ATTN G R ARRABITO 1133 SHEPPARD AVE WEST PO BOX 2000 TORONTO ONTARIO M3M 3B9 CANADA	3	MICRO ANALYSIS AND DESIGN ATTN JOHN MILANSKI DAVE DAHN SUE ARCHER 4949 PEARL EAST CIRCLE STE 300 BOULDER CO 80301
1	JONATHAN I LEE 2LT USAF BAO INTEGRATED HEADGEAR PM AFRL/HECB RESEARCH PHYSICIST 2610 SEVENTH ST WRIGHT PATTERSON AFB OH 45433-7901	1	ARTISENT INC ATTN DAVID ROGERS 374 CONGRESS ST BOSTON MA 02210
		3	US ARMY NATICK SOLDIER CTR ERGONOMICS TEAM/SSTD ATTN J B SAMPSON 100 KANSAS ST NATICK MA 01760-5020

<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>	<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>
1	RDECOM-NATICK SOLDIER CTR MANAGEMENT TEAM/PEO SOLDIER/PM FUTURE FORCE WARRIOR TECH PRO OFC ATTN C FITZGERALD 100 KANSAS ST NATICK MA 01760-5020	1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MH C BURNS BLDG 1002 ROOM 117 1ST CAVALRY REGIMENT RD FT KNOX KY 40121
1	RDECOM-NATICK SOLDIER CTR HUMAN PERF/MANPRINT/SSI LEAD FUTURE FORCE WARRIOR TECH PRO OFC ATTN C L BLACKWELL 100 KANSAS ST NATICK MA 01760-5020	1	ARMY RSCH LABORATORY - HRED AVNC FIELD ELEMENT ATTN AMSRD ARL HR MJ D DURBIN BLDG 4506 (DCD) RM 107 FT RUCKER AL 36362-5000
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR M DR M STRUB 6359 WALKER LANE SUITE 100 ALEXANDRIA VA 22310	1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MK MR J REINHART 10125 KINGMAN RD FT BELVOIR VA 22060-5828
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MA J MARTIN MYER CENTER RM 2D311 FT MONMOUTH NJ 07703-5630	1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MV HQ USAOTC S MIDDLEBROOKS 91012 STATION AVE ROOM 111 FT HOOD TX 76544-5073
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MC A DAVISON 320 MANSCEAN LOOP STE 166 FT LEONARD WOOD MO 65473-8929	1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MY M BARNES 2520 HEALY AVE STE 1172 BLDG 51005 FT HUACHUCA AZ 85613-7069
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MD T COOK BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290	1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MP D UNGVARSKY BATTLE CMD BATTLE LAB 415 SHERMAN AVE UNIT 3 FT LEAVENWORTH KS 66027-2326
1	COMMANDANT USAADASCH ATTN ATSA CD ATTN AMSRD ARL HR ME MS A MARES 5800 CARTER RD FT BLISS TX 79916-3802	1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR M DR B KNAPP ARMY G1 MANPRINT DAPE MR 300 ARMY PENTAGON ROOM 2C489 WASHINGTON DC 20310-0300
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MI J MINNINGER BLDG 5400 RM C242 REDSTONE ARSENAL AL 35898-7290	1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MJK MS D BARNETTE JFCOM JOINT EXPERIMENTATION J9 JOINT FUTURES LAB 115 LAKEVIEW PARKWAY SUITE B SUFFOLK VA 23435
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MM DR V RICE BLDG 4011 RM 217 1750 GREELEY RD FT SAM HOUSTON TX 78234-5094	1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MQ M R FLETCHER US ARMY SBCCOM NATICK SOLDIER CTR AMSRD NSC SS E BLDG 3 RM 341 NATICK MA 01760-5020
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MG R SPINE BUILDING 333 PICATINNY ARSENAL NJ 07806-5000		

<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>	<u>NO. OF COPIES</u>	<u>ORGANIZATION</u>
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MT DR J CHEN 12350 RESEARCH PARKWAY ORLANDO FL 32826-3276	1	OFC OF NAVAL RESEARCH MEDICAL/BIOLOGICAL SCI & TECH DIV ATTN R D SHILLING ARLINGTON VA 22217-5860
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MS MR C MANASCO SIGNAL TOWERS RM 303A FORT GORDON GA 30905-5233	1	DRDC TORONTO ARMY LIAISON OFC ATTN STEPHEN BOYNE 1133 SHEPPARD AVE WEST PO BOX 2000 TORONTO ON M3M 3B9 CANADA
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MU M SINGAPORE 6501 E 11 MILE RD MAIL STOP 284 BLDG 200A 2ND FL RM 2104 WARREN MI 48397-5000		<u>ABERDEEN PROVING GROUND</u>
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MF MR C HERNANDEZ BLDG 3040 RM 220 FORT SILL OK 73503-5600	1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL CI OK (TECH LIB) BLDG 4600
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MW E REDDEN BLDG 4 ROOM 332 FT BENNING GA 31905-5400	1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL CI OK S FOPPIANO BLDG 459
1	ARMY RSCH LABORATORY - HRED ATTN AMSRD ARL HR MN R SPENCER DCSFDI HF HQ USASOC BLDG E2929 FORT BRAGG NC 28310-5000	1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL HR MR F PARAGALLO BLDG 459
1	DR THOMAS M COOK ARL-HRED LIAISON PHYSICAL SCIENCES LAB PO BOX 30002 LAS CRUCES NM 88003-8002	1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL HR MB J HAWLEY BLDG 459
1	DIRECTOR UNIT OF ACTION MANEUVER BATTLE LAB ATTN ATZK UA BLDG 1101 FORT KNOX KY 40121	1	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL HR MR T HADUCH BLDG 459
2	SENSORY DEVICES INC ATTN H HOLSOPPLE W PIROTH 205 MAIN ST NEW EAGLE PA 15067	20	DIRECTOR US ARMY RSCH LABORATORY ATTN AMSRD ARL HR SD A SCHARINE BLDG 520 APG AA
1	DOUGLAS BRUNGART SENIOR COMPUTER ENGINEER HUMAN EFFECTIVENESS DIREC 2610 SEVENTH ST WRIGHT PATTERSON AFB OH 45433-7901		