Spring 5-16-2020

Development and Human Factors Evaluation of a Portable Auditory Localization Acclimation Training System

Brandon Thompson
brandon.thompson@westpoint.edu

Development and Human Factors Evaluation of a Portable Auditory Localization Acclimation Training System

Brandon Scott Thompson

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Industrial and Systems Engineering

John G. Casali, Chair
Joseph L. Gabbard
Brian M. Kleiner
Kichol Lee

May 4, 2020
Blacksburg, Virginia

Keywords: Auditory Localization, Auditory Situation Awareness, Auditory Training, Hearing Protection

Development and Human Factors Evaluation of a Portable Auditory Localization Acclimation Training System

Brandon S. Thompson

ABSTRACT

Auditory situation awareness (ASA) is essential for safety and survivability in military operations where many of the hazards are not immediately visible. Unfortunately, the Hearing Protection Devices (HPDs) required to operate in these environments can impede auditory localization performance. Promisingly, recent studies have exhibited the plasticity of the human auditory system by demonstrating that training can improve auditory localization ability while wearing HPDs, including military Tactical Communications and Protective Systems (TCAPS). As a result, the U.S. military identified the need for a portable system capable of imparting auditory localization acquisition skills at similar levels to those demonstrated in laboratory environments. The purpose of this investigation was to develop and validate a Portable Auditory Localization Acclimation Training (PALAT) system equipped with an improved training protocol against a proven laboratory grade system referred to as the DRILCOM system and subsequently evaluate the transfer-of-training benefit in a field environment.

In Phase I, a systems decision process was used to develop a prototype PALAT system consisting of an expandable frame housing 32-loudspeakers operated by a user-controlled tablet computer capable of reproducing acoustically accurate localization cues similar to the DRILCOM system. Phase II used a within-subjects human factors experiment to validate whether the PALAT system could impart similar auditory localization training benefits as the DRILCOM system. Results showed no significant difference between the two localization training systems at each stage of training or in training rates for the open ear and with two

TCAPS devices. The PALAT system also demonstrated the ability to detect differences in localization accuracy between listening conditions in the same manner as the DRILCOM system. Participant ratings indicated no perceived difference in localization training benefit but significantly preferred the PALAT system user interface which was specifically designed to improve usability features to meet requirements of a user operable system. The Phase III investigation evaluated the transfer-of-training benefit imparted by the PALAT system using a broadband stimulus to a field environment using gunshot stimulus. Training under the open ear and in-the-ear TCAPS resulted in significant differences between the trained and untrained groups from in-office pretest to in-field posttest.

Development and Human Factors Evaluation of a Portable Auditory Localization Acclimation Training System

Brandon S. Thompson

GENERAL AUDIENCE ABSTRACT

Auditory situation awareness (ASA) is essential for safety and survivability in military operations where many of the hazards are not immediately visible. Unfortunately, the Hearing Protection Devices (HPDs) required to operate in these environments can impede sound localization performance. Promisingly, recent studies have exhibited the ability of the human auditory system to learn by demonstrating that training can improve sound localization ability while wearing HPDs. As a result, the U.S. military identified the need for a portable system capable of improving sound localization performance at similar levels to those demonstrated in laboratory environments. The purpose of this investigation was to develop and validate a Portable Auditory Localization Acclimation Training (PALAT) system equipped with an improved training protocol against a proven laboratory grade system referred to as the DRILCOM system and subsequently evaluate the transfer-of-training benefit in a field environment.

In Phase I, a systems decision process was used to develop a prototype PALAT system consisting of an expandable frame housing 32-loudspeakers operated by a user-controlled tablet computer capable of reproducing similar sounds as the DRILCOM system. Phase II used a within-subjects human factors experiment to validate whether the PALAT system could impart similar sound localization training benefits as the DRILCOM system. Results showed no significant difference between the two localization training systems at each stage of training or in

training rates for the open ear and with two HPDs. The PALAT system also demonstrated the ability to detect differences in localization accuracy between listening conditions in the same manner as the DRILCOM system. Participant ratings indicated no perceived difference in localization training benefit but significantly preferred the PALAT system user interface which was specifically designed to improve usability features to meet requirements of a user operable system. The Phase III investigation evaluated the transfer-of-training benefit imparted by the PALAT system using a broadband stimulus to a field environment using gunshot stimulus. Training under the open ear and in-the-ear TCAPS resulted in significant differences between the trained and untrained groups from in-office pretest to in-field posttest.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation for Dr. John G. Casali for his willingness to assume the role of my advisor and for his confidence in my abilities to pursue a degree in Human Factors Engineering. Words cannot convey my gratitude for his mentoring and guidance over the entire course of this research endeavor. Without his support this dissertation would not have been possible.

I would also like to extend my sincere appreciation to my committee members Dr. Joseph L. Gabbard, Dr. Brian M. Kleiner, and Dr. Kichol Lee for encouraging and challenging me think beyond the military applications of this research. Their advice and support were critical in successfully completing this dissertation process. A special thank you is also owed to Dr. Lee for his expertise, accessibility, and willingness to provide hands-on training in the laboratory and for coding the computer software program that served as the backbone for this research.

To Lieutenant Colonel Kara M. Cave, thank you for serving as both a mentor and co-researcher. Your knowledge, expertise, and friendship were immeasurable throughout this journey.

I am extremely thankful for the faculty and staff in the Virginia Tech ISE program, especially the incredibly talented laboratory instrument makers in the ISE shop. Mark, Scott, Joe, Kelly, Travis, and especially Randy, thank you for sharing your tradecraft with me and for turning my ideas into a functioning system.

The experiments discussed herein were supported in part by funding from the Office of Naval Research, Noise-Induced Hearing Loss Program, with contract monitors Kurt Yankaskas and Kristy Hentchel. The views expressed herein are those of the author, and do not represent those of any branch of the United States military.

To my family and friends, thank you for your love and support for not only me but also for my family. My family and I are forever grateful for the new friendships we made in the Blacksburg community. Thank you for making our experience and time at Virginia Tech feel like home.

I cannot begin to thank my family, April, Taylor, Addyson, Kynley, and Riley enough for their unwavering love and encouragement, not only throughout this process, but during my entire military career. April, thank you for all of the sacrifices you have made to allow me to pursue my dreams. Taylor, Addyson, Kynley, and Riley, thank you for providing me with the motivation to continue to grow.

PREFACE

To meet the objective of improving and field-validating an auditory localization training protocol and system, experiments comprising this research program were conducted in three phases. Given the breadth of this objective, the experiments were covered in two dissertations. This dissertation covered Phase II and sought to develop a portable auditory localization acclimation training system capable of imparting similar auditory localization training effects demonstrated by a full scale, laboratory-grade system. LTC Kara Cave (U.S. Army) conducted Phase I whereby an auditory localization training protocol was developed to be incorporated as part of the training and testing during subsequent phases. Phase III was the combined effort of both authors of Phase I and Phase II. As such, Phase III in this document is duplicative in this dissertation and that of LTC Cave's, and is included with the knowledge and assent of the faculty who comprise both students' advisory committees.

TABLE OF CONTENTS

ABSTRACT .... **ii**

GENERAL AUDIENCE ABSTRACT .... **iv**

ACKNOWLEDGEMENTS .... **vi**

PREFACE .... **vii**

GLOSSARY .... **xii**

LIST OF TABLES .... **xvi**

LIST OF FIGURES .... **xxxix**

CHAPTER 1. Introduction **1**

1.1 Purpose **9**
1.2 Background **12**
1.2.1 Human Auditory System 12
1.2.2 Human Auditory Sensitivity 14
1.2.3 Auditory Localization 16
1.2.4 Distance Judgments 44
1.2.5 Auditory Factors in the Military 46
1.2.6 Hearing Protection Devices 51
1.2.7 Auditory Perceptual Skills Acquisition 56
1.2.8 Auditory Localization Apparatus Designs 64
1.3 Research Gaps **68**
1.4 Research Objectives and Hypotheses **70**
1.4.1 Objectives 70
1.4.2 Hypotheses 71

CHAPTER 2. Design of a Portable Auditory Localization Acclimation Training (PALAT) System **73**

2.1 PALAT Objectives **73**
2.2 PALAT System Requirements **74**
2.3 PALAT Structure **75**
2.4 PALAT Loudspeaker Evaluation **88**
2.5 PALAT Auditory Equipment **107**
2.6 PALAT Conclusions and Implications for Phases II and III **114**

CHAPTER 3. Phase II: Development and In-Laboratory Investigation of a Portable Auditory Localization Acclimation Training System ... 117

3.1 Phase II Objectives ... 117
3.2 Phase II Methodology ... 117
3.3 Phase II Experimental Design ... 119
3.3.1 Independent Variables (IVs) ... 121
3.3.2 Dependent Measures ... 137
3.3.3 Participants ... 144
3.4 Phase II Apparatus ... 146
3.5 Phase II Experimental Procedure ... 149
3.5.1 Recruitment and screening ... 150
3.5.2 Calibration and setup ... 151
3.5.3 Training and testing sessions ... 153
3.6 Phase II Results ... 161
3.6.1 Outlier Analysis ... 162
3.6.2 Phase II Objective Measures Overview and Data Graphs for Initial Visual Inspection ... 163
3.6.3 Phase II Objective Measures Statistical Analyses ... 170
3.6.4 Phase II Subjective Measures Statistical Analyses ... 232
3.7 Phase II Discussion and Conclusions ... 271
3.8 Implications for the Phase III Experiment ... 294

CHAPTER 4. Phase III In-Field Investigation of Transfer-of Training ... 296

4.1 Phase III Objectives ... 296
4.2 Phase III Methodology ... 296
4.3 Phase III Experimental Design ... 298
4.3.1 Independent Variables (IVs) ... 300
4.3.2 Dependent Measures ... 309
4.3.3 Participants ... 315
4.4 Phase III Apparatus ... 317
4.4.1 In-Office: PALAT System ... 317
4.4.2 In-Field Site ... 323
4.5 Phase III Experimental Procedures ... 332
4.5.1 Recruitment and screening ... 332
4.5.2 In-office pretest ... 332
4.5.3 Training Session (Experimental group only) ... 338
4.5.4 In-field posttest ... 341
4.6 Phase III Results ... 344
4.6.1 Outlier Analysis ... 344
4.6.2 Objective performance ... 346
4.6.3 Subjective ratings ... 369
4.7 Phase III Conclusions: In-Field Investigation of Transfer-of-Training ... 404
4.7.1 Listening Condition Conclusions ... 405
4.7.2 Training Effect Conclusions ... 406

CHAPTER 5. Implications of the Results ... 426

5.1 Limitations of the Research ... 426
5.2 Results explained as a possible function of TCAPS design variables ... 428

5.3 Recommendations for Efficiency and Effectiveness in Training with the Portable Auditory Localization Acclimation Training (PALAT) System 430
5.3.1 Recommendations for use of the PALAT in various room environments 430
5.3.2 Recommendations for use of the PALAT System as setup and deployed by a trainee 431
5.3.3 Recommendations for use of the PALAT System given time likely available and time likely required 432
5.4 Implications for military implementation 434
5.4.1 Relevance to TCAPS design, selection, and procurement 434
5.4.2 Relevance to ground combat service member duties and mission 436
5.4.3 Implications for NIHL reduction 437
5.4.4 Cost-Benefit of Implementing the PALAT system 439
5.5 Summary of Applications and Recommendations for Implementation of the PALAT System 440

Final Conclusions 444

REFERENCES 447

Appendix A. Review of Auditory Localization Apparatus Designs 458

Appendix B. Loudspeaker Alternative Screening 459

Appendix C. Subject Matter Expert Questionnaire 461

Appendix D. PALAT System Parts List and Costs 492

Appendix E. Human Subjects IRB Documents 493

Appendix F. Phase II Participant Questionnaire 503

Appendix G. Phase II Screening Form 506

Appendix H. Phase II Recruiting Flyer 507

Appendix I. Phase II Script 508

Appendix J. Phase III IRB Approval 510

Appendix K. Phase III Questionnaire 513

Appendix L. Phase III Screening Form 515

Appendix M. Remote Firing Device Wiring Diagram 516

Appendix N. Phase III Participant Flyer 517

Appendix O. Phase III Informed Consent 518

Appendix P. Phase III Participant Instructions 528

Appendix Q. Figures of statistically non-significant findings included in qualitative analysis. 532

GLOSSARY

Air Force Instruction (AFI): documented regulations and standards for members of the United States Air Force.

Analysis of Variance (ANOVA): collection of statistical procedures that compare the means of several groups to determine if they are equal.

American National Standards Institute (ANSI): organization that oversees the creation, promulgation and use of norms and guidelines for businesses including acoustical devices.

Army Regulation (AR): documented regulations and standards for members of the United States Army.

Attenuation: the reduction of sound pressure level in decibels as achieved by a device; for hearing protection devices (HPDs), this is typically taken to be the difference in decibel levels of a sound at a listener's hearing threshold, as heard with and without an HPD. This is termed an "insertion loss" attenuation measurement.

Auditory Fitness for Duty (AFFD): standards of hearing thresholds and profiles that dictate whether an individual is able to perform their duties safely and effectively.

Back-front (errors) (BF): auditory localization error where the sound is perceived as originating from in back of, or behind, the listener when the sound actually is presented from the front.

Behind-the-ear (BTE): hearing aids that are placed behind the ear with a small tube that connects the hearing aid to an earpiece in the ear canal.

Decibel (dB): logarithmic unit used to measure the intensity of sound, used in this dissertation study to refer to a decibel measurement with no frequency weighting, sometimes called dB(linear) or dBZ.

Decibel, A-weighted (dBA): decibel weighting filters that approximates the relative loudness of sounds as perceived by the human ear.

Decibel, Peak (dBP): highest level of sound pressure produced in a certain period of time, taken as the peak pressure value converted to the logarithmic decibel scale, with no time-averaging or weighting applied.

Difference Limen (DL): degree of difference needed for an observer to detect a difference between two stimuli, or a change in one stimuli, at least half the time.

Department of Defense (DoD): the executive branch department of the United States responsible for oversight of the U.S. Armed Forces.

Detection, Recognition, Identification, Localization, and Communication (DRILCOM): test battery designed by the Virginia Tech-Auditory Systems Laboratory to test auditory situation awareness.

Front-back reversal errors (FB): auditory localization error where the sound is perceived as originating from in front of the listener when the sound actually is presented from in back of the listener.

Hearing Center of Excellence (HCE): DoD agency charged with improving the prevention, diagnosis, mitigation, treatment and rehabilitation of hearing loss in the military.

Hearing Protection Device (HPD): device worn in or over the ear to protect against noise hazards.

Head-related transfer function (HRTF): the transformation of a sound wave, boosting and attenuating frequencies, as a result of the listener's size and shape of the head, torso, ears, and ear canals.

Hemi-anechoic chamber: an environment where the walls and ceiling are insulated with sound-absorbent materials while the ground consists of a hard, reflective surface.

Hertz (Hz): unit of measure for frequency of sound defined as the number of wave cycles per second.

Interaural level difference (ILD): difference in loudness and frequency distribution between the two ears.

Interaural phase difference (IPD): difference in phase of a sound wave that reaches each ear.

Interaural spectrum difference (ISD): spectral differences in sounds that arrive at the ear drum due to physical differences between the pinna and head/torso.

Interaural time difference (ITD): difference in arrival time of sound between the two ears.

In-the-ear (ITE): hearing protection device or hearing aid that is worn inside the external ear canal.

Just noticeable difference (JND): degree of difference needed for an observer to detect a difference between two stimuli, or a change in one stimuli, at least half the time.

Minimum Audible Angle (MAA): smallest discernable difference in horizontal angle between two sound sources.

Minimum Audible Field (MAF): method of measuring monaural absolute hearing threshold level in a sound field with stimulus presented by loudspeakers.

Minimum Audible Movement Angle (MAMA): the minimum horizontal angle of travel required for detection of the direction of sound movement.

Minimum Audible Pressure (MAP): method of measuring binaural absolute hearing threshold level in a sound field with stimulus presented by headphones.

Median plane: the mid-sagittal plane that bisects the body vertically through the midline, dividing the body exactly in half (left and right sides).

Midline: coplanar with the median plane, a line that bisects the body, vertically dividing the body exactly in half.

Millisecond (msec): one-thousandth of a second.

Noise-Induced Hearing Loss (NIHL): hearing impairment resulting from exposure to loud noise.

Noise-induced permanent threshold shift (NIPTS): a permanent change in the absolute auditory threshold due to overexposure to hazardous noise.

Noise-induced temporary threshold shift (NITTS): a temporary change in the absolute auditory threshold due to overexposure to hazardous noise.

Noise Reduction Rating (NRR): unit of measure in decibels of the amount of potential attenuation afforded by a hearing protection device.

Office of Naval Research (ONR): agency that coordinates, executes, and promotes science and technology research for the U.S. Navy.

Occupational Safety and Health Administration (OSHA): U.S. Department of Labor agency charged with setting and enforcing occupational safety standards.

Portable Auditory Localization Acclimation Training (PALAT) system: small, portable auditory localization apparatus used for training and testing azimuthal localization performance.

Reverberation Time 60 (RT60): the measure of time after a sound source stops that it takes for the sound pressure level to decrease by 60 decibels.

Semi-reverberant room: an environment free not insulated with sound-absorbent materials where acoustic reflections occur.

Sensation level (SL): unit of measure indicating the number of decibels above the hearing threshold level.

Signal-to-noise ratio (SNR): unit of measure expressed in plus or minus decibels of the difference between a sound signal level and a background noise level.

Sound Pressure Level (SPL): logarithmic base 10 measurement in pascal, Pa, of the ratio of the pressure of a sound wave compared to ambient atmospheric pressure (re 20 μPa), typically referenced in dB.

Tactical Communication and Protective System (TCAPS): active hearing protection device that incorporates a communication capability; the device can enhance quiet sounds and cut-off loud, hazardous noise.

Tactical Communication and Protective System-Lite (TCAPS-Lite): active hearing protection device that does not incorporate a communication capability; the device can enhance quiet sounds and cut-off loud, hazardous noise.

Transfer Function of the Open Ear (TFOE): the spectral changes of a sound wave from a free field environment to the eardrum that are produced by the external ear.

Temporary threshold shift (TTS): a temporary change in the hearing threshold relative to a baseline level.

Unity gain: the state where the electronic gain control of an active hearing protection device is set to overcome or offset the passive attenuation of the protector and provide as close to natural hearing as possible (Casali & Lee 2016a).

Virginia Tech Auditory Systems Laboratory (VT-ASL): facility that studies the principles and methods of human factors engineering, ergonomics, and acoustics applied to human hearing problems.

LIST OF TABLES

Table 1. Overall % correct in horizontal plane sound localization. Effect of ear and helmet conditions (from Abel, 2009). 22

Table 2. Localization in azimuth accuracy at midline by signal type (adapted from Blauert, 1997). 36

Table 3. Hearing Profile Criteria for the U.S. Air Force, Army, and Navy. 48

Table 4. Steady-state and impulse noise levels from various military equipment types (adapted from Melzer et al., 2012, Figure 9.1). 50

Table 5. PALAT system's specified and implied requirements. 75

Table 6. Visual field of view search distances associated with 30° and 45° azimuthal separation for military threats originating from their effective range distances (USMC, 2017). 78

Table 7. Loudspeaker design parameters and engineering specifications. 88

Table 8. Manufacturer-reported frequency response. 91

Table 9. Manufacturer-reported data by loudspeaker alternative. 100

Table 10. Subject Matter Expert Acoustics and Audiology Team Experience. 102

Table 11. Example SME criteria pairwise comparison and weighting coefficient score. 103

Table 12. Example of SME pairwise comparison of alternatives and alternative choice coefficient score. 104

Table 13. Example of SME alternative scores by performance measure and overall alternative score. 105

Table 14. Wilcoxon results pairwise comparisons between rankings of each loudspeaker alternative on alternative score (bolded text in the table indicates a significant test result at *p*<0.05.) 106

Table 15. Participant study order by sex, training system order (random assignment based on arrival order), and listening condition by session (counterbalanced using a repeating 3 x 6 Latin square). 121

Table 16. Mean noise floor and reverberation time (RT60) measurements of the DRILCOM and PALAT rooms, as measured in 1/3 octave-band frequency. 130

Table 17. Sound pressure level (SPL) of 70 dBA pink noise measured using KEMAR manikin for three sets of TEP-100 devices and mean (by left and right ear designation). Levels compared with the open ear 77.6 dBA reference level displayed as the delta (Δ), open ear – TEP-100. 134

Table 18. Sound pressure level (SPL) of 70 dBA pink noise measured using KEMAR manikin for three sets of ComTac™ III devices and mean. Levels compared with the open ear 77.6 dBA reference level displayed as the delta (Δ), open ear – ComTac™ III. 136

Table 19. Mean pure-tone hearing level thresholds (dBHL) for all participants. 145

Table 20. DRILCOM (left column) and PALAT (right column) apparatus comparison. 149

Table 21. Sample statistic *Q* for dependent measures at LU0 and LU5 by listening condition using 95% confidence interval. Dixon's r_{10} critical statistic for n=24 is Q=0.34. 163

Table 22. Mauchly's test of sphericity for full factorial ANOVA for the effect of training system and listening condition on *absolute score* at LU0 (pretest). 173

Table 23. Full factorial ANOVA table evaluating **differences in *absolute score* at LU0 (pretest)** according to training system and listening condition (*bold text indicates significant values at the α=0.05 significance level). 173

Table 24. Paired-samples *t*-test pairwise **comparisons between listening conditions on *absolute score* at LU0 (pretest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 174

Table 25. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the DRILCOM system on *absolute score* at LU0 (pretest). 177

Table 26. One-way ANOVA table evaluating **differences in *absolute score* at LU0 (pretest)** according to listening condition while using the DRILCOM system (*bold text indicates significant values at the α=0.05 significance level). 177

Table 27. Paired-samples *t*-test pairwise **comparisons between listening conditions while using the DRILCOM system on *absolute score* at LU0 (pretest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 177

Table 28. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the PALAT system on *absolute score* at LU0 (pretest). 179

Table 29. One-way ANOVA table evaluating **differences in *absolute score* at LU0 (pretest) according to listening condition while using the PALAT system** (*bold text indicates significant values at the α=0.05 significance level). 179

Table 30. Paired-samples *t*-test pairwise **comparisons between listening conditions while using the PALAT system on *absolute score* at LU0 (pretest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 179

Table 31. Mauchly's test of sphericity for full factorial ANOVA for the effect of training system and listening condition on *absolute score* at LU5 (posttest)........ 181

Table 32. Full factorial ANOVA table evaluating **differences in *absolute score* at LU5 (posttest)** according to training system and listening condition (*bold text indicates significant values at the α=0.05 significance level). 181

Table 33. Paired-samples *t*-test pairwise **comparisons between listening conditions on *absolute score* at LU5 (posttest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 182

Table 34. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the DRILCOM system on *absolute score* at LU5 (posttest). 184

Table 35. One-way ANOVA table evaluating **differences in *absolute score* at LU5 (posttest) according to listening condition while using the DRILCOM system** (*bold text indicates significant values at the α=0.05 significance level). 184

Table 36. Paired-samples *t*-test pairwise **comparisons between listening conditions while using the DRILCOM system on *absolute score* at LU5 (posttest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 184

Table 37. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the PALAT system on *absolute score* at LU5 (posttest)................................ 186

Table 38. One-way ANOVA table evaluating **differences in *absolute score* at LU5 (posttest) according to listening condition while using the PALAT system** (*bold text indicates significant values at the α=0.05 significance level). 186

Table 39. Paired-samples *t*-test pairwise **comparisons between listening conditions while using the PALAT system on *absolute score* at LU5 (posttest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 186

Table 40. Mauchly's test of sphericity for full factorial repeated-measures ANOVA for the effect of training system, listening condition, and stage of training on *absolute score*................ 188

Table 41. Full-factorial ANOVA table evaluating **differences in *absolute score* according to training system, listening condition, and stage of training (LU0 and LU5)** (*bold text indicates significant values at the α=0.05 significance level)... 189

Table 42. Paired-samples *t*-test pairwise **comparisons between listening conditions on *absolute score* across both training systems and stages of training (LU0 and LU5)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 190

Table 43. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and LU5** across all listening conditions (*bold text indicates significant values at the α=0.05 significance level)... 194

Table 44. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and LU5** across all listening conditions (*bold text indicates significant values at the α=0.05 significance level)... 194

Table 45. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and the PALAT system at LU5** across all listening conditions (*bold text indicates significant values at the α=0.05 significance level)......... 195

Table 46. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and the DRILCOM system at LU5** across all listening conditions (*bold text indicates significant values at the α=0.05 significance level)......... 195

Table 47. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system under the open ear condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).. 197

Table 48. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system under the open ear condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).. 197

Table 49. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and the PALAT system at LU5 under the open ear condition** (*bold text indicates significant values at the α=0.05 significance level). 197

Table 50. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and the DRILCOM system at LU5 under the open ear condition** (*bold text indicates significant values at the α=0.05 significance level). 198

Table 51. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system under the TEP-100 condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).. 199

Table 52. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system under the TEP-100 condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).. 200

Table 53. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and the PALAT system at LU5 under the TEP-100 condition** (*bold text indicates significant values at the α=0.05 significance level). 200

Table 54. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and the DRILCOM system at LU5 under the TEP-100 condition** (*bold text indicates significant values at the α=0.05 significance level). 200

Table 55. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system under the ComTac™ III condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level)..................................... 202

Table 56. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system under the ComTac™ III condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level)... 202

Table 57. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and the PALAT system at LU5 under the ComTac™ III condition** (*bold text indicates significant values at the α=0.05 significance level). 203

Table 58. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and the DRILCOM system at LU5 under the ComTac™ III condition** (*bold text indicates significant values at the α=0.05 significance level). 203

Table 59. **Comparisons of *slope* for each listening condition between training systems (DRILCOM and PALAT) from LU0 (pretest) through LU5 (posttest)**....................... 205

Table 60. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the DRILCOM system on *slope* from LU0 through LU5. 205

Table 61. One-way repeated-measures ANOVA table evaluating **differences in *slope* from LU0 through LU5 on the DRILCOM system by listening condition** (*bold text indicates significant values at the α=0.05 significance level). 206

Table 62. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the PALAT system on *slope* from LU0 through LU5. 206

Table 63. One-way repeated-measures ANOVA table evaluating **differences in *slope* from LU0 through LU5 on the PALAT system by listening condition** (*bold text indicates significant values at the α=0.05 significance level). 207

Table 64. Linear regression equations for *absolute score* by training system and listening condition from LU0 (pretest) through LU5 (posttest). 208

Table 65. Spearman correlation for percent accuracy while using the PALAT system and the DRILCOM system by listening condition. 209

Table 66. Mauchly's test of sphericity for full factorial repeated-measures ANOVA for the effect of training system, listening condition, and stage of training on *front-back reversal errors* (*bold text indicates significant values at the α=0.05 significance level). 214

Table 67. Full-factorial ANOVA table evaluating **differences in *front-back reversal errors* according to training system, listening condition, and stage of training (LU0 and LU5)**. Greenhouse-Geisser corrected *p*-value evaluated for significance at α=0.05 (*bold text indicates significant values at the α=0.05 significance level). 215

Table 68. Paired-samples *t*-test pairwise **comparisons between listening conditions on *front-back reversal errors*** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 216

Table 69. Mauchly's test of sphericity for full factorial repeated-measures ANOVA for the effect of training system, listening condition, and stage of training on *response time*. 220

Table 70. Full-factorial ANOVA table evaluating **differences in *response time* according to training system, listening condition, and stage of training (LU0 and LU5)** (*bold text indicates significant values at the α=0.05 significance level). 221

Table 71. Paired-samples *t*-test pairwise **comparisons between listening conditions on *response time* across both training systems at LU0 and LU5** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 222

Table 72. Mauchly's test of sphericity for DRILCOM system two-way repeated-measures ANOVA for the effect of listening condition and stage of training on *response time*. 224

Table 73. Two-way repeated-measures ANOVA table evaluating **differences in *response time* for the DRILCOM system according to listening condition and stage of training** (*bold text indicates significant values at the α=0.05 significance level). 224

Table 74. Paired-samples *t*-test pairwise **comparisons between listening conditions on *response time* on the DRILCOM system** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 225

Table 75. Mauchly's test of sphericity for PALAT system two-way repeated-measures ANOVA for the effect of listening condition and stage of training on *response time*. 226

Table 76. Two-way repeated-measures ANOVA table evaluating **differences in *response time* for the PALAT system according to listening condition and stage of training** (*bold text indicates significant values at the α=0.05 significance level). 227

Table 77. Paired-samples *t*-test pairwise **comparisons between listening conditions on *response time* on the PALAT system** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level). 227

Table 78. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 1, perceived confidence (bolded text in the table indicates a significant test result at p<0.05.) 236

Table 79. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 1, perceived confidence (bolded text in the table indicates a significant test result at p<0.05.) 237

Table 80. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 2, loudspeaker proximity impacts (bolded text in the table indicates a significant test result at p<0.05.) 239

Table 81. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 2, loudspeaker proximity impacts (bolded text in the table indicates a significant test result at p<0.05.) 240

Table 82. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 3, ease of use (bolded text in the table indicates a significant test result at p<0.05.) 243

Table 83. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 3, ease of use (bolded text in the table indicates a significant test result at p<0.05.) 244

Table 84. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 4, room effects on localization (bolded text in the table indicates a significant test result at p<0.05.) 246

Table 85. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 1, room effects on localization (bolded text in the table indicates a significant test result at p<0.05.) 247

Table 86. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 5, training improvements (bolded text in the table indicates a significant test result at p<0.05.) 249

Table 87. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 5, training impacts (bolded text in the table indicates a significant test result at p<0.05.) 250

Table 88. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 6, difficulty judging signal location (bolded text in the table indicates a significant test result at p<0.05.) 253

Table 89. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 6, difficulty judging signal location (bolded text in the table indicates a significant test result at p<0.05.) 254

Table 90. Wilcoxon results pairwise comparisons between each listening condition for Question 6, difficulty judging signal location (bolded text in the table indicates a significant test result at p<0.017.) 255

Table 91. Wilcoxon pairwise comparisons between each listening condition on DRILCOM system for Question 6, difficulty judging signal location (bolded text in the table indicates a significant test result at p<0.05.) 256

Table 92. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 7, reaction time (bolded text in the table indicates a significant test result at p<0.05.) 258

Table 93. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 7, impact on reaction time (bolded text in the table indicates a significant test result at p<0.05.) 259

Table 94. Wilcoxon results pairwise comparisons between each listening condition for Question 7, impact on reaction time (bolded text in the table indicates a significant test result at p<0.05 with Bonferroni correction α= 0.017.) 260

Table 95. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 8, User interface (bolded text in the table indicates a significant test result at p<0.05.) 262

Table 96. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 8, impact of user interface on localization (bolded text in the table indicates a significant test result at p<0.05.) 263

Table 97. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 9, Room environment impact on reaction time (bolded text in the table indicates a significant test result at p<0.05.) 266

Table 98. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 9, room environment impact on reaction time (bolded text in the table indicates a significant test result at $p<0.05$.) 267

Table 99. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 10, Loudspeaker visibility impact of localization (bolded text in the table indicates a significant test result at $p<0.05$.) 269

Table 100. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 10, impact of loudspeaker visibility on localization (bolded text in the table indicates a significant test result at $p<0.05$.) 270

Table 101. Participant study order by sex, group assignment (random assignment based on arrival order), and listening condition (counterbalanced using a repeating 3 x 6 Latin square). 300

Table 102. Participant demographics by group. 301

Table 103. Sound pressure level (SPL) of 70 dBA pink noise measured using KEMAR manikin for three sets of TEP-100 devices and mean (by left and right ear designation). Levels compared with the open ear 77.6 dBA reference level displayed as the delta (Δ), open ear – TEP-100. 305

Table 104. Sound pressure level (SPL) of 70 dBA pink noise measured using KEMAR manikin for three sets of ComTac™ III devices and mean. Levels compared with the open ear 77.6 dBA reference level displayed as the delta (Δ), open ear – ComTac™ III. 307

Table 105. Mean pure-tone hearing level thresholds (dBHL) for all participants and by group.317

Table 106. Mean noise floor and reverberation time (RT60) measurements of the DRILCOM and PALAT rooms, as measured in octave bands. 319

Table 107. Magnetic azimuth (degrees), radial distance (feet), and sound pressure level (dBA max) at center position of each remote firing device location. 324

Table 108. Weather conditions during the field test. 343

Table 109. Visual field of view search distances associated with 30° and 45° azimuthal separation for military threats originating from their effective range distances (USMC, 2017) 348

Table 110. Mauchly's Test of Sphericity and MANOVA results evaluating the effects of training group, listening condition, and stage of training on absolute score, ballpark score, and response time (bolded text indicates a significant test result at the α=0.05 significance level). 350

Table 111. Univariate ANOVA results for each dependent measure for each independent variable (bold text indicates significant values at the α=0.10 significance level). 351

Table 112. Pairwise comparisons for listening condition using the absolute correct score. 353

Table 113. Listening condition significant pairwise comparisons using the **ballpark** correct score. 354

Table 114. Listening condition significant pairwise comparisons using the **response time** score. 354

Table 115. Simple-effect *F*-tests for trained group versus untrained group at each stage of training using the absolute correct score. 357

Table 116. Simple-effects *F*-tests for each group examining pretest versus posttest performance collapsed across listening conditions using the absolute correct score. 358

Table 117. Simple-effect *F*-tests examining absolute correct score differences at **pretest** in trained versus untrained participants for each listening condition (open ear, TEP-100, and ComTac™ III) 359

Table 118. Simple-effect *F*-tests examining absolute correct score differences at **posttest** in trained versus untrained participants for each listening condition (open ear, TEP-100, and ComTac™ III) 360

Table 119. Simple-effect *F*-tests examining absolute correct score differences at pretest versus posttest for the **untrained** group. Different letters indicate significant differences at $p \leq 0.10$. 361

Table 120.Simple-effect *F*-tests examining absolute correct score differences at pretest versus posttest for the trained group. 362

Table 121. Mauchly's test of sphericity for mixed ANOVA for the effect of group, listening condition, and stage of training on **front-back errors** using the **120-degree arc criterion**. 364

Table 122. Mixed-factor ANOVA table evaluating differences in front-back reversal errors using the 120-degree arc criterion according to group, listening condition, and stage of training. 364

Table 123. Significant pairwise comparisons between listening conditions for front-back reversal errors using the 120° arc criterion with a Bonferroni adjustment........ 365

Table 124. ANOVA table for pretest prediction of posttest absolute correct scores in the trained group, TEP-100 condition........ 367

Table 125. Descriptive statistics and independent-samples *t*-tests, using the 0.033 corrected alpha level comparing group differences within each listening condition measured by the slope of the regression line from pretest to posttest for absolute correct score. 368

Table 126. Wilcoxon results comparing ratings for pretest versus posttest for each listening condition and group for Question 1, Perceived Confidence. 371

Table 127. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions on ratings of confidence for Question 1, Perceived Confidence. 372

Table 128. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 1, Perceived Confidence. 372

Table 129. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at pretest for Question 1, Perceived Confidence. 373

Table 130. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at posttest for Question 1, Perceived Confidence. 373

Table 131. Friedman test results demonstrating significant differences among listening conditions for the trained and untrained groups at pretest for Question 1, Perceived Confidence. 375

Table 132. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at pretest for Question 1, Perceived Confidence. 375

Table 133. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at pretest for Question 1, Perceived Confidence. 375

Table 134. Friedman test results demonstrating significant differences among listening conditions for the trained and untrained groups at posttest for Question 1, Perceived Confidence. 376

Table 135. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at posttest for Question 1, Perceived Confidence. 376

Table 136. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at posttest for Question 1, Perceived Confidence. 376

Table 137. Wilcoxon signed-ranks results comparing confidence ratings for pretest versus posttest for each listening condition and group for Question 2, Perceived Accuracy. 377

Table 138. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 2, Perceived Accuracy. 378

Table 139. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 2, Perceived Accuracy. 378

Table 140. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at pretest for Question 2, Perceived Accuracy. 378

Table 141. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at posttest for Question 2, Perceived Accuracy. 379

Table 142. Friedman test results demonstrating significant differences in perceived ratings among listening conditions for the trained and untrained groups at pretest, Perceived Accuracy. 380

Table 143. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at pretest for Question 2, Perceived Accuracy. 381

Table 144. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at pretest for Question 2, Perceived Accuracy. 381

Table 145. Friedman test results demonstrating significant differences in perceived accuracy ratings among listening conditions for the trained and untrained groups at posttest for Question 2, Perceived Accuracy. 381

Table 146. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at posttest for Question 2, Perceived Accuracy. 381

Table 147. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at posttest for Question 2, Perceived Accuracy. 382

Table 148. Wilcoxon signed-ranks results comparing difficulty ratings for pretest versus posttest for each listening condition and group for Question 3, Perceived Difficulty. 382

Table 149. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 3, Perceived Difficulty. 383

Table 150. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 3, Perceived Difficulty. 383

Table 151. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at pretest for Question 3, Perceived Difficulty........... 384

Table 152. Results of the Mann-Whitney *U* test evaluating group differences at posttest for the TEP-100 condition for Question 3, Perceived Difficulty. 384

Table 153. Friedman test results demonstrating significant differences in perceived difficulty ratings among listening conditions for the trained and untrained groups at pretest for Question 3, Perceived Difficulty.. 385

Table 154. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at pretest for Question 3, Perceived Difficulty. 386

Table 155. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at pretest for Question 3, Perceived Difficulty. 386

Table 156. Friedman test results demonstrating significant differences in ratings among listening conditions for the trained and untrained groups at posttest for Question 3, Perceived Difficulty. 386

Table 157. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at posttest for Question 3, Perceived Difficulty. 386

Table 158. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at posttest for Question, Perceived Difficulty. 387

Table 159. Wilcoxon signed-ranks results comparing response time for pretest versus posttest for each listening condition and group for Question 4, Perceived Reaction Time. 388

Table 160. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 4, Perceived Reaction Time. 388

Table 161. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 4, Perceived Reaction Time. 388

Table 162. Results of the Mann-Whitney *U* test evaluating group differences at pretest for each listening condition for Question 4, Perceived Reaction Time. 389

Table 163. Results of the Mann-Whitney *U* test evaluating group differences at posttest for each listening condition for Question 4, Perceived Reaction Time. 389

Table 164. Friedman test results demonstrating significant among listening conditions for the trained and untrained groups at pretest Question 4, Perceived Reaction Time. 390

Table 165. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 4 at pretest for the trained group at pretest, Perceived Reaction Time. 391

Table 166. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 4 at pretest for the untrained group, Perceived Reaction Time. 391

Table 167. Friedman test results demonstrating significant differences in ratings among listening conditions for the trained and untrained groups at posttest, Question 4, Perceived Reaction Time. 391

Table 168. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 4 for the trained group at posttest, Perceived Reaction Time. 392

Table 169. Table Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for the untrained group at posttest for Question 4, Perceived Reaction Time. 392

Table 170. Wilcoxon signed-ranks results comparing comfort ratings for pretest versus posttest for each listening condition and group for Question 5, Perceived Comfort. 393

Table 171. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 5, Perceived Comfort.................................... 394

Table 172. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 5, Perceived Comfort.................................... 394

Table 173. Results of the Mann-Whitney *U* test evaluating group differences at pretest for each listening condition for Question 5, Perceived Comfort. 395

Table 174. Results of the Mann-Whitney *U* test evaluating group differences at posttest for each listening condition for Question 5, Perceived Comfort. 395

Table 175. Friedman test results demonstrating significant differences in perceived comfort ratings among listening conditions for the trained and untrained groups at pretest, Question 5, Perceived Comfort. 396

Table 176. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 5 at pretest for the trained group (α=0.016) for Question 5, Perceived Comfort. 397

Table 177. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 5 at pretest for the untrained group (α=0.016) for Question 5, Perceived Comfort. 397

Table 178. Friedman test results demonstrating significant differences in comfort ratings among listening conditions for the trained and untrained groups at posttest for Question 5, Perceived Comfort. 397

Table 179. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 5 at posttest for the trained group for Question 5, Perceived Comfort. 397

Table 180. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at posttest for the untrained group for Question 5, Perceived Comfort. 398

Table 181. Wilcoxon signed-ranks results comparing likelihood of maintaining the current listening condition given a similar task for pretest versus posttest for each listening condition and group for Question 6, Question 6. Likelihood of Wearing Device. 399

Table 182. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 6, Likelihood of Wearing Device. 399

Table 183. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 6, Likelihood of Wearing Device. 399

Table 184. Results of the Mann-Whitney *U* test evaluating group differences at pretest for each listening condition for Question 6, Likelihood of Wearing Device. 400

Table 185. Results of the Mann-Whitney *U* test evaluating group differences at posttest for each listening condition for Question 6, Likelihood of Wearing Device. 400

Table 186. Friedman test results demonstrating significant differences among listening conditions for the trained and untrained groups at pretest for Question 6, Likelihood of Wearing Device. 401

Table 187. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at pretest for the trained group, for Question 6, Likelihood of Wearing Device. 402

Table 188. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at pretest for the untrained group, for Question 6, Likelihood of Wearing Device. 402

Table 189. Friedman test results demonstrating significant differences among listening conditions for the trained and untrained groups at posttest, for Question 6, Likelihood of Wearing Device. 402

Table 190. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at posttest for the trained group (α=0.016) for Question 6, Likelihood of Wearing Device. 402

Table 191. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at posttest for the untrained group (α=0.016) for Question 6, Likelihood of Wearing Device. 403

Table 192. Friedman two-way analysis of variance results for listening condition for the trained group by ranks for Question 7, Degree of Preparedness. 403

Table 193. Wilcoxon signed-ranks results comparing perceived degree of preparedness for each listening condition for Question 7, Degree of Preparedness.. 404

Table 194. PALAT System: Applications and Recommendations for Implementation............ 441

LIST OF FIGURES

Figure 1. Mean enemy camp detection distances across HPDs, with 95% confidence interval plotted. Mean values, shown on bars, with the same letter are not significantly different according to Tukey's test at p<0.05 (adapted from Casali et al., 2009, Figure 4)................. 4

Figure 2. Effects of HPDs at 60 dBA (left graph) and 90 dBA (right graph) background noise level on percentage correct localization; mean values shown on bars with 95% confidence intervals. Means with the same letter are not significantly different at p ≤ 0.05. (adapted from Alali & Casali, 2011, Figures 9 & 10). 5

Figure 3. The effect of listening condition on mean response time. Error bars are the 95% confidence interval about the mean. Mean number shown above bars. Top letters show Tukey's multiple comparisons test for main effects of listening conditions (adapted from Talcott et al., 2012, Figure 5)... 6

Figure 4. Network diagram of human hearing subtasks involved in achieving and maintaining auditory situation awareness (adapted from Casali & Lee, 2016a, Figure 1) Copyright 2012 by John G. Casali. 7

Figure 5. Mean absolute localization error for open ear and training electronically modified hearing protection device, in-the-ear (A) and over-the-ear (B). Within pairs of means, both across Training Day (horizontal) and across Listening Condition (vertical), means with same letter are not significantly different at p 0.05 using a paired sample *t*-test with Bonferroni correction (adapted from Casali & Robinette, 2014, Figures 3 & 4). 8

Figure 6. Number of correct responses (absolute (ABS) correct and ballpark correct) out of 36 total possible, under each testing condition by learning units (OPEN = open ear, A and B represent two TCAPS devices) (adapted from Casali & Lee, 2016b). 9

Figure 7. Cross-sectional view of the human ear (adapted from Samsung, 2016). 13

Figure 8. Human auditory envelope with areas of music (black) and speech (gray inside of black) (adapted from Scharine et al., 2009, Figure 11-3). 15

Figure 9. Anthropometric diagram with body planes and coordinate system used for sound localization. (adapted from Arazi, 2017, Figure 2; Moore, 2004). 16

Figure 10. Pinna effects on frequencies shown by variations of average sound pressure level at the human eardrum as a function of azimuth θ of incident sound for various frequencies (adapted from Shaw, 1974a, Figures 6 & 7). 20
Figure 11. Effects of listening condition on percent front-back errors. Means and 95% confidence interval error bars. Top letters are main effects of listening condition. (adapted from Talcott et al., 2012, Figure 4). 21
Figure 12. Mean absolute and algebraic azimuthal localization error for each azimuth position (adapted from Oldfield & Parker, 1984, Figure 3). 22
Figure 13. Results of applying the twist model fit to the localization data for four hearing conditions: A The control condition (both ears free); B a unilateral mold in the right ear; C a unilateral mold in the left ear; D molds bilaterally applied (adapted from Hofman & Van Opstal, Figure 4A-D). 23
Figure 14. Horizontal head-related transfer function for sound sources located at 45° and 315° in the right ear (adapted from Gelfand, 2010, Figure 3.2, derived from Shaw, 1974b). 25
Figure 15. Interaural level difference (ILD) and interaural time difference (ITD) created by a sound arriving from a 45° azimuth angle (adapted from Scharine et al., 2009, Figure 11-29, and Kapralos et al., 2008, Figure 1). 27
Figure 16. Variations in intensity level as a sound source is rotated in a horizontal plane around the head (adapted from Steinberg & Snow, 1934, Figure 2). 28
Figure 17. Interaural time differences plotted as a function of azimuth (adapted from Scharine et al., 2009, Figure 11-31). 29
Figure 18. Average Minimum Audible Angle as a function of stimulus frequency. The parameter θ is the azimuth of the reference tone pulse (adapted from Mills, 1958, Figure 5). 30
Figure 19. The concept of the cone of confusion. 31
Figure 20. Intensity Difference Limen as a function of sensation level for 4000, 7000, and 10000 Hz frequencies. Sensation level plotted in units of 10 log10(E/E0) where E0 is the threshold intensity (adapted from Riesz, 1928, Figure 3). 32
Figure 21. Frequency Difference Limen as a function of frequency for pure tones presented at 20 and 80 dB SL (adapted from Scharine et al., 2009, Figure 11-5). 32
Figure 22. Spectral changes in sound intensity across frequencies from 1000 to 12000 Hz presented from different levels of elevation (adapted from Alali, 2011). 34

Figure 23. Azimuth localization test result, measured as absolute and ballpark (within ±15°) percent correct response rate at Low signal level of 50 dBA in 40 dBA pink noise background and High signal level of 85 dBA in 75 dBA pink noise background, SNR +10 (adapted from Casali & Lee, 2016a). 41

Figure 24. Aggregate of detection on approach, detection lost on retreat and identification with percent worse labeled as compared to open ear (adapted from Clasing & Casali, 2014, Figure 7). 46

Figure 25. Effect of listening condition by noise level on percent correct response within 22.5° (“ballpark”). 95% CI error bars with means labeled above. (adaptedfrom Talcott et al., 2012, Figure 2). 55

Figure 26. Azimuth localization test results, measured as percent worse than open ear (adapted from Casali & Lee, 2016a, Figure 45). 56

Figure 27. Combined training groups localization performance for open ear and training TCAPS. Means with same letter are not significantly different at $p < 0.05$. (adapted from Casali & Robinette, 2014, Figure 2). 62

Figure 28. Localization apparatus designs, circular (left adapted from Alali & Casali, 2011), semi-circular horizontal and vertical (top right adapted from Noble et al., 1990b), sound boom (bottom right adapted from Abouchacra & Letowski, 2012). 65

Figure 29. VT-ASL DRILCOM test apparatus located in the hemi-anechoic test room. 66

Figure 30. Army Research Laboratory Sphere Room (Left) and Dome Room (Right) (adapted from Letowski et al., 2012b). 67

Figure 31. SoLoArc localization system (from Westerberg et al., 2016). 68

Figure 32. Initial PALAT system design concepts: A) Individual support poles with multiple loudspeaker mounts, B) Circular array base with vertical loudspeaker poles, and C) Expandable accordion style frame. 83

Figure 33. Initial PALAT system frame measurements and custom aluminum joint measurements. 85

Figure 34. Fully expanded PALAT system frame with 32 loudspeakers. 86

Figure 35. Collapsed PALAT system frame (left), three loudspeaker mount (top center), two loudspeaker elevation mount and single loudspeaker mount (bottom center), and elevation loudspeaker mount (right). 87

Figure 36. PALAT system entry/exit spring-loaded swivel gate loudspeaker mount. 87
Figure 37. Loudspeaker alternative finalists: A) Bose® FreeSpace® 3 Satellite Speakers, B) Cambridge Audio Minx Min 12, and C) Boston Acoustics® SoundWare XS. 89
Figure 38. On-Axis frequency response – stepped sine (dB SPL 1 watt at 1 meter)................... 93
Figure 39. On-Axis frequency response – sine sweep (dB SPL 1 watt at 1 meter). 93
Figure 40. Total harmonic distortion reference to the fundamental frequency (dB). 95
Figure 41. Total harmonic distortion reference to the fundamental frequency (%). 95
Figure 42. Sound pressure level of dissonant tone signal tone at 55 dBA.................................. 97
Figure 43. Sound pressure level deviations from DRILCOM loudspeaker for dissonant tone signal tone at 55 dBA.. 97
Figure 44. Sound pressure level of dissonant tone signal tone at 80 dBA.................................. 98
Figure 45. Sound pressure level deviations from DRILCOM loudspeaker for dissonant tone signal tone at 80 dBA.. 98
Figure 46. Mean criterion weighting coefficient for each performance measure criterion. Error bars are the 95% confidence interval about the mean. Numbers above the error bars are means. 103
Figure 47. Mean alternative scores for each loudspeaker alternative. Means above 95% confidence interval error bars. Different letters indicate significant differences at $p<0.05$ with Bonferroni adjustment. 107
Figure 48. PNG-400 Pink Noise Generator (top) mounted on top of a Stewart Audio AV30MX-2 amplifier (bottom). 108
Figure 49. Audio components mounted to a custom audio rack (top left and right) and mounted in audio equipment bag (bottom). 110
Figure 50. PALAT system wiring diagram.. 111
Figure 51. Spectral content of the pink noise masker produced by the DRILCOM system (green dash-dot line) and two alternative pink noise loudspeaker locations for the PALAT system, two Cambridge Audio Minx Min 12 loudspeakers positioned on vertical uprights at 11 o'clock and 1 o'clock (red dashed line) and one Cambridge Audio Minx Min 12 loudspeaker positioned under the trainee chair facing forward (purple dotted line) in 1/3 octave-band frequency. Overall sound pressure level of 55 dBA for the pink noise. 113

Figure 52. Microsoft® Surface Pro tablet user interface and audio case (left) and a trainee operating the PALAT system via tablet user interface (right). 114
Figure 53. Schematic of the PALAT system (left semicircle) compared to the DRILCOM system (right semicircle) with radial and signal loudspeaker separation distances. 116
Figure 54. Phase II experimental design order. 119
Figure 55. Experimental design with repeated-measures subject assignment for Phase II, with independent variables, experimental order, participant assignment, and dependent measures listed. 120
Figure 56. Schematic of DRILCOM system (only azimuthal loudspeakers shown) housed within the hemi-anechoic room. (Actual photograph appears in Figure 72.) 123
Figure 57. Spectral content of DRILCOM system pink noise (green dash dot line) and dissonant signal (blue solid line) in 1/3 octave-band frequency. Seven pure tones comprising the dissonant tone are labeled above the respective frequency. Overall sound pressure level is 70 dBA for the dissonant signal and 55 dBA for the pink noise masker. 124
Figure 58. Schematic of the PALAT system (only azimuthal loudspeakers shown) located in a semi-reverberant office room at Virginia Tech. (Actual photograph appears in Figure 72.) 126
Figure 59. Spectral content of PALAT system pink noise (purple dotted line) and dissonant signal (orange dashed line) in 1/3 octave-band frequency. Seven pure tones comprising the dissonant tone are labeled above the respective frequency. Overall sound pressure level of 70 dBA for the dissonant signal and 55 dBA for the pink noise masker. 127
Figure 60. Spectral content of dissonant signal produced by the DRILCOM system (blue solid line) and the PALAT system (orange dashed line) in 1/3 octave-band frequency. Seven pure tones comprising the dissonant tone are labeled above the respective frequency. Overall sound pressure level of 70 dBA for the dissonant signal. 128
Figure 61. Spectral content of the pink noise masker produced by the DRILCOM system (green dash dot line) and the PALAT system (purple dotted line) in 1/3 octave-band frequency. Overall sound pressure level of 55 dBA for the pink noise. 129
Figure 62. 3M™ PELTOR™ TEP-100 electronic earplug-style TCAPS device...................... 132
Figure 63. Mean frequency response of 70 dBA pink noise measured using KEMAR under open ear and TEP-100 devices on normal volume setting by 1/3 octave-band frequencies. 134

Figure 64. 3M™ PELTOR™ ComTac™ III electronic earmuff-style TCAPS device. 135
Figure 65. Mean frequency response of 70 dBA pink noise measured using KEMAR under open ear and ComTac™ III devices at high volume setting by 1/3 octave-band frequency. 136
Figure 66. DRILCOM system participant training screen (left) displaying 12-signal locations (grey circles) and PALAT system participant training screen (right) displaying 24-response options (black circles) and 12-signal locations (black circles with yellow numbers). 138
Figure 67. DRILCOM system participant testing screen (left) displaying 24-response locations (grey circles) and PALAT system participant testing screen (right) displaying 24-response options (black circles) and 12-signal locations (black circles with yellow numbers). 139
Figure 68. Absolute correct response (arrow) when the signal emanates from the 1 o'clock position. 140
Figure 69. Front and back regions (shaded regions) depicting the range of signal locations and response locations for possible front-back reversal errors. 141
Figure 70. Example of semantic differential rating scale with seven anchors and bipolar descriptors at end points. 143
Figure 71. Portable audiometric booth (left) co-located with the PALAT system in the office environment used for training and pretesting. 146
Figure 72. DRILCOM system (left) in large hemi-anechoic room and PALAT system (right) in small semi-reverberant office room on the fifth floor of Whittemore Hall on the campus of Virginia Tech. 147
Figure 73. LabVIEW™ calibration screen for DRILCOM system (left) and PALAT system (right). 152
Figure 74. DRILCOM system (left) and PALAT system (right) initialization screens. 153
Figure 75. DRILCOM system (left) and PALAT system (right) training screen displaying initiation of training trial. 156
Figure 76. DRILCOM system (left) and PALAT system (right) training screen displaying feedback for absolute correct response. 156
Figure 77. DRILCOM system (left) and PALAT system (right) training screen displaying feedback for incorrect response. 157
Figure 78. DRILCOM system (left) and PALAT system (right) localization testing screen displaying initiation of testing trial. 158

Figure 79. PALAT system user-select training screen displaying loudspeaker locations used for practice trials. 160

Figure 80. Mean *absolute scores* across all participants by training system and listening condition. Percent improvement from LU0 to LU5 displayed (right). 164

Figure 81. Mean overall *front-back reversal errors* for all listening conditions at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line). 166

Figure 82. Mean *front-back reversal errors* for the open ear listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line). 166

Figure 83. Mean *front-back reversal errors* for the TEP-100 listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line). 167

Figure 84. Mean *front-back reversal errors* for the ComTac™ III listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line). 167

Figure 85. Mean overall *response time* for all listening conditions at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line). 169

Figure 86. Mean *response time* for the open ear listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line). 169

Figure 87. Mean *response time* for the TEP-100 listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line). 170

Figure 88. Mean *response time* for the ComTac™ III listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line). 170

Figure 89. **Mean *absolute score* for each listening condition at LU0 (pretest)** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test. 175

Figure 90. **Mean *absolute score* for each listening condition by training system at LU0 (pretest)** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means. 176

Figure 91. **Mean *absolute score* for each listening condition at LU0 (pretest) while using the DRILCOM system** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a paired-samples *t*-test. 178

Figure 92. **Mean *absolute score* for each listening condition at LU0 (pretest) while using the PALAT system** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a paired-samples *t*-test...... 180

Figure 93. **Mean *absolute score* for each listening condition at LU5 (posttest)** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a paired-samples *t*-test. 182

Figure 94. **Mean *absolute score* for each listening condition by training system at LU5 (posttest)** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means. 183

Figure 95. **Mean *absolute score* for each listening condition at LU5 (posttest) while using the DRILCOM system** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a paired-samples *t*-test. 185

Figure 96. **Mean *absolute score* for each listening condition at LU5 (posttest) while using the PALAT system** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a paired-samples *t*-test...... 187

Figure 97. **Mean *absolute score* for each listening condition across both training systems and stages of training (LU0 and LU5)** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a paired-samples *t*-test. 190

Figure 98. **Mean *absolute score* by training systems and listening condition across stages of training (LU0 and LU5)** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means. 191

Figure 99. **Mean *absolute score* at LU0 (pretest) and LU5 (posttest)** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per repeated-measures ANOVA. 192

Figure 100. **Mean *absolute score* by training systems at LU0 (pretest) and LU5 (posttest)** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means. 193

Figure 101. **Mean *absolute score* across all listening conditions at LU0 and LU5 by training system**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs. 195

Figure 102. **Mean *absolute score* open ear condition at LU0 and LU5 by training system**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs. 198

Figure 103. **Mean *absolute score* TEP-100 condition at LU0 and LU5 by training system**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs. 201

Figure 104. **Mean *absolute score* ComTac^{TM} III condition at LU0 and LU5 by training system**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs. 203

Figure 105. **Mean slope of absolute score from LU0 through LU5 by listening condition on the DRILCOM and PALAT systems**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs. 207

Figure 106. Radial plots of mean absolute correct accuracy percentage by loudspeaker source location for each listening condition comparing training system, DRILCOM (solid blue line) and PALAT (dashed orange line) by stage of training, LU0, pretest, (left column) and LU5, posttest, (right column). 12 represents the position directly in front of the participant, at 12 o'clock or 0 degrees azimuth. 211

Figure 107. Radial plots of mean absolute correct accuracy percentage by loudspeaker source location for each listening condition comparing LU0, pretest, (dashed red line) and LU5, posttest, (solid purple line) by training system, DRILCOM (left column) and PALAT (right column). 12 represents the position directly in front of the participant, at 12 o'clock or 0 degrees azimuth. 212

Figure 108. **Mean *front-back reversal errors* for each listening condition across both training systems at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test. 216

Figure 109. **Mean *front-back reversal errors* by training systems and listening condition at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means. 217

Figure 110. **Mean *front-back reversal errors* at LU0 (pretest) and LU5 (posttest) across all training systems and listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a repeated-measures ANOVA. 218

Figure 111. **Mean *front-back reversal errors* by training systems at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means. 219

Figure 112. **Mean *response time* for the DRILCOM and PALAT systems across all listening conditions at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a repeated-measures ANOVA. 222

Figure 113. **Mean *response time* for each listening condition across both training systems at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test...... 223

Figure 114. **Mean *response time* for each listening condition on the DRILCOM system at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test...... 225

Figure 115. **Mean *response time* for each listening condition on the PALAT system at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test...... 228

Figure 116. Radial plots of mean response time by loudspeaker source location for each listening condition comparing training system, DRILCOM (solid blue line) and PALAT (dashed orange line) by stage of training, LU0, pretest, (left column) and LU5, posttest, (right

column). 12 represents the position directly in front of the participant, at 12 o'clock or 0 degrees azimuth. 230

Figure 117. Radial plots of mean response time by loudspeaker source location for each listening condition comparing LU0, pretest, (dashed red line) and LU5, posttest, (solid purple line) by training system, DRILCOM (left column) and PALAT (right column). 12 represents the position directly in front of the participant, at 12 o'clock or 0 degrees azimuth. 231

Figure 118. **Mean *subjective ratings* for Question 1, perceived confidence on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test. 236

Figure 119. **Mean *subjective ratings* for Question 1, perceived confidence on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ within each listening condition per a Wilcoxon signed-rank test. 237

Figure 120. **Mean *subjective ratings* for Question 2, loudspeaker proximity impacts using the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test. 240

Figure 121. **Mean *subjective ratings* for Question 2, loudspeaker proximity impacts on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ within each listening condition per a Wilcoxon signed-rank test. 241

Figure 122. **Mean *subjective ratings* for Question 3, ease of use on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test. 243

Figure 123. **Mean *subjective ratings* for Question 3, ease of use on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on

each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test. 244

Figure 124. **Mean *subjective ratings* for Question 4, room effects on localization on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test... 246

Figure 125. **Mean *subjective ratings* for Question 4, room effects on localization on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test... 247

Figure 126. **Mean *subjective ratings* for Question 5, training impacts on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test. 250

Figure 127. **Mean *subjective ratings* for Question 5, training impacts on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test. 251

Figure 128. **Mean *subjective ratings* for Question 6, difficulty judging signal location on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test... 253

Figure 129. **Mean *subjective ratings* for Question 6, difficulty judging signal location on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test... 254

Figure 130. **Mean *subjective ratings* for Question 6, difficulty judging signal location between each listening condition**, with 95% confidence intervals plotted around the mean

values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test.. 255

Figure 131. **Mean *subjective ratings* for Question 6, difficulty judging signal location between each listening condition**, with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test.. 257

Figure 132. **Mean *subjective ratings* for Question 7, impact on reaction time on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test... 258

Figure 133. **Mean *subjective ratings* for Question 7, impact on reaction time on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test.. 259

Figure 134. **Mean *subjective ratings* for Question 7, impact on reaction time between each listening condition**, with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test. 260

Figure 135. **Mean *subjective ratings* for Question 8, impact of user interface on localization on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test.................................... 262

Figure 136. **Mean *subjective ratings* for Question 8, impact of user interface on localization on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test. 264

Figure 137. **Mean *subjective ratings* for Question 9, room environment impacts on reaction time on the DRILCOM and PALAT systems across all listening conditions** with 95%

confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test.................................... 266

Figure 138. **Mean *subjective ratings* for Question 9, room environment impact on reaction time on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test. 267

Figure 139. **Mean *subjective ratings* for Question 10, impact of loudspeaker visibility on localization on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test. 269

Figure 140. **Mean *subjective ratings* for Question 10, impact of loudspeaker visibility on localization on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test. 270

Figure 141. Mean *absolute score* under open ear condition by training system by order of assignment, DRILCOM system (blue solid line) and PALAT system (orange dashed line). 278

Figure 142. Mean *absolute score* under TEP-100 condition by training system by order of assignment, DRILCOM system (blue solid line) and PALAT system (orange dashed line). 279

Figure 143. Mean *absolute score* under ComTac™ III condition by training system by order of assignment, DRILCOM system (blue solid line) and PALAT system (orange dashed line). 281

Figure 144. Phase II mean *absolute score* for six participants assigned to DRILCOM (blue solid line) or PALAT (orange dotted line) for the first training session versus mean absolute score from 11 participants tested using Etymotic ER125-GSE (black dash-dotted line) on DRILCOM in a previous study (Casali & Lee, 2020). Percent improvement from LU0 to LU5 of correct responses displayed on right side of chart.. 286

Figure 145. Frequency response radial plots of mean sound pressure level at a single 1/3 octave-band frequency. Measurements recorded using KEMAR manikin with open ear (blue dotted lines) and TEP-100 (orange solid lines) in right ear. Dissonant tone set at 55 dBA (left column) and 80 dBA (right column). 289

Figure 146. Level dependent function of the TEP-100 measured using a broadband pink noise signal; data are representative, but not considered exact for all samples, per manufacturer (Stergar et al., 2019). 290

Figure 147. Phase III experimental design order. 298

Figure 148. Experimental design for Phase III, with independent variables, experimental order, participant assignment, and dependent measures listed. 299

Figure 149. 3M™ PELTOR™ TEP-100 electronic earplug-style TCAPS device. 302

Figure 150. Mean frequency response of 70 dBA pink noise measured using KEMAR under open ear and TEP-100 devices on normal volume setting by 1/3 octave-band frequencies. 305

Figure 151. 3M™ PELTOR™ ComTac™ III electronic earmuff-style TCAPS device. 306

Figure 152. Mean frequency response of 70 dBA pink noise measured using KEMAR under open ear and ComTac™ III devices at high volume setting by 1/3 octave-band frequency. 307

Figure 153. 1/3 octave-band spectral content of pretest dissonant training signal (played on PALAT system) and posttest .22 caliber blank gunshot by 1/3 octave-band frequency. Recorded at the participant's ear in office environment (for PALAT) or at outdoor field site (for transfer-of-training test). Overall sound pressure level of 70 dBA for both signals. 309

Figure 154. Participant pretest and posttest screen displaying 24 response options (black circles) and 12-signal locations (black circles with yellow numbers). 310

Figure 155. Absolute correct response (arrow) and ballpark correct response region (grey shaded region) when the signal emanates from the one o'clock position. 311

Figure 156. Front and back regions (shaded regions) depicting the range of signal locations and response locations for possible front-back reversal errors. 313

Figure 157. Example of semantic differential rating scale. 314

Figure 158. Portable audiometric booth (left) co-located with the PALAT system in the office environment used for training and pretesting. 317

Figure 159. PALAT system apparatus located in a semi-reverberant office room at Virginia Tech. 321
Figure 160. 1/3-octave band spectral content of PALAT system pink noise (green dashed line) and dissonant signal (blue solid line) in 1/3 octave-band frequency. Eight pure tones comprising the dissonant tone are labeled above the respective frequency. 321
Figure 161. Example CSV file output of PALAT system. 323
Figure 162. Aerial view of field site layout with 12-remote firing device positions, located around the participant. 324
Figure 163. Mean 1/3-octave band spectral content of .22 caliber blank gunshot from all 12 firing device positions (dashed green line) versus ambient noise floor (dotted purple line) measured in 1/3 octave-band frequency at the participant center head position. 325
Figure 164. Field site panoramic picture (from left at 7 o’clock to right at 6 o’clock), with clock face positions identified. Only the 12 o’clock position actually included a sign during the conduct of the field experiment. 326
Figure 165. Participant view of computer tablet used to operate the field posttest and 12 o’clock reference sign to orient participant. 326
Figure 166. Remote firing device design concept sketch and final product. 327
Figure 167. One remote firing device containing three remote firing mechanisms mounted on a steel u-post located in the wooded forest at the field localization site (Left: Front view of remote firing device, Right: Profile (side) view as seen from direction of the participant), which reduced the visual signature. 328
Figure 168. Remote firing device control box. 329
Figure 169. Remote firing device block diagram with all major components. 330
Figure 170. Participant-controlled computer tablet placed on music stand at the center of the field experiment site. 331
Figure 171. PALAT system initialization screen. 333
Figure 172. Participant operating the PALAT system. 334
Figure 173. PALAT system main menu screen on computer tablet displaying training and testing protocol options. 335
Figure 174. PALAT system sequential training screen on computer tablet displaying initiation training trial. 336

Figure 175. PALAT system test screen on computer tablet displaying initiation and response options. 337
Figure 176. Participant interface on the computer tablet displaying 24 response location options and system-generated feedback for an absolute correct response. 340
Figure 177. Participant interface on the computer tablet displaying feedback provided for an incorrect response. 340
Figure 178. Example of the user-select display interface on the computer tablet displaying loudspeaker locations used for practice. 341
Figure 179. Field posttest site layout with participant standing in center of remote firing devices facing 12 o'clock target position and two investigators seated behind participant between the 6 o'clock and 7 o'clock positions. 342
Figure 180. Absolute correct score on pretest under TEP-100 listening condition for all participants. Values ordered and plotted from lowest to highest score by group. 345
Figure 181. Response time on posttest under open ear listening condition for all participants. Values ordered and plotted from lowest to highest response time by group. 345
Figure 182. Mean ballpark correct scores with 95% confidence intervals plotted. Different letters indicate significant differences at $p \leq 0.10$. 352
Figure 183. Mean absolute correct scores for each listening condition with 95% confidence intervals plotted. Different letters indicate significant differences at $p \leq 0.10$. 353
Figure 184. Mean ballpark correct scores for each listening condition with 95% confidence intervals plotted. Different letters indicate significant differences at $p \leq 0.10$. 354
Figure 185. Mean response times for each listening condition with 95% confidence intervals plotted. Different letters indicate significant differences at $p \leq 0.10$. 355
Figure 186. Mean response times from pretest to posttest. Different letters indicate significant differences at $p \leq 0.10$. 356
Figure 187. Mean absolute correct scores for each group at pretest and posttest. Different letters indicate significant differences at $p \leq 0.10$. 357
Figure 188. Mean absolute correct scores for each training group comparing pretest and posttest performance. Different letters indicate significant differences at $p \leq 0.10$. 358
Figure 189. Mean absolute correct scores for the trained versus untrained groups at pretest for each listening condition. Different letters indicate significant differences at $p \leq 0.10$. 360

Figure 190. Mean absolute correct scores at posttest for each listening condition comparing the trained and untrained groups. Different letters indicate significant differences at $p \leq 0.10$.360
Figure 191. Mean absolute correct scores for each listening condition for the **untrained** group **comparing pretest and posttest** performance. Different letters indicate significant differences at $p \leq 0.10$. 362
Figure 192. Mean absolute correct scores for each listening condition for the **trained** group **comparing pretest and posttest** performance. Different letters indicate significant differences at $p \leq 0.10$. 362
Figure 193. Mean front-back reversal errors using the 120-degree arc criterion for each listening condition. Different letters indicate significant differences at $p \leq 0.10$. 365
Figure 194. Mean absolute correct score on pretest and posttest for trained group. Regression line and equation plotted. 367
Figure 195. Slopes from pretest to posttest for mean absolute correct score for open ear by group. 368
Figure 196. Slopes from pretest to posttest for mean absolute correct score for open ear by group. 369
Figure 197. Plotted means and 95% confidence intervals for ratings at pretest and posttest for each group and listening condition for Question 1, Perceived Confidence. 371
Figure 198. Mean ratings for each group at pretest for Question 1, Perceived Confidence. 375
Figure 199. Mean ratings for each group at posttest for Question 1, Perceived Confidence. 376
Figure 200. Plotted means and 95% confidence intervals for ratings at pretest and posttest for each group and listening condition for Question 2, Perceived Accuracy. 377
Figure 201. Mann-Whitney *U* results comparing trained and untrained ratings at pretest in the open ear condition for Question 2, Perceived Accuracy 379
Figure 202. Mean ratings for each group at pretest for Question 2, Perceived Accuracy. 380
Figure 203. Mean ratings for each group at posttest for Question 2, Perceived Accuracy. 381
Figure 204. Plotted means and 95% confidence intervals for ratings at pretest and posttest for each group and listening condition for Question 3, Perceived Difficulty. 383
Figure 205. Mean ratings for each group at pretest for Question 3, Perceived Difficulty. 385
Figure 206. Mean ratings for each group at posttest for Question 3, Perceived Difficulty. 386

Figure 207. Plotted means and 95% confidence intervals for ratings at pretest and posttest for each group and listening condition for Question 4, Perceived Reaction Time. 388

Figure 208. Mean ratings for each group and listening condition at pretest for Question 4, Perceived Reaction Time. 391

Figure 209. Mean ratings of perceived reaction time for each group and listening condition at posttest for Question 4, Perceived Reaction Time. 392

Figure 210. Plotted means and 95% confidence intervals at pretest and posttest for each group and listening condition for Question 5, Perceived Comfort. 394

Figure 211. Mean ratings of perceived comfort for each group and listening condition at pretest for Question 5, Perceived Comfort. 396

Figure 212. Mean ratings of perceived comfort for each group and listening condition at pretest for Question 5, Perceived Comfort. 397

Figure 213. Plotted means and 95% confidence intervals for likelihood to maintain the same listening condition ratings at pretest and posttest for each group and listening condition for Question 6, Likelihood of Wearing Device. 399

Figure 214. Mean ratings of likelihood of wearing device given similar conditions for each group and listening condition at pretest for Question 6, Likelihood of Wearing Device. 401

Figure 215. Mean ratings of likelihood of wearing device given similar conditions for each group and listening condition at posttest for Question 6, Likelihood of Wearing Device. 402

Figure 216. Plotted means and 95% confidence intervals for perceived degree of preparedness for each listening condition ratings at posttest for the trained group for Question 7, Degree of Preparedness. 404

Figure 217. Radial plots of mean absolute correct accuracy percentage for each listening condition during pretest and posttest by group, trained (solid line) and untrained (dashed line). 407

Figure 218. Level dependent function of the TEP-100 measured using a broadband pink noise signal; data are representative, but not considered exact for all samples, per manufacturer (Stergar et al., 2019). 415

Figure 219. Frequency response radial plots of mean sound pressure level at a single 1/3 octave-band frequency. Measurements recorded using KEMAR manikin with open ear (dashed

blue lines) and TEP-100 (solid orange line) in right ear. Dissonant tone set at 55 dBA (left column) and 80 dBA (right column). 417

Figure 220. Input-output curve across frequencies using a pure-tone sweep at 60 and 90 dB SPL for the TEP-100 in the unity gain setting. Data were obtained with an Audioscan Verifit system. 429

Figure 221. Mann-Whitney U results comparing mean ratings of confidence for each listening condition at pretest for trained versus untrained groups for Question 1, Perceived Confidence. 532

Figure 222. Mann-Whitney U results comparing mean ratings of confidence for each listening condition at posttest for trained versus untrained groups for Question 1, Perceived Confidence. 532

Figure 223. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at pretest in the TEP-100 and ComTac™ III conditions for Question 2, Perceived Accuracy. 533

Figure 224. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at pretest for all listening condition for Question 3, Perceived Difficulty. 533

Figure 225. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at posttest for all listening condition for Question 3, Perceived Difficulty. 534

Figure 226. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at pretest and posttest collapsed across listening conditions for Question 4, Perceived Reaction Time. 534

Figure 227. Mann-Whitney *U* results comparing trained and untrained ratings of perceived reaction time at pretest for all listening condition for Question 4, Perceived Reaction Time. 535

Figure 228. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at pretest for all listening condition for Question 4, Perceived Reaction Time... 535

Figure 229. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at posttest for all listening condition for Question 4, Perceived Reaction Time. 536

Figure 230. Mann-Whitney U results comparing mean ratings of confidence for each listening condition at pretest for trained versus untrained groups for Question 5, Perceived Comfort. 536

Figure 231. Mann-Whitney U results comparing mean ratings of confidence for each listening condition at posttest for trained versus untrained groups for Question 5, Perceived Comfort. 536

Figure 232. Mann-Whitney U results comparing mean ratings of likelihood of wearing device for each listening condition at pretest for trained versus untrained groups for Question 6, Likelihood of Wearing Device... 537

Figure 233. Mann-Whitney U results comparing mean ratings of likelihood of wearing device for each listening condition at posttest for trained versus untrained groups for Question 6, Likelihood of Wearing Device... 537

CHAPTER 1. Introduction

The ability to rely on the hearing sense to detect, recognize, and locate potential hazards and threats is vital for survivability and mission success in military operations. In many cases, these hazards and threats are not immediately visible due to camouflage, obstacles, and environmental factors. As a result, service members must rely on auditory cues to provide situation awareness. Auditory situation awareness is the ability to detect, recognize, and localize, both horizontally and vertically, sound sources, as well as understand speech, or communication that may contain information about one's environment and dynamic situation (Hajicek, Myrent, Li, Barker, & Coyne, 2010; Lee & Casali, 2017). In a recent study involving 80 British army infantry personnel, sound localization was identified as being of high importance for unit safety and mission efficiency yet the warfighters were unsure how accurately they were able to determine the origin of small arms fire and expressed that they received no substantial training on localization (Bevis, Semeraro, van Besouw, Rowan, Lineton, & Allsopp, 2014). The military recognizes the need for sound localization as a combat multiplier but acknowledges the need to protect service members from noise hazards (Department of the Army, 2015; Donahue & Ohlin, 1993).

Hazardous noise is one of the primary occupational dangers in the military. Nearly two decades of combat operations and amplified training requirements have increased service members' exposure to noise from weapons, vehicles, aircraft, and explosions (McIlwain, Gates, & Ciliax, 2008; Jokel, Yankaskas, & Robinette, 2019). As a result, tinnitus and hearing loss were the two most prevalent service-connected disabilities of new compensation recipients in 2016 and effects over 2.8 million veterans (United States Department of Veterans Affairs, 2016). The prevalence rates of significant hearing loss among just the U.S. Army service members was 24%

in 2016 (DOEHRS-DR, 2016). The Veterans Benefit Administration does not publish annual compensation spending by disability. However, a study conducted in 2004 estimated the annualized cost for compensation payments to veterans with noise-induced hearing loss (NIHL) or tinnitus as their major form of disability to be over $850 million (Humes, Joellenbeck, & Durch, 2005). Since Humes et al. (2005) report, the number of veterans with NIHL or tinnitus service-connected disabilities has increased by 330%, potentially costing over $2 million annually using the same compensation rates per service member (United States Department of Veterans Affairs, 2006, United States Department of Veterans Affairs, 2016).

To combat noise-induced hearing loss, the Department of Defense (DoD) established the Hearing Center of Excellence (HCE) and implemented Auditory Fitness for Duty (AFFD) standards. One of the missions of the DoD Hearing Center of Excellence is working with the military service branches to develop military AFFD standards to determine what level of hearing is necessary in order to perform military tasks. Currently, the three service branches utilize a hearing profile criterion that categorizes hearing loss into four classifications; H1 = mild hearing loss, H2 = moderate hearing loss, H3 = significant hearing loss, H4 = severe hearing loss (Department of the Army, 2015; Department of the Air Force, 2013; Navy and Marine Corps Public Health Center, 2008). In 2006, the Army Hearing Program was established requiring audiograms be administered before entering service, upon separation, before and after each deployment, and annually during service (Brungart, 2014). The Marines adopted the annual audiogram requirement in 2012 (Brungart, 2014). However, no service branch has incorporated sound localization testing as part of their hearing programs due to challenges of both the expense of setting up systems and issues with room acoustics (Brungart, 2014). The increased auditory monitoring will help identify service members with threshold shifts but does not address the root cause of hearing loss, nor does it determine a warfighter's ability to localize sounds, which is critical to situation awareness.

One of the reasons for the pervasive hearing loss epidemic in the military can be attributed to service members' unwillingness to wear hearing protection devices (HPDs) due to the deleterious effect that HPDs can impose on mission performance (Abel, 2008; Bevis et al., 2014; Casali, Ahroon, & Lancaster, 2009; Giguere, Laroche, & Vailancourt, 2013; Price, Kalb, & Garinther, 1989; Vause & Grantham, 1999). The undamaged human ear is excellent at detecting, identifying, and localizing sound cues. Blauert (1997) reported that the unoccluded human ear can localize sounds in the horizontal plane within 2° azimuth when the sound is in front and within 10° azimuth when the sound is to the left or right. Unfortunately, this fine-tuned ability is decremented when the ear is covered or plugged by a hearing protector. In a focus group involving Canadian military personnel, Abel (2008) reported soldiers did not wear hearing protection during training and combat because the devices were uncomfortable or hard to fit, they believed they would not be able to hear critical communication, or that the HPDs would interfere with situational awareness. Numerous studies have shown that conventional passive hearing protection can degrade auditory tasks of detection, identification, and localization, depending upon the device, the wearer's hearing ability, and the situation (e.g. see Casali, 2010b & Casali, 2012a). As a result, the design feature of hearing protectors has been augmented in the form of active (electronic powered) HPDs that are intended to provide the same level of protection but preserve auditory detection and localization abilities as well as enable radio communication. In 2012, the military began fielding active HPDs referred to as Tactical Communication and Protective Systems (TCAPS), advertised to merge good hearing protection in concert with communication devices while enhancing the ability to locate and identify auditory warnings (Palca, 2016; PEO Soldier, 2017).

A series of in-field and laboratory studies conducted at the Virginia Tech Auditory Systems Laboratory (VT-ASL) confirmed that sound detection and localization is negatively affected by both conventional passive HPDs and active HPDs. The degree of detection degradation and errors in distance judgments was shown to vary greatly between devices in backup alarms (Alali & Casali, 2012) and military signals (Clasing & Casali, 2014; Lee & Casali, 2016; Lee & Casali, 2017). Casali, Ahroon, and Lancaster (2009) simulated combat scenarios, reconnaissance and raid missions, in an in-field experiment using passive and active HPDs worn by Army ROTC Cadets and found that during the raid mission, soldiers wearing one of the active HPDs, the CEPS device which provided about 36 dB of input-to-output gain, were able to detect noise of an enemy camp at 400 feet, doubling the 220 feet detection distance needed with the open ear (Figure 1).

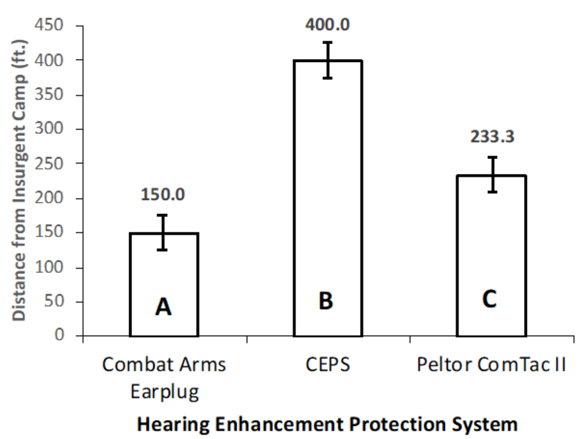


Figure 1. Mean enemy camp detection distances across HPDs, with 95% confidence interval plotted. Mean values, shown on bars, with the same letter are not significantly different according to Tukey’s test at $p<0.05$ (adapted from Casali et al., 2009, Figure 4).

However, during the reconnaissance mission, the researchers observed that all three devices limited the ability to detect certain discrete auditory threats that were clearly audible at 50 feet to the normal ear. In addition, the researchers and participants noted that all three devices tested had ergonomic issues which have proven to be an impediment to use among service members (Abel, 2008; Bevis et al., 2014; Casali et al., 2009). Similar degradation effects were discovered on the ability to localize sound sources while wearing HPDs and TCAPS.

In a vehicular backup alarm localization study using a high-fidelity laboratory simulation, wearing active HPDs and TCAPS did not improve the normal hearing listeners' ability to locate warning signals in 360° azimuth in 60 dBA and 90 dBA pink noise as compared to when wearing conventional passive earmuffs or earplugs (Alali & Casali, 2011). Alali and Casali (2011) also found that localization was consistently more degraded in the 90 dBA pink noise compared to the 60 dBA pink noise with 27% fewer correct localization responses in the higher noise condition (Figure 2).

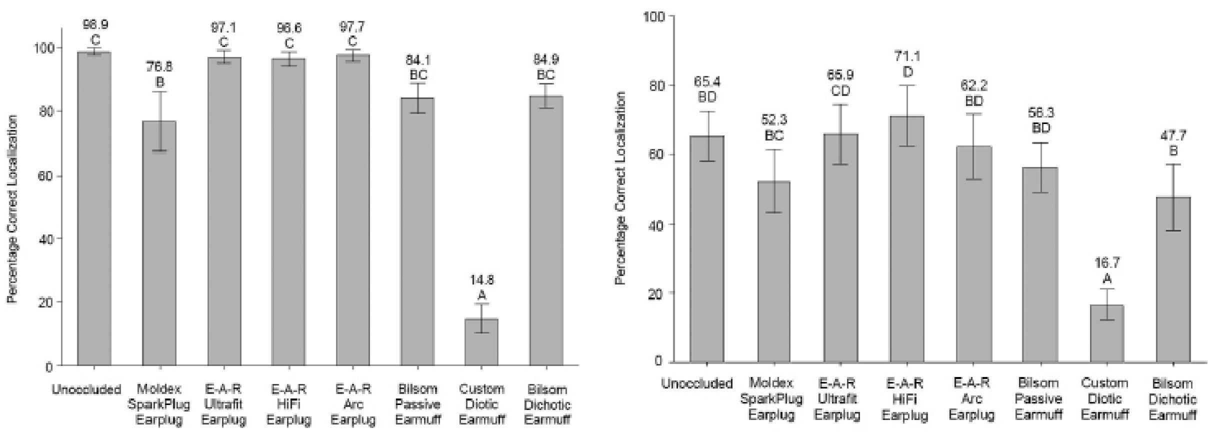


Figure 2. Effects of HPDs at 60 dBA (left graph) and 90 dBA (right graph) background noise level on percentage correct localization; mean values shown on bars with 95% confidence intervals. Means with the same letter are not significantly different at $p \leq 0.05$. (adapted from Alali & Casali, 2011, Figures 9 & 10).

Similar results were found in a field experiment to localize the azimuthal direction of actual gunshots where the performance of active HPDs and TCAPS was worse than that with the open ear (Talcott, Casali, Keady, & Killion, 2012). Wearing an active HPD or TCAPS reduced the mean percent correct response (within ±22.5°) by 18%, 20%, and 28% compared to the open ear in a quite setting (45-50 dBA ambient noise) and 29%, 24%, and 45% in the presence of high background noise (82 dBA background noise) (Talcott et al., 2012). In addition, participants' mean response time increased significantly while wearing three of the four passive or active HPDs and TCAPS compared to the open ear response time (Figure 3) (Talcott et al., 2012).

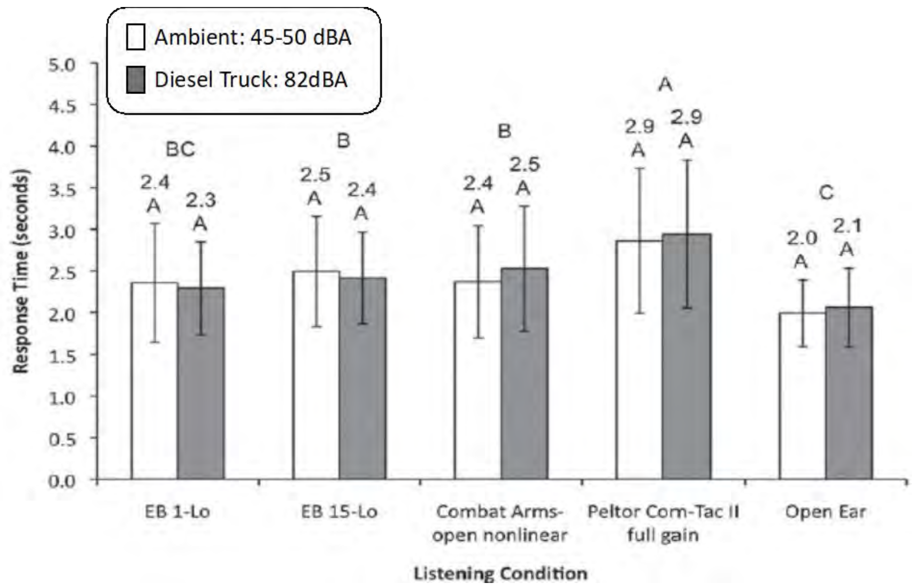


Figure 3. The effect of listening condition on mean response time. Error bars are the 95% confidence interval about the mean. Mean number shown above bars. Top letters show Tukey's multiple comparisons test for main effects of listening conditions (adapted from Talcott et al., 2012, Figure 5).

Based on the findings in the previous studies and the lack of incorporating auditory situation awareness test into AFFD, the VT-ASL developed a test battery for this purpose, funded by the Department of Defense Hearing Center of Excellence. This test battery had the primary objective of *testing* listeners, with and without TCAPS and HPDs, on the subtasks of auditory situation awareness including Detection, Recognition/Identification, Localization, and COMmunications (DRILCOM), and a secondary objective of *training* listeners to improve localization performance with the open ear and various TCAPS and HPDs. The block diagram of this DRILCOM concept is shown in Figure 4 (Casali & Lee, 2016a).

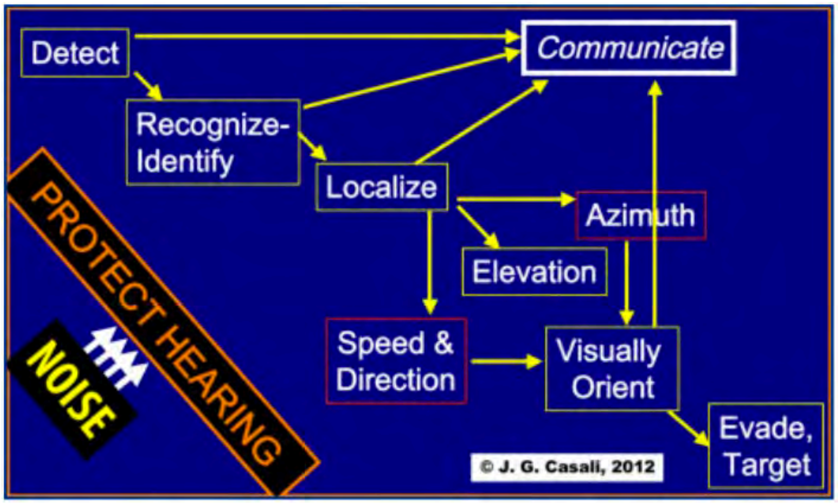


Figure 4. Network diagram of human hearing subtasks involved in achieving and maintaining auditory situation awareness (adapted from Casali & Lee, 2016a, Figure 1) Copyright 2012 by John G. Casali.

The localization portion of the test battery was designed to measure how well a user can localize a sound signal in 360° azimuthal direction and in frontal elevation. The apparatus consisted of a 3-meter diameter circular array of 12 directional loudspeakers separated by 30° increments horizontally and three additional speakers located 30° above the 330°, 0°, and 30° horizontal speakers to test elevation localization. The speaker array was housed in a large, hemi-anechoic room equipped with an investigator control station located outside of the speaker ring and a small control station with a computer monitor and mouse located in the middle of the ring for participant responses. The proof-of-concept experimental test results measured "absolute" (participant's response exactly matches signal speaker location) and "ballpark" (participant's response within ±15°) localization accuracy and confirmed detection, recognition, and localization degradation differences amongst various HPDs and TCAPS and between the devices and the open ear (Casali & Lee, 2016a). Using a DRILCOM-like experimental setup, Casali and Robinette (2014) demonstrated the plasticity of the human auditory system to learn and improve azimuthal localization abilities with the open ear and while wearing TCAPS (both in-the-ear and over-the-ear) using an acclimation-training regimen. After 12, one-hour training sessions spread out over two weeks, participants improved their localization performance for all listening conditions and reduced the mean absolute error using hearing protection to similar levels achieved after training with the open ear (Figure 5) (Casali & Robinette, 2014).

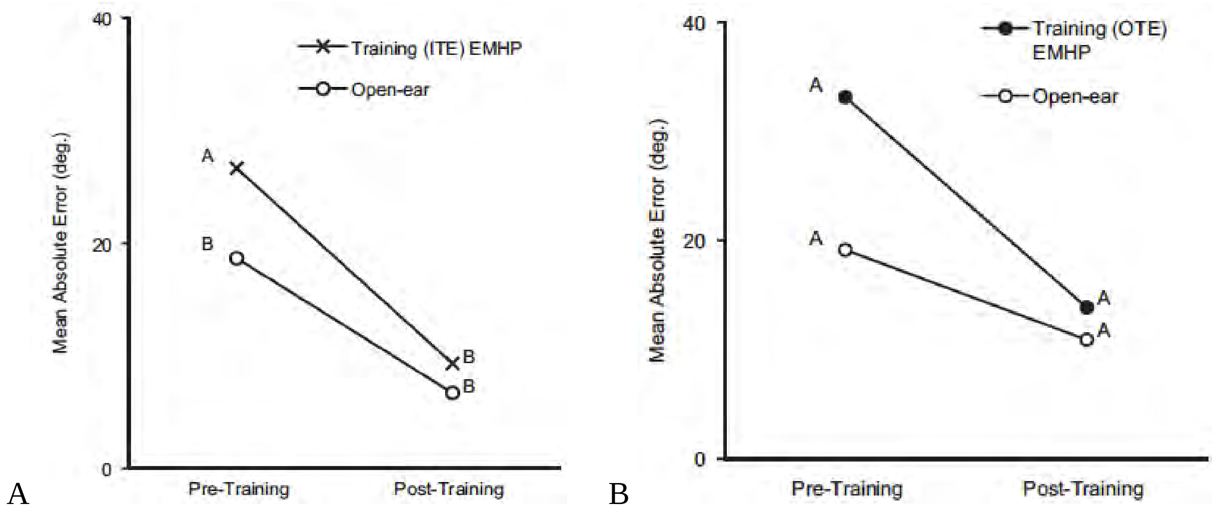


Figure 5. Mean absolute localization error for open ear and training electronically modified hearing protection device, in-the-ear (A) and over-the-ear (B). Within pairs of means, both across Training Day (horizontal) and across Listening Condition (vertical), means with same letter are not significantly different at p 0.05 using a paired sample *t*-test with Bonferroni correction (adapted from Casali & Robinette, 2014, Figures 3 & 4).

As a precursor to this proposal, Casali and Lee (2016b) developed a pilot acclimation-test system based on the azimuthal localization element of DRILCOM designed to train horizontal localization. After 12 learning units of acclimation and testing, participants improved the open ear's absolute correct performance by over 25% and demonstrated that participants using certain TCAPS can learn and perform at similar ballpark levels to the open ear with relatively little training (Figure 6) (Casali & Lee, 2016b). The study also confirmed the device-specific localization performance variations as seen between the results of TCAPS A and B, with TCAPS B performing much worse and never approaching open ear performance on either ballpark correct or absolute correct measures.

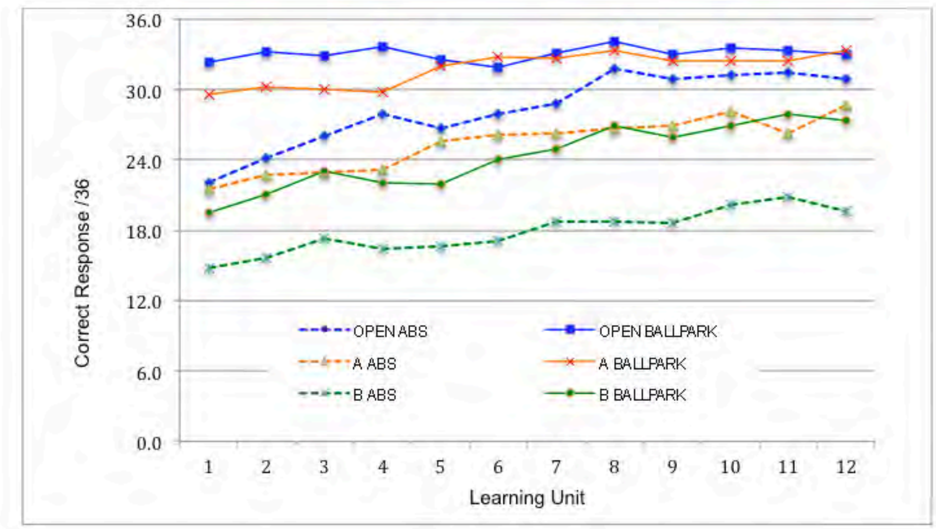


Figure 6. Number of correct responses (absolute (ABS) correct and ballpark correct) out of 36 total possible, under each testing condition by learning units (OPEN = open ear, A and B represent two TCAPS devices) (adapted from Casali & Lee, 2016b).

1.1 Purpose

The Office of Naval Research (ONR) provided a grant to develop a training protocol for a sound localization acclimation-training system, develop and evaluate a Portable Auditory Localization Acclimation Training (PALAT) system, and validate the system against a proven in-field localization test (per Talcott et al., 2012). The ability to improve sound localization and auditory situation awareness via a brief training regimen could have a profound impact on service members' survivability and mission success. Unfortunately, it is not feasible to field every military installation with the current full-scale apparatus used to test and train localization, which requires a large hemi-anechoic room and a 3-meter diameter ring of 12 loudspeakers around the participant's head. A small-scale, portable, and validated auditory localization training system is needed to fill a critical vulnerability not only in the military, but also in any industry where workers wear hearing protection and auditory situation awareness is required. Ultimately, increased confidence in auditory situation awareness while wearing TCAPS may encourage TCAPS adoption rates and thus reduce exposure to hazardous noise.

The primary purpose of this research was to develop and evaluate a Portable Auditory Localization Acclimation Training (PALAT) system that is capable of improving service members' localization ability with the open ear and while wearing a TCAPS device. The broader research effort occurred in three distinct phases: Phase I Investigation of Training Protocol, Phase II Development and In-Laboratory Investigation of the PALAT system, and Phase III In-field Investigation of Transfer-of Training. The following sections discuss the purpose of each phase of this investigation.

Phase I Investigation of Training Protocol

The purpose of Phase I was to develop an improved azimuth training protocol for use in the PALAT system. This phase of the study was conducted as a separate dissertation research effort by K. Cave (Cave, 2019). The DRILCOM system was used to develop a training regimen and to test the transfer-of-training from a dissonant tonal complex stimulus to four military relevant stimuli. The Phase I study was conducted using only the open ear listening condition. Phase I resulted in an improved training protocol consisting of five learning units using the dissonant signal that was incorporated into the PALAT system for Phase II and Phase III.

Phase II Development and In-Laboratory Investigation of the PALAT system

The purpose of Phase II was to develop and evaluate a portable localization training and testing system capable of reproducing the training effects of the full-size, laboratory grade DRILCOM system. The objective was to design the PALAT system with input from Subject Matter Experts to train and test the full battery of auditory situation awareness tasks (Detection, Recognition, Identification, Localization, and Communication). The Phase II experiment focused on validating the localization training effect of the PALAT system against the proven in-laboratory localization training effect of the DRILCOM system (per Casali & Lee, 2016a, Casali

& Lee, 2016b). The full-factorial main experiment compared the portable system against the laboratory system with the open ear, an in-the-ear TCAPS, and an over-the-ear TCAPS.

Phase III In-field Investigation of Transfer-of Training

The purpose of the Phase III in-field experiment was to evaluate the transfer-of-training effects of conducting localization training in-laboratory, using the PALAT system, on in-field localization performance. In Phase III, localization performance was compared between trained and untrained participants using an in-laboratory pretest using a dissonant signal on the PALAT system with a proven in-field posttest using blank gunshots (per Talcott et al., 2012). The Phase III experiment compared localization performance on the in-laboratory pretest and in-field posttest with the open ear, an in-the-ear TCAPS, and an over-the-ear TCAPS.

1.2 Background

1.2.1 Human Auditory System

The human auditory system is composed of the sensory organs (two ears located on the left and right side of the head) and the auditory sensory system (central auditory nervous system). Ears enable humans to hear by transmitting sound energy to the brain where it is perceived and interpreted. Human ears are comprised of three functional regions consisting of the outer, middle, and inner ear (Figure 7). The outer ear includes the pinna and external auditory canal and performs the tasks of funneling sound waves into the ear. Sound waves are modified by the pinna and auditory canal, amplifying or attenuating sound frequencies as they collect and direct the waves to the tympanic membrane, or eardrum. The collection and modification of the sound waves by the pinna produces distinctive imprints that are used for sound recognition and localization (Ward, Royster, & Royster, 2003). In addition, the shape and dimensions of the head, pinna, and auditory canal amplify frequencies in the 2000 to 4000 Hz region by 10 to 15 dB (Ward et al., 2003). Differences in sound levels measured outside the ear and just before the external auditory canal are known as the transfer function of the open ear (TFOE) (Casali, Mauney, & Burks, 1995).

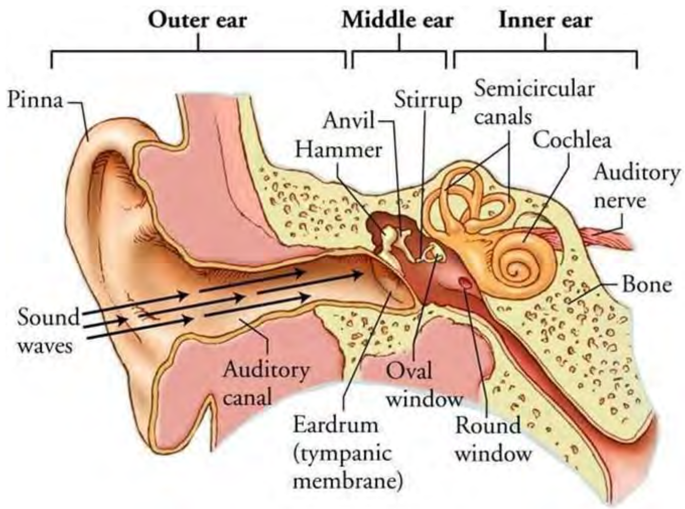


Figure 7. Cross-sectional view of the human ear (adapted from Samsung, 2016).

The middle ear is an air-filled cavity surrounded by the temporal bone and separated from the outer ear by the tympanic membrane. An Eustachian tube equalizes the air pressure within the middle ear to that of the nasopharynx at the roof of the mouth. There are three small bones, the malleus (hammer), incus (anvil), and stapes (stirrup), located in the middle ear that collectively make up the ossicles. Ossicles form a chain that connects the tympanic membrane to the oval window of the cochlea in the inner ear. The middle ear is responsible for converting, via structural-borne conduction, the airborne pressure vibration into fluid motion in the inner ear (Maroonroge, Emanuel, & Letowski, 2009). The ossicular chain lever action overcomes the impedance mismatch between the air in the outer ear and cochlear fluid in the inner ear by amplifying the sound signal up to 30 dB (Gelfand, 2010; Ward et al., 2003).

The inner ear encompasses the semicircular canals, the vestibule (utricle and saccule), and the cochlea. The first two structures comprise the vestibular system for sensing dynamic motions, accelerations, and position of the head with respect to gravity. The cochlea contains the organ of Corti, or the neural organ of hearing, and is responsible for transforming the fluid movement produced by the middle ear into nerve impulses (Ades & Engstrom, 1974). This transformation is accomplished by sensory cells (inner and outer hair cells) and supporting cells located on the basilar membrane within the organ of Corti (Maroonroge et al., 2009). These hair cells contain small hair-like projections called stereocilia that extend from each hair cell and lie just below the tectorial membrane. Fluid movement in the inner ear causes the basilar membrane to move up and down applying a shearing force on the stereocilia as they are pressed against the tectorial membrane. Neural impulses are developed as the stereocilia are bent (Maroonroge et al., 2009; Ward et al., 2003). Damage to the inner ear hair cells within the organ of Corti results in sensory hearing loss. Exposure to hazardous noise produces primarily sensory hearing loss (discussed later) (Casali, 2012b; Ward et al. 2003).

1.2.2 Human Auditory Sensitivity

Human audition, or the act of hearing, allows for the perception of speech as well as detection, recognition, and localization of sounds in a 3-dimensional, 360° spatial environment limited to an audible range of values on various parameters; this is known as the "envelope" (Maroonroge et al., 2009). The auditory sensitivity of the human ear can be defined by range of frequency, intensity, and duration, or the envelope of human hearing.

Envelope

The audible frequency bandwidth of hearing for the normal human ear ranges from 20 Hertz (Hz) to 20 kHz (Scharine, Cave, & Letowski, 2009). Sounds below 20 Hz, known as infrasound, can be heard by some individuals down to 2 Hz but are atonal (Moller & Pedersen, 2004: Gelfand, 2010; Scharine et al., 2009). The dynamic range of hearing extends from approximately 0 dB SPL to 140 dB SPL (re 20μPa) (Figure 8) (Scharine et al., 2009). Sound levels higher than 140 dB are audible but inflict pain and damage the auditory mechanism (Gelfand, 2010). The threshold of hearing is frequency dependent with the most sensitive range between 2000 and 5000 Hz (Gelfand, 2010). Hearing threshold becomes less sensitive for frequencies outside of this range, both lower and higher, requiring increased intensity levels to be audible (Sivian & White, 1933; Berger, 1981; Schechter, Fausti, Rappaport, & Frey, 1986).

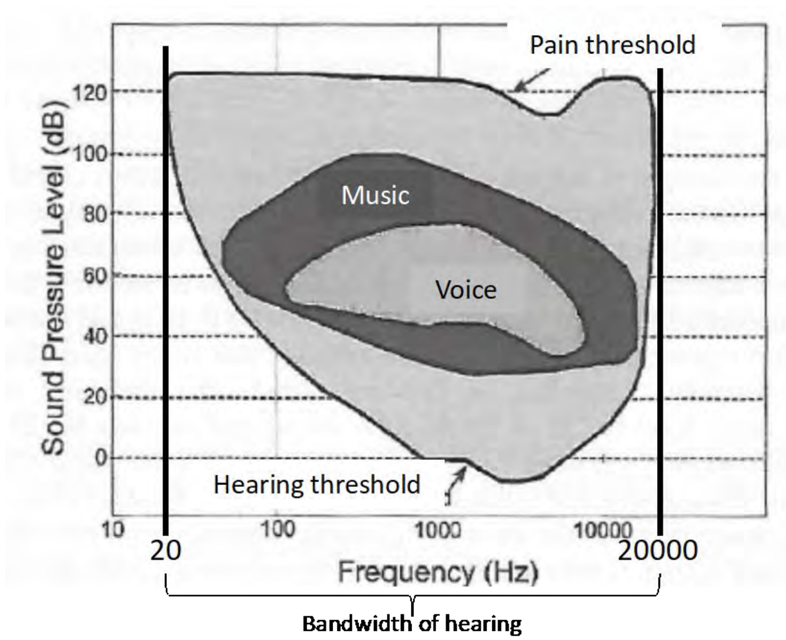


Figure 8. Human auditory envelope with areas of music (black) and speech (gray inside of black) (adapted from Scharine et al., 2009, Figure 11-3).

The binaural (hearing with two ears) threshold of hearing is approximately 2 dB lower than the monaural (single ear) threshold given both ears have similar sensitivity (Fletcher & Munson, 1933; Killion, 1978). Hearing threshold also depends on how the sound is presented. Minimum audible field (MAF) threshold refers to the threshold of hearing for sound waves arriving at the ear in free-field environment whereas the minimum audible pressure (MAP) threshold refers to the threshold of hearing from a stimulus arriving from an earphone occluding the ear canal (Scharine et al., 2009). The average difference between the MAF and MAP thresholds is 6 dB to 10 dB except for frequencies between 1500 Hz to 4000 Hz where the difference grows significantly up to almost 20 dB (Scharine et al., 2009). Frequency and intensity sensitivity are also affected when the duration of a sound is less than half a second. The threshold for tonality ranges from 60 milliseconds (msec) at 50 Hz, to 15 msec at 500 Hz, and 10 msec above 1000 Hz (Gelfand, 2010).

1.2.3 Auditory Localization

Auditory localization, defined as the ability to determine the direction and distance of a sound source, is critical to survivability and mission success in the military (Bevis et al., 2014; Price et al., 1989; Casali & Tufts, in press). Auditory localization is typically described in terms of the placement of a sound, based on a listener's hearing of it, in azimuth, elevation, and distance. Azimuth refers to the horizontal direction as an angle (θ), where θ = 0° is directly forward of the listener, θ = 90° is to the right, θ = 180° is directly behind (backward), and θ = 270° is to the left (Figure 9).

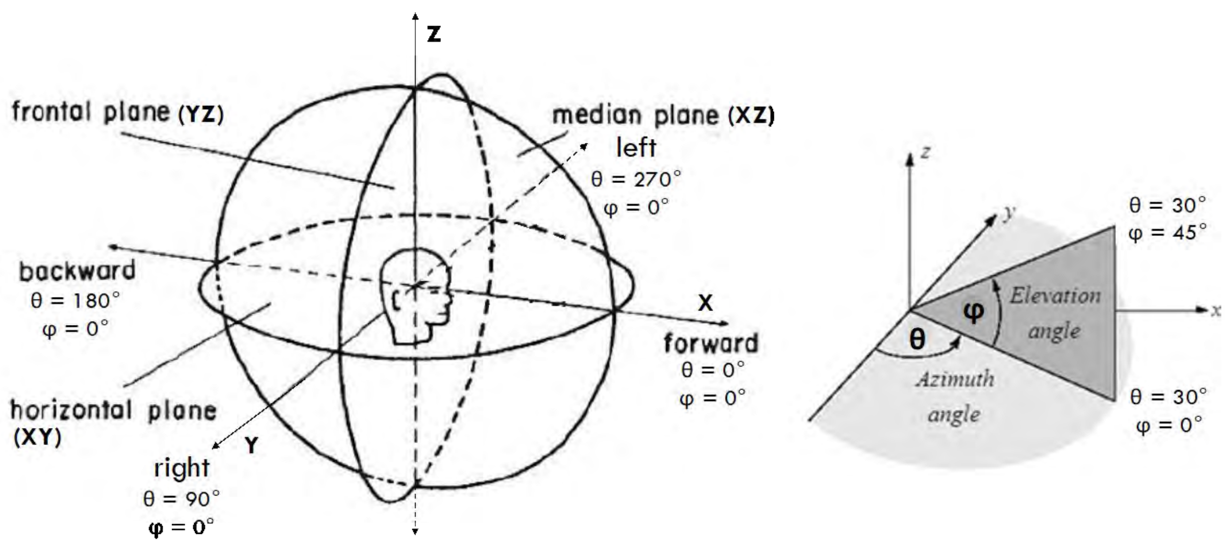


Figure 9. Anthropometric diagram with body planes and coordinate system used for sound localization. (adapted from Arazi, 2017, Figure 2; Moore, 2004).

The term median plane is used to describe the mid-sagittal plane that bisects the body vertically through the midline, dividing the body exactly in half (left and right sides). The median plane lies on the XZ plane in Figure 9, coplanar to azimuths 0° and 180°. Elevation, or vertical direction, is measured as an angle ϕ from the origin (centered in between the ears) to a height on the Z axis (Figure 9). When elevation is measured coplanar to the median plane, $\phi = 0°$ is directly in front of the listener, $\phi = 180°$ is directly behind, and $\phi = 90°$ is directly above (Gelfand, 2010). Distance refers to a judgment of how far the sound source originates from the listener. Distance can be measured in absolute judgment using some unit of distance, in simple discrimination judgment (closer-farther), or in sequential ratio judgment (half as far, twice as far) (Letowski & Letowski, 2012).

Localization Metrics

Auditory localization is normally measured in literature in terms of directional judgement accuracy, response times in locating sounds, and subjective workload measurements associated with the degree of difficulty to localize sounds. Accuracy judgements can be measured in both horizontal (azimuth) and vertical (elevation) dimensions. Localization accuracy typically entails measuring the angular distance between the presented sound source location and perceived location. The angular distance judgments can be measured in either relative localization (discrimination task) or absolute localization (identification task) (Letowski & Letowski, 2012). Relative judgments are used to compare one sound source location with another that is presented either simultaneously or sequentially. Absolute judgments involve localizing only one sound (Letowski & Letowski, 2012). In both relative and absolute judgements, errors in localization accuracy can be reported in terms of degrees or percentages. Two common types of error rates found in literature are local errors occurring with ±45° of the mean and reversal errors, or confusion errors, occurring at angles greater than ±90° and usually close to ±180° (Letowski & Letowski, 2012). Reversal errors are usually reported as front-back, meaning reporting a sound coming from the front that is actually presented in back of the listener, or back-front, reporting sounds in back when presented in front. Right-left and left-right reversals are also found in literature when the external ear is blocked preventing interaural cues (discussed later herein). While not as common as front-back or right-left reversals, spatial quadrant discrimination is another azimuthal measurement that is used to calculate localization accuracy. In spatial quadrant discrimination, the azimuthal field is divided into four equivalent angles typically bisected by the midline and a line perpendicular to the listeners ears. The Right-Front (RF) quadrant encompasses azimuthal angles from 0° to 90°, Right-Back (RB) quadrant from 90° to 180°, Left-Back (LB) quadrant from 180° to 270°, and Left-Front (LF) quadrant from 270° to 0° (Abel, Tsang, & Boyne, 2007; Butler, 1986). The spatial quadrant metric is not as precise as the absolute judgment or ballpark judgement that was used in this study but can still be used to identify the percentage of time a listener is correctly cued to the general direction of the sound. This basic auditory task allows soldiers to orient toward the threat and aides in cueing the visual modality.

Response times are used to measure how long it takes a listener to discern and indicate the location of a sound. This measurement is useful when comparing the effects of different listening conditions such as the open ear versus wearing a hearing protection device. Likewise, subjective workload measurements are used to compare the degree of difficulty associated with localizing sounds under different listening conditions. This study used absolute judgment, response times, and subjective measures in evaluating azimuthal localization under three listening conditions: open ear, an in-the-ear TCAPS, and an over-the-ear TCAPS.

Pinna Effects

The human auditory system is sensitive to the direction of the incoming acoustic signal. If a sound wave is generated from directly in front of or directly behind the listener, coplanar with the median plane, then the two ears receive the signal at approximately the same time (Maroonroge, Emanuel, & Letowski, 2009). If the sound wave is generated from the side of the head then the ear closest to the signal will receive the sound first and with greater intensity; these two cues correspond to the interaural time difference and interaural level difference, respectively, to be described later (Emanuel, Maroonroge, & Letowski, 2009). The binaural and monaural directional cues are what allows the human listener to localize sounds. In order to understand the human auditory localization ability, it is important to identify the role that the human anatomy plays, specifically the pinna and head.

The pinna is shaped like an irregular funnel that is attached to the head at a 15° to 30° angle (Emanuel et al., 2009). The complex shape of depressions and ridges of the pinna filter the high frequency aspects of a signal, above 4000 Hz, in a way that depends on the direction of the sound (Gelfand, 2010). Depending on the angle of the sound arrival, different ridges of the pinna reflect the sound causing changes to the acoustic spectrum of the sound (Emanuel et al., 2009). These changes in acoustic spectrum caused by the pinna are the dominant monaural cues for sound localization. Physical differences between the left and right pinnae for each individual, and between individuals, create interaural spectral differences (ISDs) in the localization cues (Letowski & Letowski, 2012). For mid- and high-frequency sounds, the spectral variations caused by the pinna are important directional cues for determining elevation and front-back distinctions (Blauert, 1997). Low frequency wavelengths are too long for the pinna to reflect in a way to provide localization cues (Emanuel et al., 2009). The spectral variations introduced by the pinna provide auditory cues to determine elevation and front-back distinction. Front-back (FB) and back-front (BF) errors are the most common type of reversal errors and occur most frequently when the sound is located directly on the median plane (Letowski & Letowski, 2012). At frequencies above approximately 1000 Hz, the pinna changes the spectrum of the sound waves reached at the eardrum in a direction-specific way. Humans utilize these direction-specific cues for front-back distinction. Shaw (1974a) measured variations in sound pressure level at the eardrum as a function of azimuth θ of incident sound for various frequencies. Low frequencies, at 300 Hz and 500 Hz in Figure 10, presented very little difference in sound pressure levels when presented at different azimuths around the head. However, as frequencies increased, a ±10 to ±15 dB variation was measured depending on the azimuth that the sound was presented.

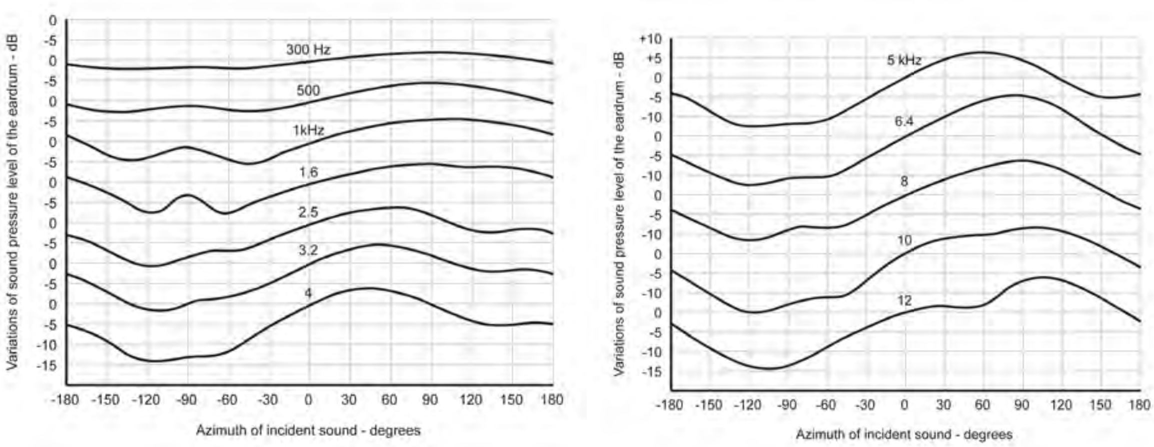


Figure 10. Pinna effects on frequencies shown by variations of average sound pressure level at the human eardrum as a function of azimuth θ of incident sound for various frequencies (adapted from Shaw, 1974a, Figures 6 & 7).

Vause & Grantham (1999) showed front-back confusions increased from 3% with the open ear to 4.2% when a U.S. Army-issue Kevlar helmet was worn and up to 19% when Kevlar and earplugs were worn. The increase in front-back confusion was a result of the Kevlar helmet blocking sound waves, preventing the waves from being reflected and modified by the pinna. As a result, listeners were not able to use pinna effects to distinguish the azimuth from which the sound was being presented resulting in higher front-back confusions percentages. Oldfield and Parker (1984) similarly reported front-back confusion rates of 12.5% and 26% when the pinna is occluded. Talcott et al. (2012) in an open-field experiment testing localization accuracy of gunshots reported mean front-back errors increased from 10% with the open ear to 26% to 30% using earplug style HPDs and to 31% using an over-the-ear (or pinna-covering) muff style HPD. All four HPDs significantly increased the percentage of front-back errors in both quiet and noisy background listening conditions (Figure 11) (Talcott et al., 2012).

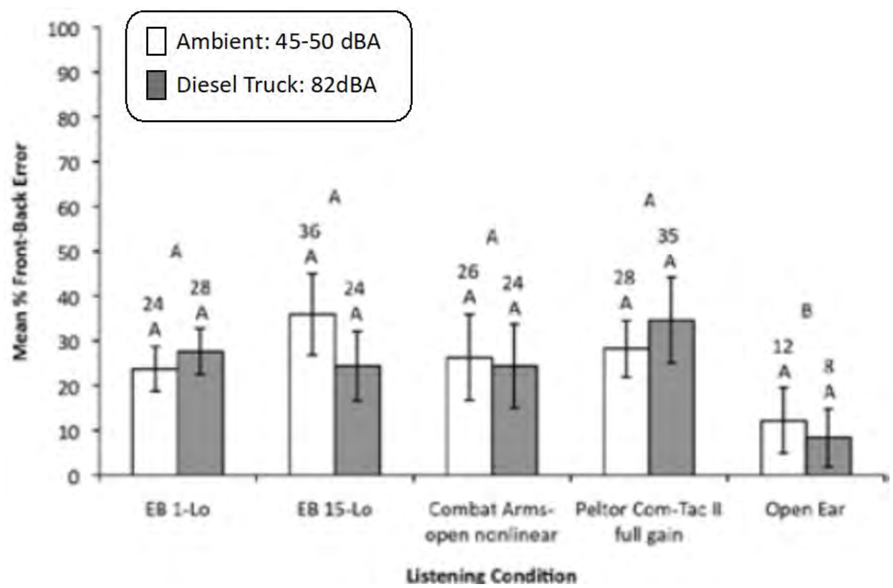


Figure 11. Effects of listening condition on percent front-back errors. Means and 95% confidence interval error bars. Top letters are main effects of listening condition. (adapted from Talcott et al., 2012, Figure 4).

The pinna is concave-shaped toward the front of the head enabling a greater funneling effect for sound presented directly in front of the listener. As a result, listeners tend to report sounds that are presented behind them as originating from the front (front given back or front-back error) more often than back-front errors (Abel, Boyne, & Roesler-Mulroney, 2009; Muller & Bovet, 1999; Shaw, 1974b). Oldfield and Parker (1984) showed that localizing sounds emanating from behind the head are impaired by the forward directed pinna (Figure 12).

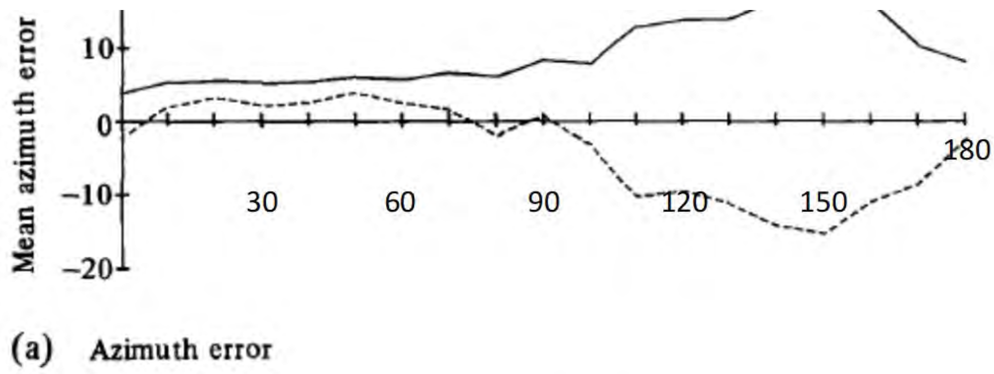


Figure 12. Mean absolute and algebraic azimuthal localization error for each azimuth position (adapted from Oldfield & Parker, 1984, Figure 3).

Localization accuracy is greatly decreased when the pinna is blocked or covered (Abel, 2009; Alali & Casali, 2011; Talcott et al., 2012; Hajicek et al., 2010; Oldfield & Parker, 1984; Scharine et al. 2007; Vause et al., 1999). Abel (2009) found that the percentage of correctly localized sounds decreased from 93.6% correct bareheaded to 79.7% when the ear was fully covered by a helmet with ears unoccluded and to 77.5% wearing a helmet with ear occlusion (Table 1).

Table 1. Overall % correct in horizontal plane sound localization. Effect of ear and helmet conditions (from Abel, 2009).

	Helmet			
Ear	None	Up	Half	Full
Unoccluded	93.6 (5.2)*	91.0 (6.7)	86.1 (6.0)	79.7 (5.3)
Occluded	83.4 (9.3)	85.7 (8.4)	81.3 (7.5)	77.5 (6.9)

*Mean (SD), N=10

Scharine et al. (2007) conducted a similar experiment differing ear coverage by 0%, 50%, and 100% using helmets and found as ear coverage increased, localization performance decreased. This supports findings of service members' perception that wearing military equipment to include helmets and hearing protection reduces situational awareness (Abel, 2008; Bevis et al., 2014). Unfortunately, in many occasions this loss of situational awareness causes military service members to forgo the use of hearing protection.

Pinna effects play a larger role in aiding with elevation localization than with azimuth localization. Hofman and Van Opstal (2003) used rubber molds placed inside the folds of the external ear, leaving only the auditory canal open, to test participants' localization accuracy under four listening conditions, natural open ear, mold in left ear, mold in right ear, and mold in both ears. The molds degraded localization accuracy in elevation on the side of the mold fitted ear but had less effect on azimuth localization (Figure 13). Figure 13A shows a nice symmetric grid of localization performance from -30° to +30° azimuth and -30° to +30° elevation. Figure 13B and 13C show the degradation on elevation localization accuracy due to the inability to use pinna cues on the side of the mold fitted ear indicated by the asymmetrical grid. When both ears were filled with the mold, elevation localization performance was completely abolished. Whereas the azimuth localization performance in the mold fitted ear experienced very little effect (Hofman & Van Opstal, 2003).

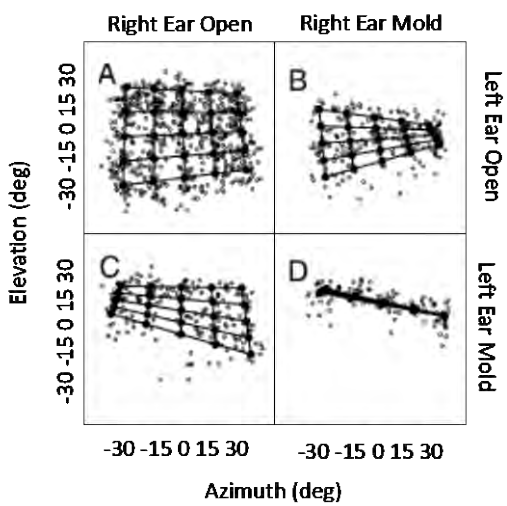


Figure 13. Results of applying the twist model fit to the localization data for four hearing conditions: A The control condition (both ears free); B a unilateral mold in the right ear; C a unilateral mold in the left ear; D molds bilaterally applied (adapted from Hofman & Van Opstal, Figure 4A-D).

Head Shadow Effects

The head casts an acoustical shadow when it is between the sound source and the ear receiving the sound. Except for sounds presented at 0° and 180° on the median plane, one ear will always be affected by head shadow. Sound waves diffract around the head. Low frequency sounds, below 1500 Hz, are capable of more diffraction and as a result arrive at both ears with similar intensity, or level. Frequencies above 3000 Hz with smaller wavelengths cannot diffract around the head as easily causing a decrease in intensity, or level, that is perceivable at the far ear (Gelfand, 2010). The difference in level presented at the two ears is known as the interaural level difference (ILD) and provides localization cues (discussed later).

The directional function of the pinna and head are known as the head-related transfer function (HRTF) (Gelfand, 2010; Scharine & Letowski, 2005). Figure 14 shows the HRTF for the right ear with sound sources presented at 45° (represented by near ear) and 315°, or 45° to the left of median plane, (represented by far ear) at various horizontal azimuths around the head (Gelfand, 2010). As previously discussed, the differences in sound level and spectrum provide localization cues. At low frequencies, the difference in sound intensity is small because the wavelengths are able to diffract around the head, yielding essentially no Interaural Level Difference (ILD), whereas higher frequency wavelengths are blocked causing greater ILDs and providing localization cues. (ILD will be further discussed in a later section herein.)

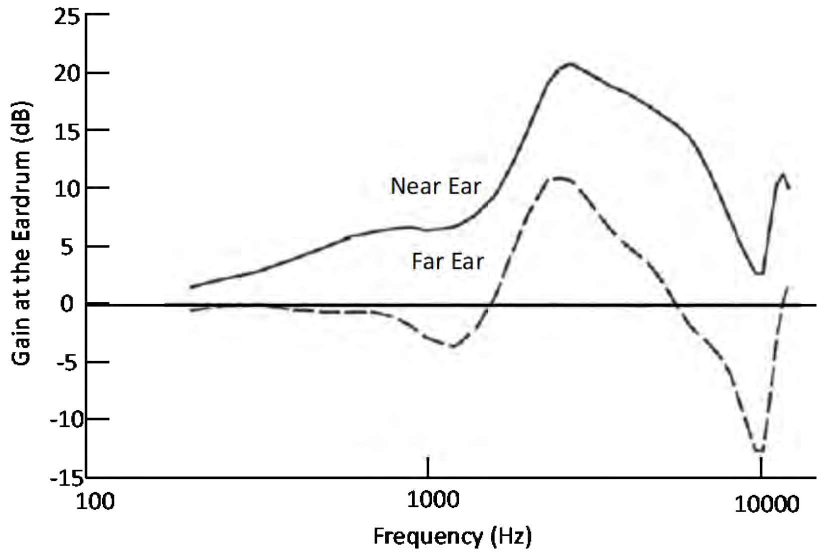


Figure 14. Horizontal head-related transfer function for sound sources located at 45° and 315° in the right ear (adapted from Gelfand, 2010, Figure 3.2, derived from Shaw, 1974b).

<u>*Head Movement Effects*</u>

Head movement can be used to overcome a lack of localization cues, especially found in low frequency signals. Head movement causes momentary changes in the sound spectrum at each ear. As a result, localization errors are larger when the listener's head is fixed than when the head is allowed to move (Muller & Bovet, 1999; Thurlow & Mergener, 1970). Listeners tend to turn their heads to the sound source except when the sound is emitted directly in front or behind, in which case the listener turns the head to both sides (Muller & Bovet, 1999). The main localization effect of head movement is to reduce front-back errors. Rotating the head allows a signal presented directly in front or behind to present interaural time differences and interaural level differences between the two ears (discussed later). Perrett and Noble (1995) found lateral head movements can be used to distinguish frontal elevation of low frequency sounds. Scharine and Letowski (2005) reported that head movement is only beneficial for sounds of duration greater than 400 to 500 milliseconds. Shorter sounds evidently disappear before the head movement is able to rotate and capture the sound, based on Scharine and Letowski's (2005) conclusions.

Muller and Bovet (1999) found that the pinna effects and head movement had an additive effect on localization ability in the horizontal plane. When one is taken away, there is a 10% loss in localization accuracy (Muller & Bovet, 1999). In addition, when the pinna was removed (filled with a mold), head movements were larger but localization accuracy did not reach the same level of performance with pinna effects (Muller & Bovet, 1999). However, when participants were able to use both pinna effects and head movement cues, localization accuracy was highest at every azimuth except for directly behind the participant (Muller & Bovet, 1999).

Binaural Localization Cues

Sound localization in the horizontal plane uses monaural (single ear) and binaural (two ears) cues. The two binaural cues are interaural level difference (ILD) and interaural time difference (ITD). ILD is the difference in the intensity of sound arriving at the two ears. ITD is the difference in time of arrival of a sound wave at the two ears. Figure 15 illustrates a sound arriving from a 45° azimuth angle creating a level difference in the right (near ear) and left (far) ear, ILD, and an arrival time difference between the right and left ear, ITD (Scharine et al., 2009). The sound wave arrives earlier at the near ear and with more intensity. ILD and ITD aid in localizing frequencies below 1500 Hz and above 3000 Hz, respectively, in the horizontal plane that originate outside the median plane (Casali & Lee, 2020). Frequencies between 1500 Hz and 3000 Hz are too high in frequency to provide time differences and have wavelengths that are too long to provide intensity differences between the two ears (Middlebrooks & Green, 1991).

Sounds that emanate on the median plane, 0° and 180°, arrive at both ears at the same time and with the same intensity limiting localization cues to pinna effects and spectral cues.

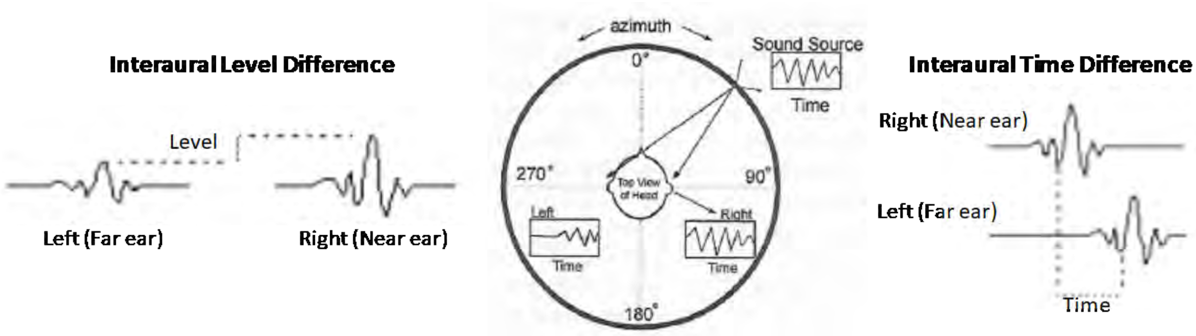


Figure 15. Interaural level difference (ILD) and interaural time difference (ITD) created by a sound arriving from a 45° azimuth angle (adapted from Scharine et al., 2009, Figure 11-29, and Kapralos et al., 2008, Figure 1).

Interaural Level Difference (ILD)

The interaural level difference (ILD), also referred to as interaural intensity difference (IID), is caused by the baffling effect, or acoustic shadow, cast by the head. The amount of baffling is dependent upon the size of the head and the frequency of the sound (Scharine, Cave, & Letowski, 2009). Low frequency wavelengths can be several times larger than the head and are able to diffract around the head resulting in negligible level differences between the two ears. At high frequencies, the acoustic shadow of the head causes differences in sound levels as much as 35 dB between the two ears (Middlebrooks & Green, 1991). The shadow results in lower intensity levels at the ear further from the sound source, that is, at the ear blocked by (in the shadow of) the head. The largest ILD occurs when the sound emanates perpendicular to one ear, or along the interaural axis at 90° or 270°. The ILD decreases as the position of the sound source lies closer to the median plane (Scharine & Letowski, 2005). Steinburg and Snow (1934) reported the variations in intensity level for various frequencies as a sound source is rotated in the horizontal plane and found ILD as large as 30 dB for 10 kHz (Figure 16).

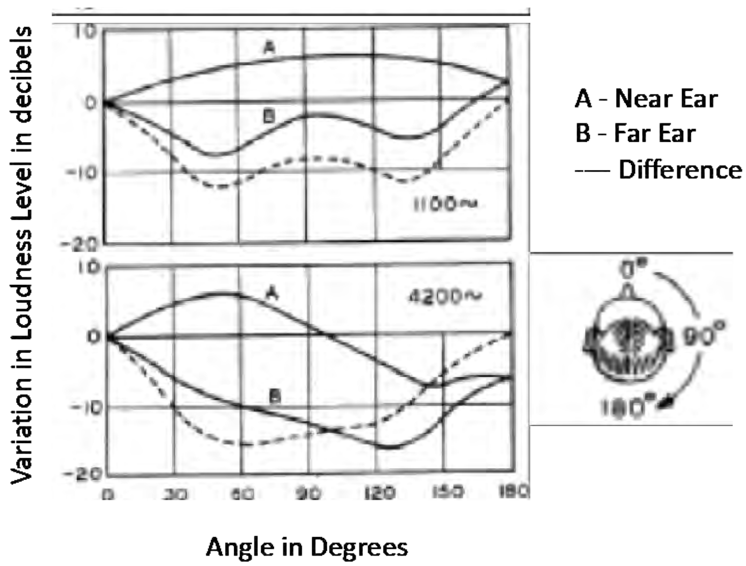


Figure 16. Variations in intensity level as a sound source is rotated in a horizontal plane around the head (adapted from Steinberg & Snow, 1934, Figure 2).

Interaural Time Difference

The interaural time difference (ITD) is caused by the delay in arrival time of the sound wave at the far ear. Sounds that originate on the median plane arrive at the ears at the same time and provide no ITD cues. However, sounds presented at any angle off the median plane arrive at the listener's ears at different times. Scharine and Letowski (2005) estimate the maximum achievable ITD is about 0.8 msec for a head diameter of 0.1 m and a sound wave velocity of 340 m/s when the sound wave is presented perpendicular to one ear, at 90° or 270°. The smallest perceivable ITD is about 0.2 msec to 0.3 msec when the sound is presented from 2° to 3° away from the median plane (Figure 17) (Scharine et al., 2009). This ITD is a result of the minimum perceived difference in azimuth equal to about 2° to 3° (Scharine et al., 2005). ITDs occur for clicks (short duration sound, less than 200 msec, that lose pitch quality), onset of a sound (beginning of a sound), and non-periodic sounds (sounds with non-repeating patterns).

Figure 17. Interaural time differences plotted as a function of azimuth (adapted from Scharine et al., 2009, Figure 11-31).

Interaural phase difference (IPD), as opposed to ITD, is the term more commonly used for continuous periodic sounds, such as sounds produced by pitched musical instruments, since the time delay of the sound arriving at the far ear is equal to a phase shift between the sounds arriving at the two ears (Scharine et al., 2005). The phase delay, as opposed to time delay, occurs with longer wavelengths associated with low frequencies, below about 1000 Hz, and requires both ears to hear the same cycle of a sound with no more than a 180° phase shift (Scharine et al., 2005). The strongest IPD cues occur in frequencies between about 500 Hz and 750 Hz (Scharine et al., 2005).

Minimum Audible Angle

The minimum audible angle (MAA) is the smallest angular separation of two sounds that is detectable by the human auditory system, also known as localization accuracy. Mills (1958) found an MAA of 1° for a sound emitting directly in front of the listener, 0°, at a frequency of 500 Hz to 750 Hz. The MAA increases as the source approaches 90° with the average MAA at 90° more than 40° (Mills, 1958). The smallest MAA were found in frequencies between 250 Hz and 1000 Hz. There was a significant increase in MAA for frequencies between 1000 Hz and

3000 Hz followed by a decrease from 3000 Hz to 6000 Hz and a smaller increase in MAA around 8000 Hz (Figure 18) (Mills, 1958).

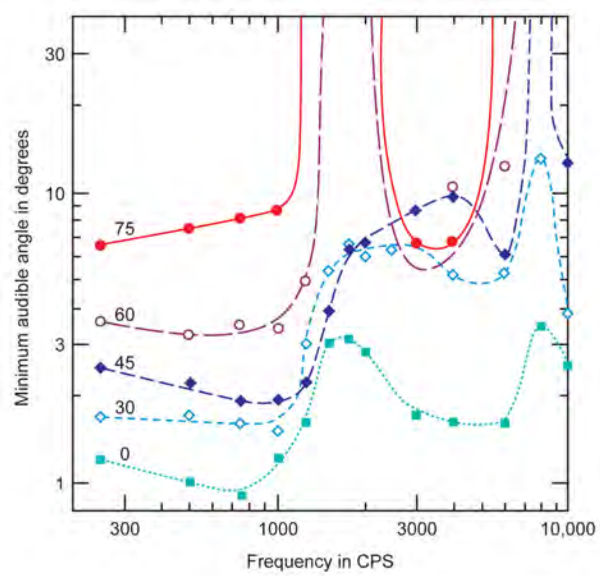


Figure 18. Average Minimum Audible Angle as a function of stimulus frequency. The parameter θ is the azimuth of the reference tone pulse (adapted from Mills, 1958, Figure 5).

Cone of Confusion

Binaural cues play a great role in horizontal localization accuracy but are only marginally useful for vertical localization and front-back differentiation. This is a result of locations within a conical region, known as the cone of confusion, where interaural level and time cues are the same (Letowski & Letowski, 2012). The cone of confusion is caused by the left-right head symmetry. The imaginary cone of confusion extends outward from the ear along the interaural plane. Sound sources anywhere along the circumference of the cone present the same interaural differences (Figure 19). Studies show the cone of confusion is the source for location errors in both the vertical and front-back directions (Makous & Middlebrooks, 1990, Oldfield & Parker, 1984). As discussed earlier, head movements are one method for resolving front-back (cone of confusion) errors (Muller & Bovet, 1999; Perrett & Noble, 1995).

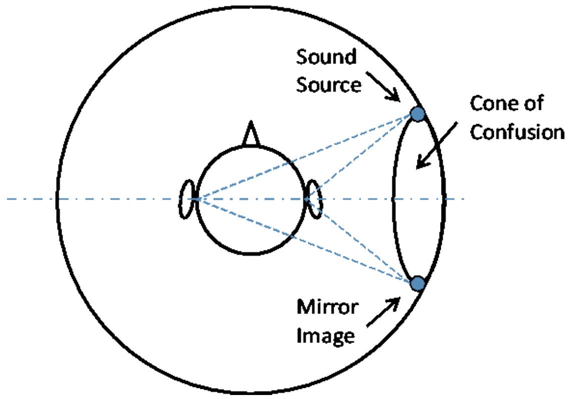


Figure 19. The concept of the cone of confusion.

Difference Limen and Just Noticeable Difference

Difference Limen (DL) is the smallest change in the value of a sound stimulus's parameter that can be detected. The two most frequently used difference limens, also referred to as the just noticeable difference (JND), are intensity DL and frequency DL (Scharine et al., 2009). Intensity DL is the smallest change in sound intensity level that is required to perceive a change in loudness. The relationship between the size of the differential threshold and the size of the stimulus for intensity DL follows Weber’s Law, or the change in stimulus magnitude is always a constant fraction of the stimulus magnitude (Scharine et al., 2009). The intensity DL is about 0.5 dB to 1.0 dB within a wide range of intensities greater than 20 dB above the hearing threshold (Riesz, 1928; Scharine et al., 2009). Intensity DL can be as small as 0.2 dB for pure tones in quiet and sound levels exceeding 50 dB SPL and as high as 3 dB for sounds in a natural environment (Figure 20) (Pollack, 1954; Riesz, 1928; Scharine et al., 2009).

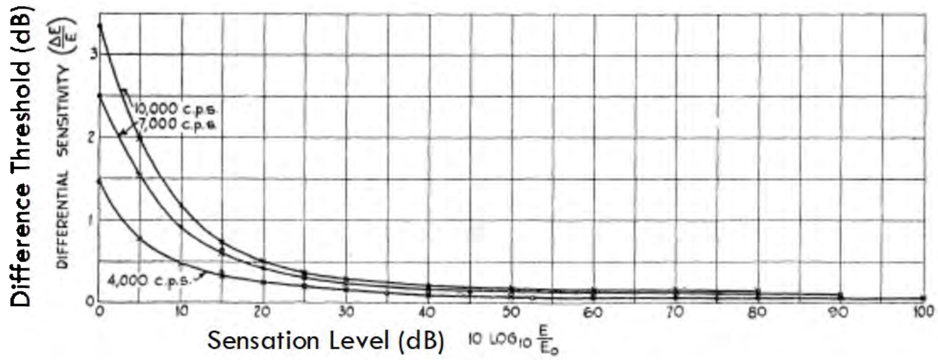


Figure 20. Intensity Difference Limen as a function of sensation level for 4000, 7000, and 10000 Hz frequencies. Sensation level plotted in units of 10 log10(E/E0) where E0 is the threshold intensity (adapted from Riesz, 1928, Figure 3).

Frequency DL is the minimum change in frequency required to detect a change in pitch. For low frequencies below 500 Hz, the frequency DL is relatively independent of frequency. The frequency DL at 1000 Hz is about 1 Hz and follows a logarithmic function as frequency increases with the frequency DL of about 10 Hz at 4000 Hz (Figure 21) (Scharine et al., 2009). The frequency DL decreases as the intensity of sound increases. Interaural difference limen are related to interaural cues and can be measured in intensity, frequency, and time. Rowland and Tobias (1967) found interaural DLs decrease as a function of level and as a function of frequency.

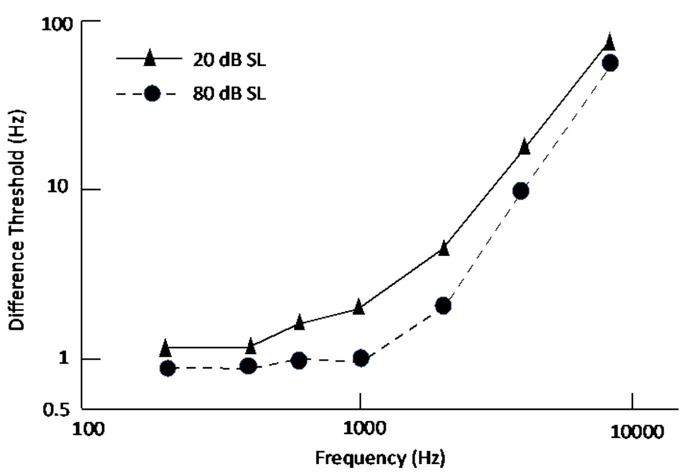


Figure 21. Frequency Difference Limen as a function of frequency for pure tones presented at 20 and 80 dB SL (adapted from Scharine et al., 2009, Figure 11-5).

Monaural Cues

Monaural cues are directionally dependent spectral cues that can be detected by one ear. Monaural cues are primary cues for sound localization in elevation and in front-back differentiation (see Pinna Effects herein) (Scharine et al., 2009). Although monaural cues are detectable with one ear, these cues occur in both ears simultaneously providing enhancements in localization ability.

Location and Spectral Cues

Localization of sound in the horizontal plane is best when the sound is presented directly in front of the listener (Blauert, 1997). Localization errors increase as the sound source is moved away from the median plane in either direction. Localization accuracy of sound in the vertical plane is best directly in front of the listener and decreases with elevation (Blauert, 1997). Blauert (1969/1970) found the spectral cues from the sound source provide vertical localization ability. Blauert recorded the spectral cues for a sound presented from the front and presented it directly behind the participants, and vice versa. All 10 participants reported the location of the sound signal from their originally recorded location (Blauert, 1969/1970).

Similarly, Butler and Musicant (1993) used spectral cues from broadband noise bursts previously recorded at the listener’s ear to test localization accuracy and found that monaural spectral cues contribute significantly to the accuracy of binaural localization. The basis of the monaural cue contribution is the spatial referents of stimulus frequency (Butler & Musicant, 1993).

Oldfield and Parker (1984) measured monaural localization for both azimuth and elevation using a white noise source. They found that monaural cues were much more effective in localizing elevation. The absolute error in elevation using monaural conditions was 12° compared to 9° under binaural conditions, whereas the absolute error in azimuth using monaural conditions was 30° to 40° compared to 5° to 10° under binaural conditions (Oldfield & Parker, 1984). From this study, it is clear that spectral cues induced by the structural characteristics of the pinna and head provide the major cues for elevation discrimination whereas binaural cues of interaural level difference and interaural time difference provide the major cues for azimuth discrimination. Figure 22 demonstrates the pinna effects on spectral cues by showing the fluctuations in sound intensity across frequencies from 1000 to 12000 Hz presented from different levels of elevation (Alali, 2011).

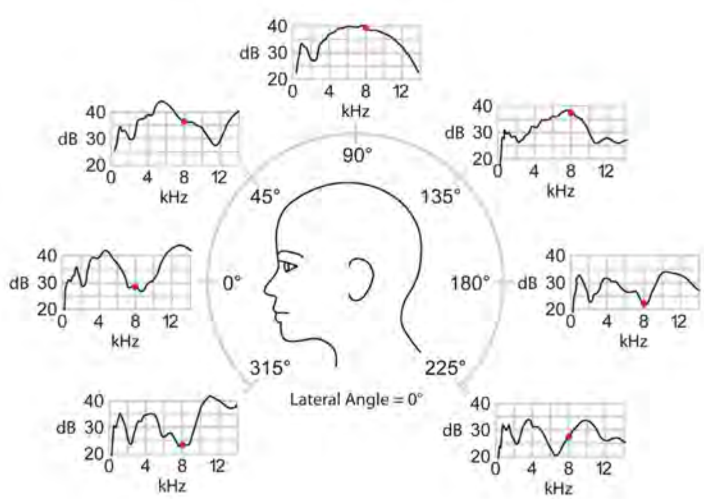


Figure 22. Spectral changes in sound intensity across frequencies from 1000 to 12000 Hz presented from different levels of elevation (adapted from Alali, 2011).

Stevens and Newman (1936) found that frequency had a significant effect on error in horizontal localization and percent front-back confusion. In their experiment, localization was most accurate below 1000 Hz and above 4000 Hz with the largest localization errors occurring with frequencies between 2000 Hz and 4000 Hz. Other studies have found similar, but narrower frequency bandwidth ranges to avoid for localizing spectral cues between 1500 Hz to 3000 Hz

(Blauert, 1997; Moore, 1997; Casali & Tufts, in press). In addition, front-back errors occurred at 35-40% in low frequency stimuli but were reduced to about 20% in high frequency stimuli (Stevens & Newman, 1936).

Numerous studies have reported that monaural spectral cues help resolve front-back confusion in the horizontal plane (Butler, 1986; Makous & Middlebrooks, 1990; Middlebrooks & Green, 1991; Muller & Bovet, 1999; Vause & Grantham, 1999). Butler (1986) tested monaural and binaural horizontal localization using an 8000 Hz-centered noise burst whose bandwidth was set at 2000, 4000, 6000, and 8000 Hz. The binaural localization accuracy exceeded monaural accuracy for sounds presented along the median plane but showed no significant difference in sounds presented in the middle section of the arc (Butler, 1986). In addition, binaural localization significantly improved at frequencies of 4000, 6000, and 8000 Hz compared to 2000 Hz where results were similar to monaural localization results (Butler, 1986). However, monaural localization only slightly improved as the frequency increased (Butler, 1986). When broadband noise was employed, binaural localization increased in accuracy for all azimuths compared to monaural (Butler, 1986). Makous and Middlebroooks (1990) tested localization of broadband sound sources varied in both azimuth and elevation using 150 msec sound bursts in a free-field. The smallest errors were about 2° in the horizontal plane and 3.5° in the vertical plane with the maximum errors occurring for more peripheral stimulus at about 20°.

Table 2 shows a summary of several localization studies reporting the localization accuracy associated with various signal types at the midline (Blauert, 1997).

Table 2. Localization in azimuth accuracy at midline by signal type (adapted from Blauert, 1997).

Reference	Type of signal	Localization accuracy
Klemm (1920)	Impulse	0.75° - 2°
Mills (1958)	Sinusoids	1.0° - 3.1°
Stiller (1960)	Narrowband Noise	1.4° - 2.8°
Blauert (1970)	Speech	1.5°
Haustein &Schirmer (1970)	Broadband Noise	3.2°

Duration

The duration of a sound is an important factor in localization because the ear is not capable of integrating the spectral information of extremely short sounds (less than about 100 msec) (Scharine & Letowski, 2005). Blauert (1997) observed that durations less than 100 msec resulted in the highest elevation localization errors. Pollack and Rose (1967) showed the average error in horizontal localization of a stationary sound source increased by 60% when the duration of the noise source was reduced from 50 msec to 20 msec. Thurlow & Mergener (1970) found that participants need about 2 seconds to achieve maximum localization performance. Longer sounds are easier to localize, especially if the listener is able to move their head, because it allows the listener more time to gain spectral information. Scharine and Letowski (2005) reported that head movement is only beneficial for sounds of duration greater than 400 to 500 milliseconds. Shorter sounds disappear before the head movement is able to rotate and capture the sound. Noble, Murray, and Waugh (1990b) observed that head movement had minimal effect on localization of a 500 msec sound but considerable improvement on localization when the sound duration increased to 1.5 sec. In addition, Rakerd and Hartmann (1985) observed that listeners' experienced difficulty localizing low-frequency tones with slow onsets.

Intensity

The sound intensity level of a signal has a greater effect on the localization ability in elevation than in azimuth. Davis and Stephens (1974) tested vertical localization effects from sound intensities of 10, 30, 50, and 70 dB SL (sensation level) and found that the mean absolute error decreased as sound intensity increased. In a follow-up experiment, Hebrank and Wright (1975) tested sound intensities of 40, 60, and 80 dB SL and compared results with Davis and Stephens (1974). For white noise intensities greater than 40 dB SL, localization accuracy is independent of intensity level with mean localization errors reported between 0.2° and 0.7°. However, intensities less than 40 dB SL resulted in decreased accuracy with decreasing sound level with mean localization errors reported at 1.1° for 10 dB SPL (Hebrank & Wright, 1975). Vliegen and van Opstal (2004) observed vertical localization accuracy decreased with decreasing sound levels down to the 36 dB SPL, which was above the hearing threshold for all participants. Sabin, Macpherson, and Middlebrooks (2005) confirmed that sound localization is inaccurate at sound levels near the detection threshold but improves as level increases.

Althsuler and Comalli (1976) (as cited by Sabin et al., 2004) tested the effect of sound level on horizontal localization using a narrow-band noise source by having participants report whether the sound source was presented on the midline (coplanar with the median plane directly in front or behind the listener) or to the left or right of the midline. The number of locations incorrectly classified as emanating from the midline increased as the sound level decreased (Altshuler & Comalli, 1976).

Movement and Doppler Phenomenon

Movement of an object can both aid and hinder localization ability. Numerous acoustical components of sound cues enable the detection and localization of moving objects to include sound intensity changes, interaural differences, signal duration, sound source velocity, and Doppler shift (Ericson, 2000). The movement of a sound source is typically measured in minimum audible movement angle (MAMA). The MAMA measures the minimum arc that is required to detect movement in a given direction. The MAMA increases with increased velocity, increased frequency, and increased displacement from the midline (Chandler, Grantham, & Leek, 1993). Grantham (1986) reported MAMAs of 5° for stimuli presented at 0° azimuth and 30° for stimuli present at ±90° azimuth. Localization of slow movement uses the same auditory cues as head movement but sound sources moving at quicker velocities are localized using Doppler cues.

Doppler shift is a phenomenon that occurs due to a change in frequency as a result of movement of an object relative to the observer (or observer movement toward a stationary object). The sound waves from the moving source are compressed as the object moves toward the observer, which has the effect of increasing the sound frequency. As the sound source moves away from the observer, the wavelengths spread out resulting in a lower frequency (Letowski & Letowski, 2012). This phenomenon applies both to sound sources approaching a listener from front and moving away from the listener behind, as well as, a sound source that passes the listener in a perpendicular manner (i.e. car passing a pedestrian waiting to cross the road). The degree of frequency increase and drop depends on the speed of the sound source in relation to the observer and the distance between them. The Doppler shift causes observers to perceive a rise in pitch as the sound source approaches and a decrease in pitch as the object moves away (McBeath & Neuhoff, 2002).

Rosenblum, Carello, and Pastore (1987) tested three types of acoustic cues: amplitude change, interaural temporal differences, and Doppler effect, to identify which variable contributed the most to localization of a moving sound source. The Doppler effect was shown to be the least preferred acoustic cue for relative importance for localization and demonstrated the most inaccurate location predictions with participant reactions (button push) occurring nearly a full second prior to the actual arrival (Rosenblum et al., 1987). Two studies performed by Getzmann, Lewald, and Gurski (2004 & 2007) found in all test cases the perceived final location was placed in front of the actual location attributed to the perceived increase in pitch of the approaching sound due to Doppler shift. In the study, the continuous noise sources resulted in more displacement at lower velocities while the pulsed noise resulted in more displacement at higher velocities (Getzmann et al., 2004). Localization accuracy in Doppler shift experiments are typically measured using relative judgments where listeners estimate the location of the sound at a given point in time in relation to a static target. This differs from the non-Doppler shift studies where localization measurements are judged in absolute terms within an angular range (e.g. ±15°).

Perrott and Musicant (1977) tested static and moving sounds of three velocities and observed forward displacement for both onset and offset of the moving sounds. In addition, the degree of displacement increased with higher velocities for onset. However, the offset location did not significantly increase with higher velocities (Perrott & Musicant, 1977). This displacement effect also occurred in studies where the sound source was moving directly toward or away from the observer (Neuhoff, 2001).

Environmental Effects

Military service members are required to perform operations in environments littered with noise from vehicles, aircraft, generators, weapon fire, explosions, radio communication devices, and shouting. In addition, many operations now take place in urban terrain where sound waves reflect off structures. Localization ability is difficult in listening conditions with high background noise and reverberation, so the effects of these conditions will be reviewed next.

<u>*Background Noise*</u>

Scharine and Letowski (2005) define background noise as sounds created by external sources through vibrations and reflections of sounds. Localization of a sound of interest is difficult in the presence of high levels of background noise because the signals of all sounds are hard to distinguish and often masked. Masking is an increase of threshold of a desired signal in the presence of an interfering signal (Casali J. G., 2012b). The greater the background noise, the harder it is to detect and localize the desired signal. Abouchacra and Letowski (2001) found localization ability decreased as sensation level decreased due to noise. In addition, back-front errors occurred more often than front-back errors but were both dependent upon the signal-to-noise ratio (SNR) and were largest when the sound source was located about 135° (Abouchacra & Letowski, 2001). In addition, the results supported the need for a +9 dB SNR for accurate localization of a speech sound source (Abouchacra & Letowski, 2001). Getzmann (2003) reported that the presence of a distracter sound caused participants to shift the position of a perceived target sound away from the distractor in both azimuth and elevation. (Getzmann, 2003). Casali and Lee (2016a) tested azimuthal localization under two noise conditions presenting a low noise condition consisting of a 50 dBA signal with 40 dBA background pink noise (+10 SNR) and a high noise condition consisting of an 85 dBA signal with 75 dBA background pink noise (+10 SNR). The mean absolute correct rate (exact azimuthal match) and mean ballpark correct rate (within ±15°) were higher for all listening conditions (open ear, HPDs, and TCAPS) under low noise conditions (Figure 23) (Casali & Lee, 2016a).

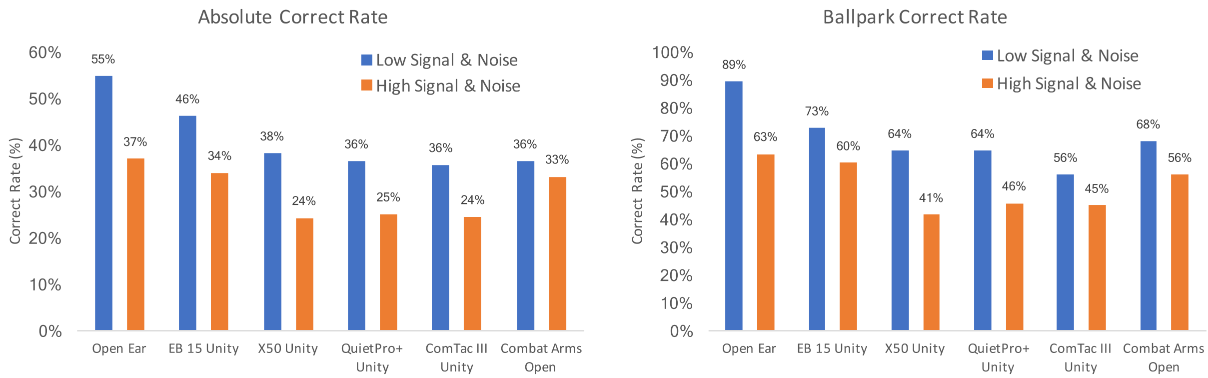


Figure 23. Azimuth localization test result, measured as absolute and ballpark (within ±15°) percent correct response rate at Low signal level of 50 dBA in 40 dBA pink noise background and High signal level of 85 dBA in 75 dBA pink noise background, SNR +10 (adapted from Casali & Lee, 2016a).

Reverberation

Sounds in urban environments bounce back and forth with relatively small loss in sound energy (Scharine & Letowski, 2005). The reflections provide false or ambiguous sound localization cues based more on the shape and reflectivity of the environment than the location of the original sound source. Sound reflections are classified as early reflections (arriving within 50 msec of direct sound), late reflections (arriving after 50 msec of direct sound), and echoes (late reflections that are distinguishable as separate events from direct sound). Early reflections increase sound intensity but do not impact localization (Scharine & Letowski, 2005). Late reflections combine with other sounds to create reverberation, the product of all sound reflections arriving at a given point (Scharine & Letowski, 2005). Reverberant sounds can degrade localization ability in both azimuth and elevation. Rakerd and Hartmann (1985) found that reflections off ceilings and walls have a greater negative effect on localization than floor reflections. They also reported that narrowband sounds and sounds with very slow rise times are the hardest to localize (Rakerd & Hartmann, 1985). In addition to types of reflections, reverberation times are used to measure reverberation within various environments.

Reverberation time is defined as the length of time required for a sound to decay 60 dB from the initial level. Ideally reverberation time for classrooms and offices should be about 0.6 seconds (RT60 = 0.6 sec), at frequencies between 250 – 4000 Hz (Ward et al., 2003). However, many rooms with hard surface floors and walls, including military buildings, have higher reverberation times. Hartmann (1983) tested localization accuracy in four room conditions including an absorbing room (RT60 = 1 sec), reflecting room (RT60 = 4 sec), low ceiling room (RT60 = 2 sec), and mirror reversed room (RT60 not reported), at 250 – 3000 Hz for all rooms. Hartmann (1983) observed that the low ceiling room had less localization errors than the absorbing or reflecting room. The author noted that very early reflections are more likely to be confused with the direct sound and as a result are more likely to negatively affect localization accuracy (Hartmann, 1983). Hartmann (1983) hypothesized that floors and ceilings reflect sound waves in an azimuth that agrees with the azimuth of the direct sound and actually reinforces the perception of the sound source azimuth. However, side walls reflect sounds in a way that do not agree with the direct sound azimuth resulting in localization decrement (Hartmann, 1983). It is expected that the PALAT system will be employed in various room settings included semi-reverberant environments. As a result, experiments performed to validate the training impacts of the PALAT system were conducted in a semi-reverberant office with furniture.

Listener Effects

Movement

Movement can facilitate and decrement localization ability. The effects of movement depend on direction of the movement of the listener, sound source, or both, as well as the velocity of the movement. Listener movements when a sound source is stationary present similarly to moving sound with a stationary listener (as discussed earlier). Ericson (2000) found the acoustical cues used in localization by a moving listener include sound intensity changes, interaural differences, signal duration, velocity, and Doppler shift. Summarizing previously discussed findings on movement of sounds, listener movement results in early arrival prediction to the sound source (Rosenblum et al., 1987), predicted locations ahead of, or in the direction of motion, the actual location (Getzmann et al., 2004; Getzmann & Lewald, 2007), and increased displacement of location with velocity (Getzmann et al., 2004).

Another listener movement characteristic that effects localization ability is listener head movement (discussed earlier).

Hearing Loss

Studies have shown that asymmetrical (unilateral) hearing loss decreases azimuthal localization performance. Viehweg and Campbell (1960) compared localization ability between normal hearing participants and participants with unilateral hearing impairments using speech played at 30 dB above the better ear threshold. Results showed that participants with unilateral hearing loss had a higher number of localization errors and higher size of error (45° azimuth = 1 error size) than normal listeners in both quiet and in noise (Viehweg & Campbell, 1960). Newton and Hickson (1981) tested localization of normal hearing adults and participants with otological or neurological disorders using a 500 Hz pure tone and a narrow band centered at 500 Hz. Those participants with unilateral hearing impairment and middle ear conditions had difficulty localizing both pure tones and noise as shown in their abnormal mean errors (Newton & Hickson, 1981). The authors hypothesized the loss in localization ability was due to the loss of binaural cues of time and phase (Newton & Hickson, 1981).

Symmetrical hearing loss up to 40 dB is reported to have little effect on localization in the horizontal plane (Blauert, 1974; Letowski & Letowski, 2012). However, Nobel, Byrne, and

Lepage (1994) found a significant decrease in horizontal localization performance for participants with conductive hearing loss. The authors reported that deficits in localization accuracy in different regions of auditory space could be related to different types of hearing loss. Nobel et al. (1994) showed associations between vertical plane discrimination and high frequency sensitivity. In addition, the author reported associations between front-back discrimination and mid-to-high frequency sensitivity (Noble et al., 1994). Participants with bilateral high frequency hearing loss performed worse than participants with conductive hearing loss (Letowski & Letowski, 2012; Noble et al., 1994). Participants involved in this investigation were screened for normal hearing with bilateral symmetry.

1.2.4 Distance Judgments

The ability to estimate distance to a sound source is less accurate than the ability to determine azimuthal direction (Zahorik, Brungart, & Bronkhorst, 2005). Distance estimates are usually overestimated for close sounds (less than one meter) and underestimated for far sounds (greater than one meter (Zahorik et al., 2005). The primary auditory cues for determining the distance of a sound source from the listener are sound intensity, spectral changes, reverberation, and motion (Scharine et al., 2009). Mershon and King (1975) found that distance judgments for familiar sounds can be made based primarily on sound intensity cues by comparing previous knowledge of the sound at different distances in certain environments. The distance of unfamiliar sound sources is difficult to estimate using sound cues alone. Sound intensity is the primary distance cue in free-field settings, relying on the physics imposed by the inverse square law where sound decreases by 6 dB every time the distance doubles (Coleman, 1963). In closed spaces, reverberation of sounds may act as the primary distance cue because the inverse square law does not hold due to the short distances that sound is allowed to travel. Spectral changes of sounds due to the environment, such as the effects of humidity that attenuates primarily high frequencies, provide distance judgment cues for sounds traveling long distances in free-field (Scharine & Letowski, 2005). Again, this phenomenon requires some familiarity of the sound and shows the importance of using military specific sounds for training and testing service members on auditory distance and localization.

Detecting sounds as early as possible is vital to the safety of service members and mission success of military operations. Early warning of hazards and enemy threats allows for evasion or element of surprise (firing first). Alali and Casali (2012) found that HPDs decrease detection distances of vehicular backup alarms compared to the open ear. More specifically, as the attenuation of hearing protectors increases, detection distances tend to decrease (Alali & Casali, Auditory backup alarms: distance-at-firstdetection via in-situ experimentation on alarm design and hearing protection effects, 2012). Casali et al. (2009) simulated combat scenarios, reconnaissance and raid missions, in an in-field experiment using passive and active HPDs and found that some active HPDs, equipped with electronic gain features, provide the ability to detect sounds at greater distances than the open ear but that there are significant variations in distance detection between devices and device-specific effects based on the detection task and environment. Clasing and Casali (2014) confirmed variations in HPDs/TCAPS detection ability compared to the open ear using military relevant signals of foreign spoken language, an AK-47 charging, and gunshot. Four of the five HPDs/TCAPS tested performed worse than the open ear for detection distance for all sounds and one TCAPS provided some early warning compared to open ear for detecting foreign language and AK-47 charging (Figure 24) (Clasing & Casali, 2014).

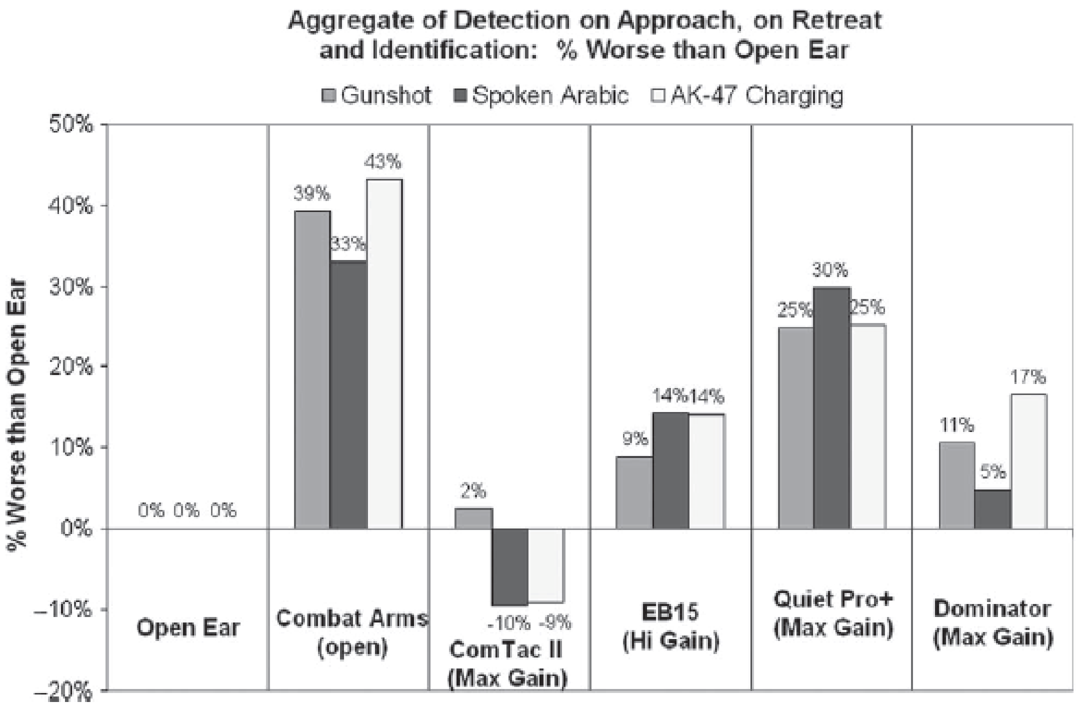


Figure 24. Aggregate of detection on approach, detection lost on retreat and identification with percent worse labeled as compared to open ear (adapted from Clasing & Casali, 2014, Figure 7).

1.2.5 Auditory Factors in the Military

At his change of responsibility, General Milley, Chief of Staff of the U.S. Army, stated:

> "It is on the ground where the United States Army, the United States Marine Corps, and the United States Special Operations Forces must never ever fail. And to succeed in that unforgiving environment of ground combat, we must have forces that have both capacity and capability, both size and skill. They must be manned. They have to be equipped and they **better be trained**" (Milley, 2015, pg 1).

Unfortunately, the equipment, capabilities, and training of auditory localization in the military are not where they need to be to ensure success in ground combat. In fact, training for auditory localization is nonexistent, or minimal at best (Casali & Robinette, 2014). Situational awareness is critical to the survivability and mission success in military operations (Abel, 2009; Talcott et al., 2012; Hajicek et al., 2010; Lee & Casali, 2017, McIlwain & Gates, 2008). Service members, especially ground combat forces, must be able to shoot, move, and communicate. All three of these military tasks require good hearing in order to accurately detect, identify, recognize, and **locate** enemy threats.

Military Auditory Requirements

Hearing readiness is the process to ensure that every Soldier, Sailor, Airman, and Marine has the necessary hearing capability and personal protective equipment for readiness, deployment, lethality, and survivability (U.S. Army Public Health Center, 2017). Price et al. (1989) developed mathematical models in an attempt to quantify the importance of service members' ability to hear. The models used hearing thresholds for normal hearing, poor hearing, and poor hearing plus temporary threshold shift (TTS) and compared them to spectra of military sounds. Effects were calculated for detection of sounds of enemy and ability to communicate. Normal hearing is highly effective in noise but even the modest hearing losses and/or wearing of hearing protectors had a profound effect on military performance (Price et al., 1989). Price et al. (1989) estimated hearing loss could result in a reduced area that could be monitored acoustically by more than 30-fold or cut warning times by a factor of more than 100. Understanding the effectiveness of good hearing, the military has set clear guidelines for noise exposure, the requirement of hearing protection, and the standards for testing noise emissions and attenuation capabilities of military equipment (Department of Defense, 2015). The Department of Defense steady state noise limits shall be less than 85 dBA and peak pressure levels of impulsive noise will be less 140 dBP (dB Peak), at the ear (protected or unprotected). All personal protective equipment, vehicles, weapons, and facilities must follow American National Standards Institute (ANSI) respective standards for noise levels and noise reduction (Department of Defense, 2015).

The military has established auditory fitness for duty (AFFD) standards that set the hearing thresholds and profiles that dictate whether a service member is able to perform their duties safely and effectively (Brungart, 2014; Casali & Tufts, in press). The three service branches all use a profiling system for hearing with categories H1 through H4 based on pure tone thresholds (Table 3) (Department of the Air Force, 2013; Department of the Army, 2015; Navy and Marine Corps Public Health Center, 2008).

Table 3. Hearing Profile Criteria for the U.S. Air Force, Army, and Navy.

	Air Force AFI 48-123	Army AR 40-501	Navy TM 620.51.99-2
H-1	Unaided hearing loss in either ear with no single value greater than: 25 dB at 500 1000 2000 Hz, 35 at 3000 Hz, 45 at 4000 Hz, and 45 at 6000 Hz.	Audiometer average level for each ear not more than 25 dB at 500, 1000, 2000 Hz with no individual level greater then 30 dB. Not over 45 dB at 4000 Hz.	Unaided hearing loss in either ear with no single value greater than: 25 dB at 500 1000 2000 Hz, 35 at 3000 Hz, 45 at 4000 Hz, and 45 at 6000 Hz
H-2	Unaided hearing loss in either ear with no single value greater than: 35 dB at 500 1000 2000 Hz, 45 at 3000 Hz, and 55 at 4000 Hz; no requirement for 6000 Hz.	Audiometer average level for each ear at 500, 1000, 2000 Hz, or not more than 30 dB, with no individual level greater than 35 dB at these frequencies, and level not more than 55 dB at 4000 Hz; or audiometer level 30 dB at 500 Hz, 25 dB at 1000 and 2000 Hz, and 35 dB at 4000 Hz in better ear. (Poorer ear may be deaf.)	Unaided hearing loss in either ear with no single value greater than: 35 dB at 500 1000 2000 Hz, 45 at 3000 Hz, and 55 at 4000 Hz; no requirement for 6000 Hz.
H-3	Any loss that exceeds the values noted above, but does not qualify for H-4.	Speech reception threshold in best ear not greater than 30 dB HL, measured with or without hearing aid; or acute or chronic ear disease.	Any loss that exceeds the values noted in the above definition.
H-4	Hearing loss sufficient to preclude safe and effective performance of duty, regardless of level of pure tone hearing loss, and despite use of hearing aids.	Functional level below H3.	Hearing loss sufficient to preclude safe and effective performance of duty, regardless of degree of pure tone hearing loss, or unknown hearing loss values. The H-4 profile indicates an incomplete follow-up or a requirement for a Medical Evaluation Board.

In 2006 and 2012, the Army and Marine Corps respectively, began requiring audiograms be administered before entering service, upon separation, before and after each deployment, and annually during service (Brungart, 2014). However, none of the service branches test service members on their ability to perform auditory tasks including sound localization as part of their

AFFD (Brungart, 2014). Brungart (2014) states, "the challenges are due to both the expense of setting up these systems and to issues with room acoustics and other factors that might cause performance to vary from test location to test location."

Communication

One of the reasons hearing is so important in the military is the necessity to communicate. In a landmark study, Peters and Garinther (1990) studied the effect speech intelligibility on armor crew performance. Results proved that degraded speech intelligibility resulted in a significant decline in target identification, kills, kills with one shot, time to complete mission, and crew survivability rates. On the modern battlefield, service members must be able to communicate with troops in close proximity on the ground by voice commands and over the radio with adjacent forces and combat support elements. In a recent focus group study, military personnel reported they were often expected to understand speech without visual cues such as in low visibility situations and when using a radio to communicate (Bevis et al., 2014). Tactical missions require stealth and the ability to communicate without being detected and during intense combat operations where noise overwhelms natural hearing ability. Unfortunately, there is often very little time or warning for when stealth operations become overloaded with noise.

Noise Induced Hearing Loss

Noise is both a combat multiplier and hazard on the battlefield (Donahue & Ohlin, 1993). Noise can mask troop movement and provide the element of surprise or prevent detection of enemy threats. However, regardless of the scenario, the pervasive noise in the military environment is one of the most prevalent occupational hazards (Donahue & Ohlin, 1993). Overexposure to noise can result in a temporary threshold shift (TTS) or can permanently injure the neural structure of the ear, resulting in noise-induced hearing loss (NIHL) (Casali & Gerges, 2006). Military service members in training and combat operations are frequently susceptible to both high levels of continuous noise from generators, vehicles, and weapons and extremely high level impulse noises from explosions. McIlwain et al. (2008) reported that explosions were the single largest cause for injury in Operation Iraqi Freedom and accounted for 47% of all medical evacuations. Explosions and high-caliber gunshots present extremely intense acoustic impulse noise that can immediately damage the conductive chain in the ear or cause neural damage by dislodging hair cells in the organ of Corti (Casali & Gerges, 2006). Table 4 shows examples of hazardous noise levels that service members are exposed to during training and in combat operations (Melzer, Scharine, & Amrein, 2012).

Table 4. Steady-state and impulse noise levels from various military equipment types (adapted from Melzer et al., 2012, Figure 9.1).

Noise type	Source	Noise Level
Steady-state	HMMWV	94 dB in vehicle (moving)
	M1A2 Abrams (tank)	115 dB in vehicle (moving)
	CH-47 helicopter	107 dB in cockpit
Impulse	M16A2 rifle	157 dB at shooter's ear
	M249 machine gun	159.5 dB at gunner's ear
	Javelin missile	172.3 dB at gunner's ear
	Multi-role anti-armor antipersonnel weapon system (MAAWS)	190 dB at gunner's ear

Both noise-induced temporary threshold shift (NITTS) and noise-induced permanent threshold shift (NIPTS) can result from acoustic trauma from sudden intense noise but are more likely to occur over time from exposures that are repeated over a long period of time (Casali & Gerges, 2006). These repeated exposures have a cumulative effect on hearing sensitivity (Casali & Gerges, 2006).

In a recent interview, the Chief Scientist for the Audiology and Speech Center at Walter Reed National Military Medical Center, stated that according to the Center for Disease Control, "veterans are 30% more likely to suffer from hearing impairment than nonveterans, and those who have served since September 11^{th}, 2001, are four times as likely to have hearing loss than their civilian counterparts" (Brungart, 2014). The number of new compensation recipients for tinnitus and hearing loss have risen from 92260 in 2010, to 149429 in 2016, and 63583 in 2010, to 77622 in 2016, respectively (United States Department of Veterans Affairs, 2016). Many service members who suffer from hearing loss continue to serve without knowing or without reporting until they are screened. The military AFFD standards and annual audiograms will help identify hearing loss but the primary preventative measure relies on the use of hearing protection devices.

1.2.6 Hearing Protection Devices

Hearing protection devices are designed to reduce the level of noise at the ear to prevent exposure to noise-induced hearing loss. HPDs in the military come in several forms and offer different degrees of attenuation and amplification of sound signals. Military HPDs that allow for integration with communication systems are referred to as Tactical Communications and Protection Systems (TCAPS). Military HPDs that provide the same level-dependent attentiation benefits but do not incorporate a communication capability are called TCAPS-Lite, referred to herein after as TCAPS. HPDs can be classified into two broad categories, passive (conventional) and active (electronically augmented). Casali (2010a; 2010b) conducted an extensive classification and technical overview of HPDs for both passive and active.

Conventional, or passive, HPDs in the military typically consist of earplugs, earmuffs, or helmets that enclose the ears. The conventional HPDs usually offer adequate protection according to DoD noise attenuation standards but result in degraded auditory performance (e.g. Abel, 2008; Bevis et al., 2014; Casali, 2010a; Talcott et al., 2012). Conventional passive HPDs attenuation is linear or the same regardless of sound level, at least up to extremely high levels where the HPD behaves nonlinearly if it exhibits resonance, sudden loss of seal, or other acoustical-dynamic effects (Casali J. G., 2010a). To fix this issue, the military fielded non-linear or level dependent passive HPDs including the Combat Arms earplug, now in its fifth design generation. The level dependent passive HPDs provide moderate attenuation for sound levels up to about 110 dB and then sharply increased attenuation for sound over 110 dB (Casali J. G., 2012a). The Combat Arms earplug still performed poorly in detection and localization tests (Talcott et al., 2012; Clasing & Casali, 2014).

Active, or electronically augmented, HPDs or TCAPS in the military typically consist of earmuffs or earplugs that feature a microphone mounted on the external surface and small loudspeakers mounted within the earmuffs or internal to the earplug's body (Casali J. G., 2010a). The electronics are designed to boost certain frequency ranges of sounds that include speech and warning signals but reduction of gain for incident high-intensity sounds, usually those that are above about 85 dBA, though this varies among devices (Casali J. G., 2010a). The active TCAPS devices typically include a full amplifier shutoff level of about 110 dBA in order to prevent overexposure to explosions and loud impulse noises (Casali J. G., 2010a).

HPDs and TCAPS are evaluated and measured on their ability to provide protection from continuous and impulsive noise exposure. Every HPD is tested for at least its passive spectral noise attenuation under the S3.19-1974 standard promulgated by the American National Standards Institute (ANSI), under an EPA federal regulation (40CFR211) (ANSI, 1974; EPA, 2002). Each HPD receives a Noise Reduction Rating (NRR) score that indicates the level of noise reduction provided by the device. While this is extremely beneficial for gauging the attenuation performance of hearing protection in passive mode, it overlooks the impacts that the HPD has on auditory detection, recognition, identification and localization, for which not military or ANSI test standard exists. Furthermore, it also does not quantify the effects of the level-dependent electronics of the TCAPS or HPD on attenuation provided when the sound level changes over time, or when gunfire is encountered; this requires a much more complex test using ANSI S12.42-2010 (ANSI, 2010).

Effects of HPDs and TCAPS

The deleterious effects of conventional hearing protection devices on the ability to detect and localize sounds is well documented (e.g. Abel, 2008; Alali & Casali, 2012; Bevis et al., 2014; Casali, 2012; Letowski et al., 2014; Noble et al., 1990b; Vause, 1999). Earmuff style HPDs were particularly disruptive to sound localization because they prohibited pinna effects which have been shown to be a primary source for elevation localization and front-back discrimination (Abel et al., 2007). The degraded auditory ability caused service members to forgo the much-needed hearing protection in order to increase situational awareness (Abel, 2008; Bevis et al., 2014).

As a result of poor performance and lack of use, considerable improvements have been made in developing augmented HPDs and TCAPS that are designed to maintain or enhance the listener's auditory performance, to include detection and localization (Casali & Lee, 2016b). Studies of TCAPS show improved detection and recognition performance is possible with active HPDs due most likely to electronic amplification, or gain, features (see Distance Judgments section herein) (Casali et al., 2009; Clasing & Casali, 2014; Lee & Casali, 2016; Lee & Casali; 2017). However, detection performance varied greatly between active HPD devices, indicating more analysis is needed to ensure optimal performance in various military settings (Alali & Casali, 2012; Casali et al., 2009; Clasing & Casali, 2014; Lee & Casali, 2016; Lee & Casali; 2017). Giguere et al. (2013) found normal hearing listeners showed improved speech recognition compared to the open ear when wearing both an over-the-ear and in-the-ear TCAPS, by 35% and

15% respectively. However, active HPDs and TCAPS continue to negatively impact the localization subtask of auditory situation awareness.

While advertised to improve situation awareness, studies testing the effects of TCAPS on localization show degraded performance resulting in greater localization errors and significantly more front-back confusions. Abel et al. (2007) tested localization of active HPDs using a 75 dB SPL, 300 msec broadband noise and found unoccluded listening resulted in significantly higher localization with 94.1% accuracy followed by Nacre QuietPro® (previous Marine Corps TCAPS) with 71.1% and Racal Slimgard II with 69.2%. More alarming were the localization accuracy results with the Nacre in passive mode, 51.7%, and Rascal using active noise reduction (ANR), 36.1% (Abel et al., 2007). Service members could very easily use either of these devices in the wrong configuration due to a lack of training or by accident, severely degrading localization capability without even being aware. In a follow-on study, Abel et al. (2009) tested the Nacre QuietPro® while wearing a helmet with varying ear coverage and found similar degraded localization accuracy results for all helmet types using the Nacre compare to the open ear.

A series of experiments performed at the Virginia Tech Auditory Systems Laboratory (VT-ASL) resulted in similar findings. In a field experiment testing horizontal localization of actual gunshots, Talcott et al. (2012) found the Peltor™ Com-Tac II earmuff ranked lowest in localization but all four active HPDs performed significantly worse than the open ear in both absolute and ballpark localization accuracy (Figure 25). Wearing an active HPD or TCAPS reduced the mean percent correct response (within ±22.5°) by 18%, 20%, and 28% compared to the open ear in a quite setting (45-50 dBA ambient noise) and 29%, 24%, and 45% in the presence of high background noise (82 dBA background noise) (Talcott et al., 2012). In addition, participants' mean response time increased significantly while wearing three of the four passive or active HPDs and TCAPS compared to the open ear response time (Talcott et al., 2012).

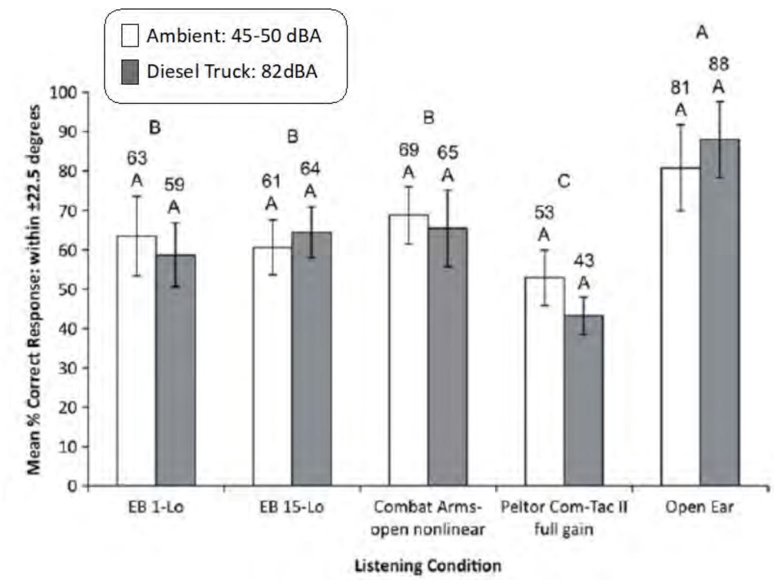


Figure 25. Effect of listening condition by noise level on percent correct response within 22.5° ("ballpark"). 95% CI error bars with means labeled above. (adapted from Talcott et al., 2012, Figure 2).

Casali & Lee (2016a) tested localization accuracy, measuring both absolute and ballpark (within ±15°), using three TCAPS and an active in-the-ear HPD in low (50 dBA) and high (85 dBA) pink noise. The TCAPS were significantly outperformed by the open ear for all horizontal localization tests and the active HPD was outperformed but not by a statistically significant degree (Casali & Lee, 2016a). Figure 26 shows the percent worse performance in azimuthal localization task with a ballpark (within ±15°) measure of accuracy compared to the open ear (Casali & Lee, 2016a). In addition, mean response times increased for all TCAPS devices and open ear condition when presented with the higher background noise (Casali & Lee, 2016a).

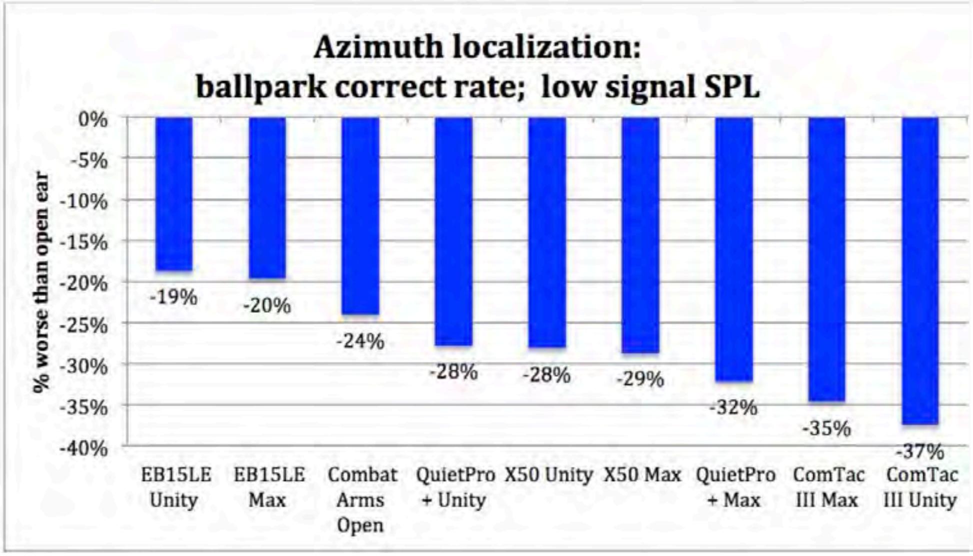


Figure 26. Azimuth localization test results, measured as percent worse than open ear (adapted from Casali & Lee, 2016a, Figure 45).

In a promising experiment, Casali & Robinette (2014) demonstrated the plasticity of the human auditory system to learn and demonstrated improved azimuthal localization abilities with the open ear, and also with TCAPS (both in-the-ear and over-the-ear) using an acclimation-training regimen. This experiment is discussed in detail later herein.

1.2.7 Auditory Perceptual Skills Acquisition

Heald and Nusbaum (2017) define auditory perception as an active cognitive process that incorporates learning (Heald & Nusbaum, 2017). They further define skills as something that is learned and has to be practiced as opposed to an ability that is a natively human biological endowment, or a starting place from where skills can be learned (Heald & Nusbaum, 2017). Thus, listeners should be able to improve their auditory abilities, to include localization, through an active cognitive learning process or acclimation-training regimen.

Auditory Learning

The human auditory system is a continuous sensory mechanism that is constantly learning and adapting. The ability to interpret binaural and monaural cues relies on the size and shape of the head and contours of the pinna. Each individual's interpretation of these localization cues is different. Yet humans are able to perform localization tasks even as they grow, learning and adapting to changes in auditory cues. Clifton, Clarkson, Gwiazda, Bauer, and Held (1988) conducted a study and found that large shifts in interaural time differences among infants, children, and adults suggests that the human auditory system is able to recalibrate the associations between interaural time differences and spatial location (Clifton et al., 1988).

Held (1955) tested the adaptive learning ability of the human auditory system to changes in interaural time and level cues using pseudophones. The pseudophones used consisted of two matched hearing aids that presented sound signals that were attenuated and rotationally displaced around the vertical axis of the head in order to produce illusory auditory localization by changing the relationship between the respector and the actual direction of the sound. Held (1955) conducted a study using pseudophones to displace the interaural axis by 22° and observed that after one hour, adults were able to partially correct localization errors induced by the pseudophones. This experiment showed that the human auditory localization ability could adapt to changes in interaural cues. Hofman, Van Riswick, and Van Opstal (1998) confirmed the adaptive learning of the human auditory system by modifying the outer ear of participants using a mold placed in the concha effectively changing spectral cues derived from pinna effects. Hofman et al. (1998) observed that localization of sound elevation was immediately and dramatically degraded but steadily improved as the participants were able to relearn localization within six weeks. In a similar study, Van Wanrooij and Van Opstal (2005) demonstrated that participants fitted with binaural pinna molds were able to regain normal localization performance within several weeks.

Auditory learning has been shown in several studies to occur through practice, either in natural settings or during repeated testing. Noble and Byrne (1990a; 1991) reported evidence of localization acclimatization as a result of hearing aid wearers daily activities. In the study, hearing aid wearers were tested for localization accuracy using their own hearing aids and two other versions. The behind-the-ear (BTE) group consisted of listeners who regularly used hearing aids that were placed behind the pinna with a microphone in a tube that was placed in the concha. The in-the-ear (ITE) group consisted of listeners who regularly used hearing aids that were placed inside the external auditory canal. The listeners were tested using their own devices and then with the opposite style hearing aid which they were not familiar. The authors found that the BTE group localized better with BTE while the ITE group localized better with their ITE hearing aids (Noble & Byrne, 1990a; 1991). These results showed that humans are able to relearn interaural cues while using hearing aids. Abel et al. (2007) determined that participants improved localization performance just by practicing localization exercises over the course of 16 testing periods. Participants showed an improvement of 8.1% in localization accuracy while wearing a TCAPS compared to only a 5% improvement with the open ear (Abel et al., 2007).

Auditory Training

Localization training regimens have been shown to improve sound localization performance in listeners with normal hearing and impaired localization abilities (Wright & Zhang, 2006). Mikaelian (1969) trained auditory localization in participants using pseudophones to shift the interaural axis by 30° and confirmed the ability to compensate for distortion after only 20 minutes of training. However, Mikaelian (1969) reported that auditory adaptive learning did not occur as quickly as visual learning indicating that training sessions for auditory localization may take more time. Abel and Paik (2004) tested localization training effects on two groups, one with ability to see visual cues and one blindfolded. The training regimen consisted of five, daily, 30-minute training sessions with three stimuli, a 500 Hz tone, a 4000 Hz tone, and a broadband noise. The blindfolded group demonstrated a greater range of improvement in localization accuracy (Abel & Paik, 2004). The training session consisted of one block of forced-choice speaker identification trials for each stimuli of 15 random speaker presentations for a total of 120 trials. The broadband signal produced the highest localization accuracy scores (Abel & Paik, 2004).

Bauer, Matuzsa, Blackmer, and Glucksberg (1966) conducted an adaptive localization training experiment where participants simulated partial hearing loss by wearing a plastic earplug for up to 3 days. A control group was administered localization tests every six hours without feedback of localization accuracy until they could achieve pretest equivalent accuracy scores. Following the same pretest procedure, the experimental group underwent a training regimen that provided immediate feedback on localization accuracy during each session. Localization errors were immediately repeated from the same speaker to train the participant. The training regimen effectively reduced the time to achieve pretest localization accuracy from 65 hours in the control group to an average of 5 hours (Bauer et al., 1966)

Dufour, Ratelle, Leroux, and Gendron (2005) developed an auditory localization training program to enhance localization ability in new users of bilateral cochlear implants. A single participant was tested and trained using a semi-circular horizontal array consisting of 11 speakers. The participant was tested and trained in three sitting positions: front facing directly toward the middle speaker, semi-circular array to the left and semi-circular array to the right. The test stimuli consisted of a 65 dBA broadband noise signal for 1.5 second duration presented from each speaker twice for each seating position. Training stimuli consisted of ecological sounds (traffic, pedestrian, etc.) for a 3 second duration or longer. The participant listened to the signal and moved their head to face toward the source wearing a head mounted laser to indicate response direction. The study consisted of two training periods totaling 21 lessons, each lasting one hour in duration. Progress in the form of quicker response time, greater localization accuracy, and fewer front-back confusions were indicated during training. Changes made from pretest to posttest to accommodate the participant's comfort during training, such as standing instead of sitting during training, caused the results to be inconclusive. However, the authors and the participant expressed confidence in the participant's increased localization abilities (Dufour et al., 2005).

A recent study tested the effects of a localization training procedure presented by spatialized auditory stimuli using the participant's customized HRTF presented over headphones (McMullen & Wakefield, 2017). The training consisted of both axial training, where participants were trained on front-back resolution, and random-source training, where targets were placed at a randomly determine range and azimuth. Participants were tasked with placing a cursor on a screen in a virtual auditory environment at the location where the sound was perceived. McCullen and Wakefield (2017) found that the training procedure maintained or significantly improved localization accuracy whereas the performance of the non-trained group did not improve.

Determining the optimal training regimen is important to ensure effective auditory training. Russell (1977) performed a localization training study to identify listeners' ability to adapt to impairments to normal hearing posed by the use of a passive earmuff. One group of participants were tested for three consecutive days and provided feedback while a second group received two localization training sessions each day for a five-day period. The authors found that feedback from testing provided to the first group resulted in increased localization accuracy but not at a significant level. The group that received training improved localization accuracy from about 50% to about 70% after only five days of training (Russell, 1977). The author concluded that “listeners cannot adapt to earmuffs” despite the fact that an approximate 20% increase in localization accuracy was realized after a very short training regimen (Russell, 1977).

Auditory localization training while wearing TCAPS has recently been reported in two studies with promising results. Casali and Robinette (2014) conducted a study with two groups of participants, one wearing in-the-ear TCAPS and one wearing over-the-ear TCAPS. Both groups were administered localization pretests with both open ear and while wearing the TCAPS. Each group then received 12 one-hour training sessions consisting of three, 15-minute auditory localization tasks that provided feedback on accuracy. The localization training increased localization accuracy by 17% in the ITE group and 19% in the OTE group, and by 18% averaged between the two TCAPS groups (Figure 27) (Casali & Robinette, 2014). The results showed that localization training while wearing either TCAPS resulted in an increase in localization accuracy when using the same TCAPS that was close to the improved open ear levels. However, training on the ITE TCAPS did not benefit localization performance with the OTE TCAPS, and vice versa, demonstrating that training was exclusively beneficial for the "trained TCAPS" only and that no crossover benefit occurred between devices.

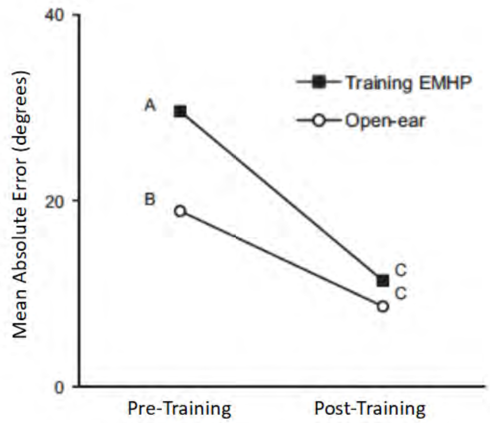


Figure 27. Combined training groups localization performance for open ear and training TCAPS. Means with same letter are not significantly different at $p < 0.05$. (adapted from Casali & Robinette, 2014, Figure 2).

The authors defined the results within a military context stating, “10 degrees of separation would enable the listener to auditorily discern whether a shooter was standing at the front or the back of a typical length tractor-trailer located perpendicular to his or her position at a distance of approximately 400 feet away” (Casali & Robinette, 2014).

In a more recent effort that extended the application of the DRILCOM auditory situation awareness test battery that was developed for the Department of Defense Hearing Center of Excellence, Casali and Lee (2016b) adapted DRILCOM to provide a pilot acclimation-test system to test azimuthal localization, and that system exists at present. The test is conducted with 12 loudspeakers that are 30° apart mounted at head level in a circle around a seated listener. In the initial experiment, participants received 12 learning units, of about one-hour each, which were divided into three training sessions. The protocol is as follows. First, a training session where signals are presented from each speaker in a clockwise or counterclockwise direction and location of the sound is presented along with the sound signal. Second, a training session is randomized and the direction is not indicated. In both sessions, response feedback is provided after the participant selects their response. Finally, a test session is conducted where 24 targets are shown on a monitor (representing the 12 speakers with 30° separation and 12 dummy speakers in-between the actual speakers). A sound signal is played through each of the 12 speakers 3 times in a randomized order. No feedback is provided during the test, as it is currently devised (Casali & Lee, 2016b). The results from the initial experiment using this setup and protocol was that participants wearing the previous Army TCAPS, Invisio® X50, were able to achieve ballpark accuracy (within ±15° of true location) approximately equal to that achieved with the open ear after only 5 learning units, at an 89% correct level (Casali & Lee, 2016b). Absolute accuracy levels wearing the Invisio® X50 increased at similar rates to the open ear, improving from 60% to 80% correct, but requiring 12 learning units (Casali & Lee, 2016b). A second, proprietary, TCAPS, with a very different external microphone design only achieved a 13% improvement in absolute accuracy and never approached the open ear performance with training, proving the importance of an auditory localization training system to be incorporated into military training and auditory fitness for duty programs (Casali & Lee, 2016b). Another conclusion from the initial experiment was that the DRILCOM-based system could be used to eliminate certain devices from consideration for deployment, in view that the second TCAPS, due to its unresponsiveness to training effects, provided poor localization performance even after a lengthy training regimen.

As a guide to successful military training in general, Wolfle (1946) detailed important principles of learning that are necessary to incorporate in any training plan. The author first introduced two overarching principles that were common knowledge. First, overlearning is important but is often disregarded due to limited training time (Wolfle, 1946). If time permits, there is value to continued training even after the level of localization accuracy has plateaued. Second, skills are lost during periods without practice (Wolfle, 1946). This means localization training may need to be repeated after long periods of time without using hearing protection. Wolfle recommends the following additional principles be incorporated into military training:

1. Distribution of practice – spread training sessions out as opposed to training for longer hours in a shorter time period.
2. Active participation – include the learner in the training experience of other learners.
3. Variation of material – vary drills and material to increase learning.
4. Accurate records of progress – accurate progress records aid instructors in tailoring training and motivate trainees.
5. Knowledge of results – timely feedback motivates trainees and identifies errors to prevent them from being practiced during subsequent training.
6. Systematic lesson plans – detailed instructions enable novice trainers to provide effective training.

These principles prescribed by Wolfe (1946) were used in designing the training protocol and the PALAT software programming. The PALAT system spread training sessions out over time and required participants to take breaks between training sessions for Phase II and Phase III of this investigation. The learner self-guided themselves through the training after the initial learning unit and a thorough demonstration. Performance feedback was immediately given to the participants on localization accuracy for all tests. Lastly, an additional learning unit was added to the improved training protocol found in Phase I in order to train past the point of proficiency in the open ear listening condition to increase learning.

1.2.8 Auditory Localization Apparatus Designs

Research, including some of the studies discussed above, has clearly evidenced the benefits of localization as a combat multiplier and the ability to improve localization accuracy using a training system. However, the majority of the localization testing apparatus designs require specialized soundproof rooms, are too expensive to field at every military installation, are too large to fit in unit training rooms, or are too fragile to be portable systems for field deployment. A review of static apparatus design features used in testing auditory localization was conducted and key attributes from each system were considered when designing the PALAT system.

Common horizontal localization apparatus design types include circular (halo), semi-circular, spherical (elevation and horizontal), and sound booms (Figure 28).

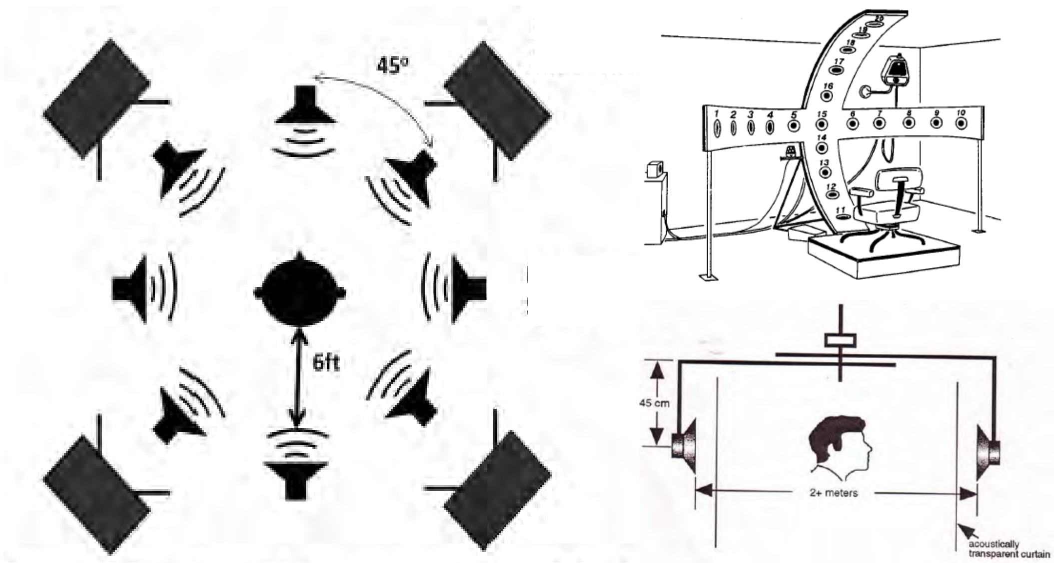


Figure 28. Localization apparatus designs, circular (left adapted from Alali & Casali, 2011), semi-circular horizontal and vertical (top right adapted from Noble et al., 1990b), sound boom (bottom right adapted from Abouchacra & Letowski, 2012).

The Virginia Tech Auditory Systems Laboratory has a 3-meter diameter circular, horizontal and vertical (front) localization apparatus consisting of 12 speakers (horizontal) and 3 additional vertical speakers (placed above the 330°, 0°, and 30° horizontal speaker) housed in a hemi-anechoic room (Figure 29). The speakers are mounted on a circular steel pipe ring located approximately 1.14 meters in height above the floor. The speakers and metal ring are covered with acoustically transparent black fabric to conceal the location and number of speakers present. The investigator control station is located outside of the speaker ring and consists of a desktop computer used to initiate the auditory tests and provide data capture and recording. A small control station with a computer monitor and mouse is located in the middle of the ring to allow the participant to control the experiment and respond. The system uses Behringer Behritone

C50A powered speakers to deliver the auditory signal. In addition, an Optimus 1850 compact disc player delivers background pink-noise through a QSC CX1102™ power amplifier to four JBL SoundPower SP215-6 loudspeakers (Casali & Lee, 2016a).

Figure 29. VT-ASL DRILCOM test apparatus located in the hemi-anechoic test room.

The U.S. Army Research Laboratory has two auditory situation awareness configurations that are used to test sound localization. The sphere room is approximately 15 ft by 15 ft and houses a spherical localization device containing 57 loudspeakers radially separated by 25° (Letowski et al., 2012b). The dome room is approximately 19 ft by 24 ft and houses a localization apparatus with a horizontal plane of speakers separated by 2° and two vertical arcs of speakers separated by 10° (Figure 30) (Letowski et al., 2012b). This system has been employed in several experiments which have been covered above in this literature review.

Figure 30. Army Research Laboratory Sphere Room (Left) and Dome Room (Right) (adapted from Letowski et al., 2012b).

The SoLoArc, developed and located at St. Olaf College, is a semicircular array consisting of 37 speakers placed every 5° measuring over 8 feet in diameter (Figure 31) (Westerberg, Balhorn, Tyshynsky, Olson, Brichetto, Gaston, & Loebach, 2016). The device is interfaced to a PC using 2, 96-channel NI USB I/O devices and is controlled via Matlab allowing for stimuli of any frequency and intensity to be delivered from each speaker (Westerberg, et al., 2016). Westerberg et al. (2016) pilot tested localization performance using SoLoArc and found that localization accuracy was comparable to results obtained using ITD and ILD stimulus presented through headphones. While the SoLoArc is capable of performing auditory localization tests, the main purpose of the device is to utilize current technologies such as functional near-infrared spectroscopy and electroencephalography to collect psychophysiological data during task performance (Westerberg et al., 2016).

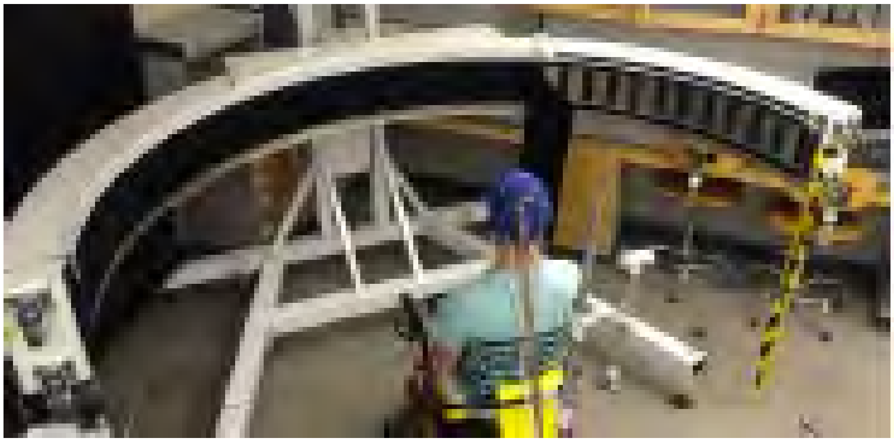


Figure 31. SoLoArc localization system (from Westerberg et al., 2016).

A review of auditory localization study apparatus design features is shown in Appendix A. Of note, an extensive literature search and review yielded no results for a *portable* auditory situation awareness or auditory localization device at the time of initiation of the design process of the PALAT system.

1.3 Research Gaps

An extensive review of literature identified a need for research on auditory localization skill acquisition and the development of a portable system for imparting localization skills that are transferable to the real-world environment. Numerous studies have shown the necessity of auditory situation awareness as a combat multiplier on the battlefield (e.g. Casali et al., 2009; Donahue & Ohlin, 1993; Hajicek et al., 2010). However, the United States military currently has no formalized auditory localization training program and has not incorporated auditory localization testing into their auditory fitness for duty programs. Service members are equipped with Tactical Communications and Protective Systems (TCAPS) that typically degrade localization ability without proper training and are sent to operational environments unaware of the deleterious effects. Casali and Lee (2016b) demonstrated that the INVISIO X50™ and the Nacre-Honeywell Quiet Pro+™, the respective TCAPS for the Army and Marine Corps at the time of the study, decremented ballpark localization accuracy by approximately 30% compared to the open ear; this should be of great concern and factored in when future generations of TCAPS are selected for deployment. As a result, service members who perceive this significant loss of situation awareness are faced with the difficult decision to operate handicapped by equipment or to forego hearing protection and risk being exposed to and injured by hazardous noise, especially severe gunfire emissions which can cause immediate neural hearing damage (Abel, 2008; Bevis et al., 2014). Fortunately, recent studies on auditory localization while wearing TCAPS have shown the ability to improve localization accuracy through training in relatively short periods of time (Casali & Lee, 2016b; Casali & Robinette, 2014). Yet, there is an operational gap between laboratory localization training and testing apparatus designs and the obvious need for a requirement to field a portable training aid for military units. As a result, the Office of Naval Research (ONR) provided a contract to develop and validate an auditory localization training protocol and a portable auditory localization training system to increase service members' auditory situation awareness while wearing TCAPS, and which would double as a means to test and evaluate TCAPS prior to fielding. This investigation addressed the current void for a Portable Auditory Localization Acclimation Training (PALAT) system as part of a larger initiative to develop an operational training and test-evaluation system for both a laboratory and portable setup. This study was also the first of its kind to evaluate the transfer-of-training effect of auditory localization by training and testing participants on a localization task in a semi-laboratory setting using a dissonant tonal signal, and subsequently testing in a field setting using a military-relevant signal (live gunshot blanks). Previous experiments did not incorporate lab and field environments within the same study nor did participants train on signals different from the test stimuli.

1.4 Research Objectives and Hypotheses

The overarching objective for this study was to develop and validate a highly portable, objective, measurement- and feedback-intensive, and user-operable system that instills and improves auditory localization skills in the service member over a regimented series of learning units. This system contemplates application for skills acquisition/improvement both with the open ear and while wearing a TCAPS device. The first phase, hereinafter referred to as Phase I, of this sequential research plan, which was completed by another doctoral student from the U.S. Army, hereinafter referred to as “K. Cave dissertation,” was to develop a training *protocol* that improves auditory localization for use in both a full-scale laboratory and a portable configuration. This dissertation study focused on the second phase, hereinafter referred to as Phase II, of designing and developing the actual PALAT *system*, to comprise a portable, tablet-based version of the localization training system in order to train service members during any available time period. A final phase, Phase III, also performed as part of this dissertation in cooperation with K. Cave dissertation, was to validate the localization training protocol performed using the portable PALAT system against an actual, proven in-field localization test (per Talcott et al., 2012). With these key requirements in mind, the specific objectives of this study were:

1.4.1 Objectives

1. Develop a portable auditory localization training system to train and test service members’ ability to localize with both the open ear and with a TCAPS or other hearing protection device.

    a. Convert laboratory localization training system software to run in real-time on a laptop PC or tablet with a simple user interface that provides training feedback and testing results.

b. Develop an audio presentation array consisting of a small amplifier and an optimal number of loudspeakers to present localization training and test signals.

c. Design and build a portable localization training system, no bigger than 7 feet in diameter and able to breakdown and fit into a hand-carried shipping trunk.

d. Upon completion of Phase I, incorporate the improved localization training strategy into the portable localization training system.

2. Evaluate the skills acquisition training effectiveness of the portable localization training system via human participant experimentation over a sequential experimental design, and thereafter, compare the results against the full-scale laboratory grade system to determine the correspondence of PALAT results to a full-scale laboratory DRILCOM-based system which has been optimized for maximum training benefit.

3. Upon completion of Phase II, evaluate the validity of the transfer-of-training effect from the results obtained in the laboratory setting with the PALAT system to an in-field, real world environment experiment.

1.4.2 Hypotheses

1. A Portable Auditory Localization Acclimation Training (PALAT) system is technologically and economically feasible and can be developed using primarily off-the-shelf components, to enable a laptop-controlled system which can be intuitively operated by a military trainee, and which provides acoustically-accurate localization cues that will impart training benefits in a non-laboratory indoor environment, such as a barracks.
2. Participants' testing scores using the PALAT system will demonstrate similar learning benefit to that imparted by the full-scale DRILCOM-based system, and do so with a similar number of learning unit sessions.

3. The PALAT system training effects will transfer to the field environment, as evidenced by improved localization ability for gunshot signals encountered in an in-field, real world test environment, both with the open ear and with a TCAPS device.
4. Training on the PALAT system and testing in the field environment will be sensitive to auditory localization performance differences with the open ear and with an in-the-ear and over-the-ear TCAPS.

CHAPTER 2. Design of a Portable Auditory Localization Acclimation Training (PALAT) System

2.1 PALAT Objectives

The primary objective of the overarching study was to develop an innovative, portable auditory localization acclimation training system that incorporated an improved training strategy to train and test service members' ability to localize with both the open ear and with a TCAPS device. A series of auditory localization studies at Virginia Tech led to the development of a test battery and full-scale laboratory training and testing system termed "DRILCOM," after the four major elements of auditory situation awareness: Detection, Recognition/Identification, Localization, and pass-through COMmunications (Casali & Lee, 2016a). The DRILCOM system along with localization portion of the test battery demonstrated the ability to measure and train localization acquisition skills. It also proved sensitive enough to detect localization ability differences between listening conditions, namely between the open ear and with a HPD/TCAPS and between types of HPDs/TCAPS (Casali & Robinette, 2014; Casali & Lee, 2016a; Casali & Lee, 2016b). The DRILCOM system highlighted the detrimental impacts that TCAPS devices impart on auditory situation awareness. Focus group studies of military service members and interviews with senior military researchers in the U.S., Canada and Great Britain confirmed the need to train localization under TCAPS conditions but underscored the challenges of developing localization tests and equipment that could consistently train localization in non-laboratory environments (Abel, 2008; Bevis et al., 2014; Brungart, 2014). With these results in mind, the Office of Naval Research issued a research grant to design and evaluate a system capable of providing the requisite auditory localization acquisition skills necessary to improve auditory situation awareness. The Portable Auditory Localization Acclimation Training (PALAT) system was developed to fill the operational gap in auditory situation awareness training by providing the military with a highly portable, objective, measurement- and feedback-intensive, and user-operable system that instills and improves auditory localization skills in service members over a regimented series of learning units. The investigator hypothesized that a portable auditory localization acclimation training system was feasible and could be developed using primarily off-the-shelf components. Furthermore, it was hypothesized that this system would provide acoustically-accurate localization cues similar to the DRILCOM system to enable a military trainee to operate the system to train and improve auditory localization acquisition skills in a reasonable amount of time.

2.2 PALAT System Requirements

The PALAT system was intended to be assembled and operated by a military service member in a small room, test and train throughout 360-degrees of azimuthal auditory localization, and store in a large shipping trunk that could be hand-carried. More specifically, the portable system was required to setup in a space no more than 7-feet wide by 7-feet long by 7-feet high. The entire system, complete with small directional loudspeakers, structure, laptop/tablet user interface, and audio equipment had to breakdown and fit in a large storage case. An optimal number of loudspeakers were required to be attached in equidistance spacing around a portable structure. Furthermore, the loudspeakers needed to be height-adjustable to align with the trainee’s ear height when seated in the middle of the array. Where possible, the system should use commercially off-the-shelf products; this was desired to alleviate a long design phase and control costs. The overriding goal of the PALAT system was to provide a user-friendly device that allowed trainees to improve auditory localization skills with the open ear and while wearing a TCAPS device. Given the demonstrated success of the DRILCOM system, the PALAT system was required to incorporate the improved DRILCOM training protocol developed in Phase I and provide a similar learning benefit over the same number of learning units. The following table lists specified and implied tasks of the PALAT system.

Table 5. PALAT system's specified and implied requirements.

Structure Requirements
Shall be no larger than 7 feet wide by 7 feet long by 7 feet high.
Shall be able to assemble by 1 or 2 trainees.
Shall breakdown and store in a large shipping trunk that can be hand-carried.
Shall form a "halo" configuration.
Shall support a ring of loudspeakers.
Shall attach loudspeakers in equidistance spacing measured angularly from the center position.
Shall attach loudspeakers an equidistance radially from the center.
Shall position the centerline of the loudspeakers 4 feet above the floor.
Shall be height adjustable to align with the trainee seated ear height.

Loudspeaker requirements
Shall be attached to a portable frame.
Shall be controlled by a laptop or tablet computer via LabVIEW™ software.
Shall receive power from a small amplifier or contain an internal amplifier.
Shall be housed in a durable enclosure.
Shall be highly directional.
Shall have a flat frequency response from 250 Hz – 10000 Hz (as close as possible to ±3 dB).
Shall have a frequency range of 250 Hz – 10000 Hz.
Shall be able to produce sound pressure levels of 85 dBA at 1 meter.
Shall be housed in a full enclosure containing no air portholes.
Shall contain single or coaxial drivers.
Shall be capable of reproducing the DRILCOM dissonant tonal complex signal.

System requirements
Shall incorporate a laptop or tablet computer user interface.
Shall provide a masking noise during training and testing.
Shall be calibrated by the trainee.
Shall import the training protocol developed in Phase I.
Shall impart similar learning benefits as DRILCOM system over the same number of learning units.

2.3 PALAT Structure

The investigator conducted a thorough literature review and evaluated 18 unique localization apparatus designs used in published studies to discover design features that could be used to develop a portable system (Appendix A). The extensive literature search and review yielded no results for a portable auditory situation awareness or auditory localization device. The majority of the localization testing apparatus designs were located in facilities equipped with sound-absorptive materials on interior surfaces to reduce reverberation. In addition, the majority of the test apparatus were too large to fit in the required space. Common horizontal localization apparatus designs included circular (halo), semi-circular, spherical (elevation and horizontal), and sound booms. Of the five circular designs, the number of loudspeakers used varied from 8 to 36 loudspeakers resulting in respective azimuthal angular separation between loudspeakers, measured from the center point, varied from 45-degrees to 10-degrees. Without a defined standard, the investigator first determined the optimal number of loudspeakers for the PALAT system.

Optimal number of Loudspeakers

The azimuthal angular separation between loudspeakers was one of a few design parameters that was not constrained by system requirements. The size and portability requirements of the PALAT system dictated the maximum distance of the loudspeaker from the trainee and the general size of the loudspeaker due to weight and size constraints. There were no known studies that attempted to identify the optimal azimuthal separation angle for testing or training auditory localization. The decision to use 24 azimuthal loudspeakers resulting in a 15° azimuthal angular separation was; 1) based on human auditory capabilities, 2) to allow for evaluation and validation of the PALAT system against the proven full-scale DRILCOM system, 3) to allow for future testing at increased azimuthal angular accuracy, and 4) because of the impacts of directing visual field of view based on auditory localization.

As covered in the literature review section, the ability to localize sounds in the horizontal plane is best when the sound is presented directly in front of the listener (Blauert, 1997).

Localization errors increase as the sound source is moved away from the median plane in either direction. Localization accuracy is also dependent upon the duration and spectral, or frequency, content of the signal. Oldfield and Parker (1984) tested azimuthal localization accuracy using white noise presented from loudspeakers with 10° angular separation and found that absolute error was approximately 4° to 6° from directly in front of the listener (0° coplanar with the median plane) to the frontal plane perpendicular to the ears (90° and 270°) (Figure 12). Localization accuracy decreased as the sound source moved behind the listener and was worst at approximately 20° at a horizontal azimuth of 160° or 200° (Oldfield & Parker, 1984). These findings suggested the need for an azimuthal angular separation less than 20° to accurately train and assess localization in front of the trainee but verified that an azimuthal angular separation too much below 20° may not be worth the benefit cost tradeoff of increased accuracy behind the trainee and extra weight and complexity of the portable system.

Another major factor that contributed to the number of PALAT loudspeakers was the requirement to validate the system against the proven full-scale laboratory DRILCOM system. The DRILCOM system consistently demonstrated the capability to test and train localization in open ear conditions and while wearing TCAPS using a 30° azimuthal angular separation. Using the same 30° azimuthal separation allowed for direct comparison of results and reduced confounding results based on angular separation differences. Previous studies on the DRILCOM system used 12 loudspeakers to train localization accuracy but provided 24 response locations allowing results to be measured for ballpark accuracy at ±15°. The ballpark measurements proved to provide useful in measuring localization accuracy while wearing TCAPS devices. As a result, the decision was made to incorporate 24 loudspeakers in the PALAT system, resulting in an azimuthal angular separation of 15°, but to train and test using only 12 loudspeakers during

Phase II and Phase III. Participants were still presented with 24 response locations during Phase II and Phase III to allow for ballpark accuracy measurement and analysis. This allowed for direct comparison with the DRILCOM system but provided the ability for more localization accuracy precision in future studies.

The last azimuthal separation factor involved the associated point of origin, or on-the-ground, separation distance at effective range of military relevant signals. One of the primary purposes of auditory localization is to orient the listener to the direction of the sound and cue the visual modality effectively reducing the response time in target identification (Wickens, Hollands, Banbury, & Parasuraman, 2013). The wider the azimuthal separation angle at the listener the broader the visual field of view search as the distance increases from the listener. This becomes a problem in military operating environments where relevant military threats originate from great distances, and reduction of visual search time is at a premium. Table 6 shows a comparison of the resulting visual field of search distances associated with 30° and 45° azimuthal separation for military threats originating from their effective range distances.

Table 6. Visual field of view search distances associated with 30° and 45° azimuthal separation for military threats originating from their effective range distances (USMC, 2017).

Military Threat	Effective Range (m)	Field of view search distance at effective range (m)		
		30°	45°	Difference (Δ)
AK-47 gunshot	300	155	229	74
Rocket Propelled Grenade (RPG)	500	258	383	125
AK-74 Sniper rifle	800	414	612	198
PKM machine gun	1000	517	756	239
82mm mortar launch	3000	1553	2296	743
107mm rocket launch	> 5000	2588	3826	1238

In military operations, time is of the essence. Military service members often have seconds or fractions of a second to identify and locate a threat after enemy contact is initiated. The on-the-ground differences or the additional amount of visual search area shown in the difference column of Table 6 could impact survivability. An increased horizontal scanning area of 125 meters when scanning the rooftops of a village in an urban environment looking for the origin of an RPG equates to a significant number of houses and rooftops. Likewise, an increased horizontal scanning area of 743 meters on the side of a mountain range for an enemy mortar team could drastically increase the time required to locate the enemy target. The impacts of using 15° or 30° instead of 45° azimuthal separation increases the number of loudspeaker positions and with it potentially the training and testing time to present repeated signals from each loudspeaker. The training time impacts are unknown to achieve similar localization accuracy. However, the potential extra training time was deemed worthwhile, given the need to validate the PALAT system with the DRILCOM system and the resulting impacts on additional visual scanning area.

Size of the PALAT system

PALAT system requirements limited the size to a 7-foot diameter in order maximize portability and flexibility. The decision was made to maximize the size allowed for due to potential issues associated with near field effects. The final PALAT system design consisted of a 2-meter diameter placing the loudspeakers approximately one meter from the trainee's ear. From an acoustical perspective, a one-meter distance from the loudspeaker places the trainee right at the edge of where the near field transitions into the far field for low frequencies. The acoustical near field is the region closest to the sound source where the sound pressure and the velocity of the wave particles are not in phase, meaning there is no simple relationship between the sound pressure and sound intensity (Hansen, 2001). Within the near field, the inverse square law does not hold up and the 6 dB decrease in SPL for every doubling of distance traveled does not occur. SPL measurements within the near field fluctuate making it difficult to obtain accurate measurements (Driscoll & Royster, 2003). The far field begins where individual sound waves combine to form uniform propagating waves and SPL measurements become predictable using the inverse square law (Driscoll & Royster, 2003). The distance of the near field depends on the frequency and characteristics of the sound source dimensions. It is hard to define an exact equation for the size of the near field but is described in general terms by one wavelength or three times the largest dimension of the sound source, whichever is largest (Hansen, 2001). The PALAT system loudspeakers were initially constrained to be between 2 to 4 inches in diameter to limit weight and maximize portability. As a result, the largest measurement to estimate the near field was the wavelength measurement. Equation 1.0 was used to calculate the length of the wavelength

$$\lambda = \frac{c}{f} \tag{1}$$

where λ is the wavelength in meters, c is the speed of sound in meters per second (343 m/s), and f is the frequency in Hz. At a 1-meter distance, the trainee could be in the near field for low frequencies around 343 Hz and below. The dissonant tonal complex includes two frequencies that may be in the near field, 104 Hz and 295 Hz. The PALAT loudspeakers are not be able to produce the same sound pressure level as the DRILCOM loudspeakers at the 104 Hz frequency due to the smaller size of the driver. The PALAT loudspeakers chosen were tested to verify the frequency response of the dissonant tonal complex compared to the DRILCOM system. The PALAT loudspeakers were able to produce sound pressure levels of the 295 Hz frequency tone within 3 dBA of the same sound pressure levels of the 737 Hz tone. These two tones would thus provide the necessary interaural timing difference cues for localization. Localization results would be evaluated to identify any differences between the two systems as a result of the potential near-field effects for the 104 Hz and 295 Hz frequencies.

Few studies have tested near field effects on auditory localization due partly to the challenges associated with defining the distance of the near field region. Some of the earliest studies use mathematical models of sound waves and assumed the head to be a perfect sphere. Duda and Martens (1998) measured the head-related transfer function (HRTF) using a 10.9 cm radius bowling ball from multiple near field distances and compared the results with previous theoretical mathematical model calculations. The primary findings were that both the theoretical and experimental data confirmed that variations of low frequency interaural level differences (ILDs) occur at close distances in the near field within five times the size of the radius of the sphere, or 0.5 meter (Duda & Martens, 1998). At distances greater than 0.5 meter, low frequencies are able to bend around the surface of the sphere resulting in no distinguishable difference in sound pressure level. In a follow-on study, Brungart and Rabinowitz (1999) measured the HRTF of frequencies from 200 Hz to 15000 Hz using a Knowles Electronic Manikin for Acoustic Research (KEMAR) for sound signals located within the near field from 0.12 meter to 1 meter to identify the effects of ILDs and ITDs. The near field had little effect on ITDs. However, ILDs increased by up to 20 – 30 dB as distance moved from 1 meter to 0.12 meter. In addition, the study found that ILDs occurred for low frequencies at the closest distances of 0.12 meter and 0.25 meter. The study showed that there was little variation between the 0.5 meter and 1 meter measurements for ILDs and ITDs and that measurements at the 1 meter distance were equal to the spherical head model predictions in previous studies (Duda & Martens, 1998; Brungart & Rabinowitz, 1999). These two studies suggest that sound signals presented from the PALAT system loudspeakers located approximately 1 meter from the listener will provide binaural localization cues consistent with sounds presented from the far field.

Frame of the PALAT system

Several design options were evaluated for the PALAT system frame including a series of individual poles with multiple loudspeaker mounts, a circular base with mounting posts for vertical loudspeaker poles, and an expandable accordion style frame (Figure 32). Out of all of the collapsible frame designs, the expandable accordion style frame was determined to be the most efficient and easiest to consistently assemble and disassemble by a trainee. The individual support poles (tripods) increased the risk of the trainee not properly aligning the poles and loudspeakers resulting in a non-equidistant spacing of the horizontal array. The circular base mount (similar to an upside-down trampoline frame) resolved this issue but resulted in numerous pieces that had to be assembled and disassembled by the trainee. The accordion frame solved both of these issues by expanding at a constant radial rate maintaining equidistant loudspeaker separation and spacing and reducing the number of assembly-required parts.

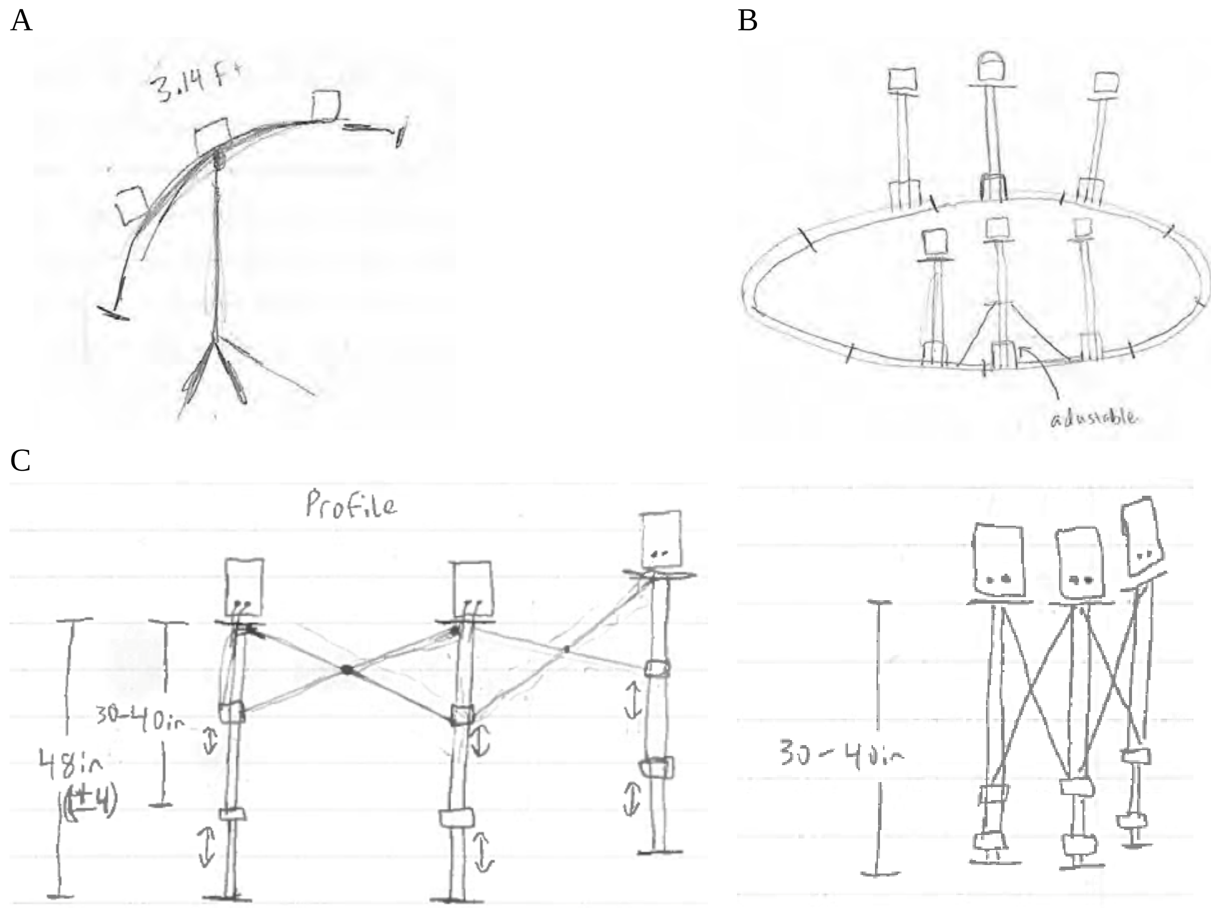


Figure 32. Initial PALAT system design concepts: A) Individual support poles with multiple loudspeaker mounts, B) Circular array base with vertical loudspeaker poles, and C) Expandable accordion style frame.

The investigator researched commercially off-the-shelf products that contained expandable accordion style frames as well as viable materials for the frame structure. Based on research and a few prototype designs, a pop-up canopy tent frame was selected as the building blocks for the PALAT structure. The lightweight steel tubes of the canopy tent frame offered the strength needed to support the loudspeakers and the rigidity to maintain the proper loudspeaker spacing during expansion. Three small 4-foot by 6-foot canopy tents were purchased and disassembled to provide 12 upright supports to hold the loudspeaker mounts and 12 scissor joints to allow for expansion and contraction of the system. The investigator decided to flip the accordion scissor joints to the bottom of the structure to allow the trainee to step over the bars as opposed to having to crawl or duck under the supports (Figure 33). In order to support the 24 azimuthal loudspeakers and 10 elevation loudspeakers, every other support pole starting at the 12 o'clock position was fitted with an aluminum mount that held three loudspeakers at the proper 15° azimuthal angular separation. The 11 o'clock and 1 o'clock support poles were custom fitted each with two interlocking poles that secured 5 loudspeakers spaced at 15° elevation angular separation for a total of 10 loudspeakers that could be used for elevation localization testing. Two of the loudspeakers, one at 11 o'clock and one at 1 o'clock, were used for both azimuthal and elevation testing. The remaining support poles each housed a single loudspeaker mount for a total of 32 loudspeakers, 24 azimuthal loudspeakers and 8 additional elevation loudspeakers. The investigator calculated the measurements for the frame and worked with Randy Waldron, Laboratory Instrument Maker in the Grado Department of Industrial and Systems Engineering at Virginia Tech, to fabricate the aluminum joints and loudspeaker mounts. Figure 33 displays the initial design schematics for the height adjustable support poles, scissor joints, and custom fabricated aluminum joints to create the 12-sided, dodecagon, accordion frame.

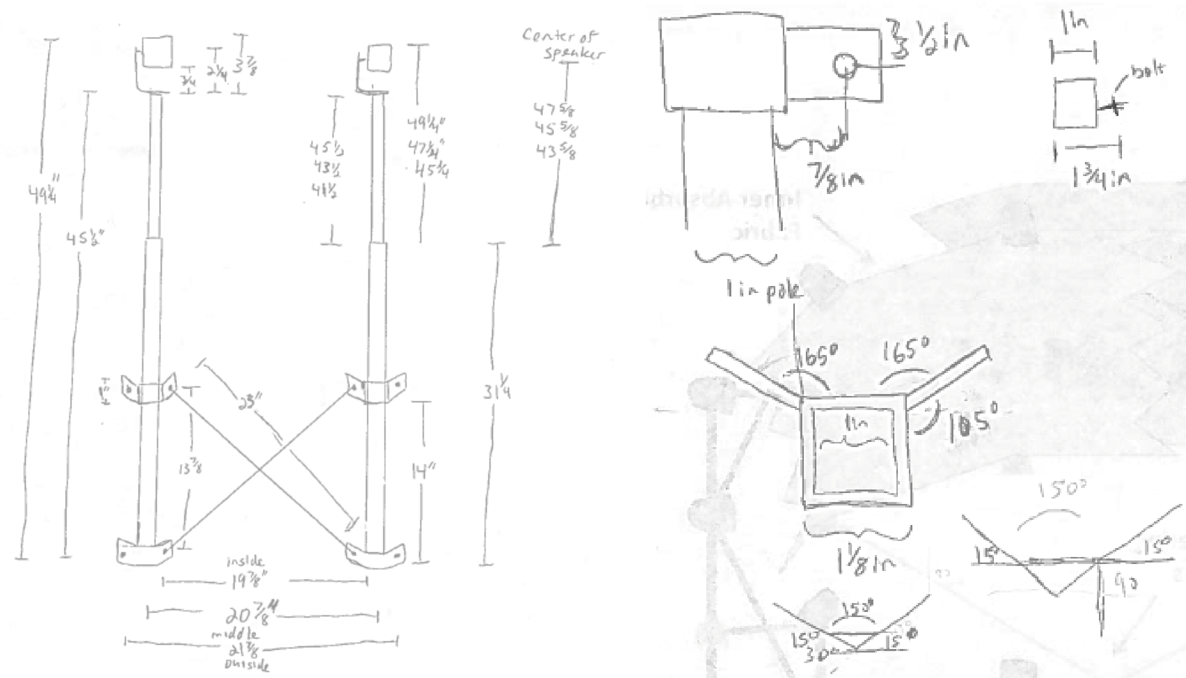


Figure 33. Initial PALAT system frame measurements and custom aluminum joint measurements.

The fully expanded PALAT system frame along with vertical elevation testing loudspeaker poles is shown in Figure 34. The collapsed PALAT system frame along with each type of loudspeaker mount is shown in Figure 35. When fully expanded, the 2-meter diameter frame (6.5 feet) places the azimuthal loudspeakers approximately 10.5 inches apart preventing most participants from fitting between the loudspeakers for access or egress. The DRILCOM system requires the participant to crawl under the halo ring frame. The investigator wanted to avoid this entry technique on the PALAT system since the scissor joints at the base of the system reduced the entry space under the loudspeakers. To provide an easier entrance method, the investigator designed a spring-loaded swivel mount for the 6 o'clock support pole. The entry/exit swivel mount held three loudspeakers and allowed the participant to rotate the loudspeaker mount in either direction opening up the full 21-inch entry/exit window into the system (Figure 36). The spring-loaded mount ensured the loudspeakers returned to the proper alignment. All of the 12 upright support poles contained a smaller center square tube that could be adjusted from a storage height of approximately 36 inches (with loudspeakers on the single mounts) up to three loudspeaker heights set at 43.5, 45.5, or 47.5 inches above the floor to align with the trainee's seated ear height.

Figure 34. Fully expanded PALAT system frame with 32 loudspeakers.

Figure 35. Collapsed PALAT system frame (left), three loudspeaker mount (top center), two loudspeaker elevation mount and single loudspeaker mount (bottom center), and elevation loudspeaker mount (right).

Figure 36. PALAT system entry/exit spring-loaded swivel gate loudspeaker mount.

2.4 PALAT Loudspeaker Evaluation

Loudspeakers were one of the most critical components of the PALAT system. The PALAT system was required to replicate similar auditory localization training results performed on the full-scale, laboratory grade DRILCOM system. This required the PALAT system loudspeakers to accurately reproduce the DRILCOM auditory signals across the audible frequency spectrum with a flat frequency response at sound pressure levels up to 85 dBA. The loudspeakers also needed to be durable and lightweight to meet military training requirements. As a result, the investigator evaluated over 20 small, commercially-available loudspeakers against 15 loudspeaker design parameters and engineering specifications (Appendix B). Table 7 displays the list of 15 loudspeaker design parameters and engineering specifications used to screen loudspeaker alternatives.

Table 7. Loudspeaker design parameters and engineering specifications.

Design Parameters	Engineering Specifications
Transducer (driver) properties	**Frequency Response**
Enclosure (cabinet)	Frequency Range
Dimensions	**Sensitivity**
Weight	Maximum SPL
Electrical impedance	Directional Response
Audio connectors	**Harmonic Distortion**
Mounting/Suspension types	Crossover Frequency
Power rating (Long-term & Maximum)	

The top three loudspeakers for the PALAT system were down-selected from the feasible alternatives and purchased to conduct in-depth testing and evaluation. The top three loudspeakers chosen were: Bose® FreeSpace® 3 Satellite Speakers, Cambridge Audio Minx Min 12, and Boston Acoustics® SoundWare XS. The Bose® FreeSpace® 3 Satellite Speakers, Cambridge Audio Minx Min 12, and Boston Acoustics® SoundWare XS were specifically chosen due to their high ratings across all design parameters but also because each loudspeaker exhibited a unique design element that may have offered benefits to the PALAT system (Figure 37). The

Bose® FreeSpace® 3 Satellite Speakers used a traditional 2.5 inch full-range single cone transducer, or driver (Bose, 2017). The Cambridge Audio Minx Min 12 used a 2.25 inch flat Balanced Mode Radiator (BMR) driver (Cambridge Audio, 2017). The Boston Acoustics® SoundWare XS used a 2-way, coaxial driver with a 2.5 inch woofer and 0.5 inch tweeter mounted directly in front of the woofer via a bridge mount (Boston Acoustics, 2017).

Figure 37. Loudspeaker alternative finalists: A) Bose® FreeSpace® 3 Satellite Speakers, B) Cambridge Audio Minx Min 12, and C) Boston Acoustics® SoundWare XS.

Upon receiving the three alternative loudspeakers, the investigator conducted acoustical testing on three performance measurements on all three loudspeakers in the VT-ASL facilities: 1) frequency response, 2) total harmonic distortion, and 3) ability to reproduce DRILCOM's localization signals used during Phase I of the overarching investigation. Data were also collected on three additional performance measures (sensitivity, power rating, and impedance), on two portability measures (weight and 3-axis physical dimensions), and on two durability/usability measures (loudspeaker driver type and wire terminal connector type). The resulting data for the 10 performance measurement criteria were then summarized into a SME questionnaire. The questionnaire first included a discussion of each of the performance metrics and the desirable objective performance standards needed for the PALAT system. The SMEs were asked to conduct a pairwise comparison between each of the 10 criteria to determine the criterion weighting coefficient, or priority of each criterion. Following the performance measure description and desired value for each criterion, the resulting data for each loudspeaker alternative were presented. To avoid bias from brand name or loudspeaker preference, the three loudspeaker alternatives were referenced throughout the questionnaire as **Alternative A**, **Alternative B**, and **Alternative C** respectively representing the **Bose® FreeSpace® 3 Satellite Speakers (A)**, **Cambridge Audio Minx Min 12 (B)**, and **Boston Acoustics® SoundWare XS (C)**. The data, graphs, and tables throughout the questionnaire were color-coded to match the alternatives. The SMEs were then asked to conduct a pairwise comparison between each of the alternatives to determine the alternative choice coefficient, or order of rank for each criterion. The following sections describe the performance measure testing conducted for three of the performance measure criteria, the desired PALAT loudspeaker performance for each criterion, and the resulting data measured or collected for the three loudspeaker alternatives for each criterion. The actual loudspeaker alternative names are displayed in the sections below, **FreeSpace 3** for **Bose® FreeSpace® 3 Satellite Speakers**, **Minx Min 12** for **Cambridge Audio Minx Min 12**, and **SoundWare XS** for **Boston Acoustics® SoundWare XS**, along with the color-coded data, graphs, and tables. The SME questionnaire with Alternative A, B, and C can be found in Appendix C.

Frequency Response

The frequency response is the range of frequencies over which a loudspeaker produces a sound pressure level that remains within a specific ±dB tolerance level of its nominal sensitivity level. Typically, the tolerance level is set at ±3 dB for mid- to high-frequencies and ±6 dB for low frequencies depending on the size and quality of the loudspeakers (Emanuel, Maroonroge, & Letowski, 2009). Frequency response is an output measure based on a constant level input of pure tone frequencies (Borwick, 2001). The measurement is given by a stated frequency range in Hz within a dB SPL tolerance range, e.g. 200 Hz – 16000 Hz (±3 dB). Frequency response is often measured on-axis, i.e., directly in front of the loudspeaker, at a distance of 1 meter (m) with 1 Watt (W) of power. A flat frequency response over a broad frequency spectrum means the loudspeaker is capable of reproducing the input sound accurately.

The PALAT system needed to reproduce military relevant sounds that spanned the frequency spectrum including the low frequency sounds of explosions, mid frequency sounds of small-caliber gunshots, and high frequency sounds of the whistle of incoming rocket propelled grenades or mortars, and the clicks emitted by charging of a rifle (Clasing & Casali, 2014). In addition to presenting military relevant sound across the audible spectrum, to train localization, the PALAT system *must have been able* to accurately produce sounds that provided interaural time difference cues below 1500 Hz and interaural level difference cues above 3000 Hz (Casali & Tufts, in press). *Based on the requirements above, the PALAT system was required to have a flat frequency response of 250 Hz – 10000 Hz within ±3 dB.*

Table 8 below displays the manufacturer-reported frequency response range within ±3 dB at 1 watt measured at 1 meter. Although 1 watt at 1 meter and a ±3 dB tolerance is the standard, Minx Min 12 manufacturer did not specify ±3 dB and SoundWare XS did not specify 1 watt at 1 meter (Bose, 2017; Cambridge Audio, 2017; Boston Acoustics, 2017).

Table 8. Manufacturer-reported frequency response.

	Freespace 3	**Minx Min 12**	**SoundWare XS**
Frequency response	**210 Hz – 16 kHz**	**120 Hz – 20 kHz**	**150 Hz – 20 kHz**

An on-axis frequency response test was conducted by the investigator at the Virginia Tech – Auditory Systems Laboratory (VT-ASL) using a manual-stepped pure tone sinusoidal signal from 100 Hz to 20000 Hz. The test was conducted in the VT-ASL anechoic chamber (200

Hz low-frequency cutoff) with a 1-inch Larson-Davis Model LD2575 measurement microphone (SN: 1280) and Larson-Davis 900B Preamp (SN: 2394) placed 1 meter from the cone of the loudspeaker. Measurements were recorded using a Larson-Davis Model 2900 spectrum analyzer (SN: A0280). The microphone was calibrated at 94.0 dBA (1000 Hz tone) using a Quest QC-20 Calibrator (SN: QOA070051). The audio signal was generated using Audacity® 2.2.0 and presented via a MacBook Pro laptop with a Kemo® Electronic 12W audio amplifier. The output voltage was manually measured and set to produce 1 W. Of note, Minx Min 12 and SoundWare XS were 8 Ohm loudspeakers and the output voltage was set to ~2.83 Volts root mean square (Vrms) at 1000 Hz. Freespace 3 was a 6 Ohm loudspeaker and the output voltage was set to ~2.45 Vrms at 1000 Hz. The volume and output voltage were not adjusted during the testing in order to try and maintain a constant voltage as specified by industry standards (AES 2-2012, 2012). Some deviations in frequency response may be attributable to frequency response limitations of the computer soundcard or amplifier. However, the tests were consistent across all alternatives. Figure 38 displays the measured sound pressure level (dB SPL) in every 1/3-octave frequency band from 100 Hz – 20000 Hz.

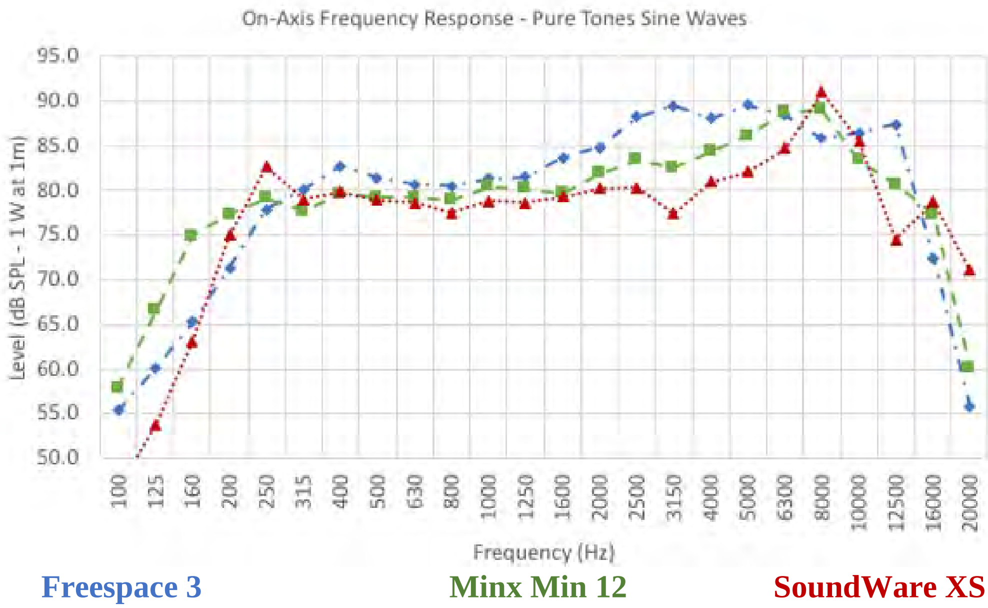


Figure 38. On-Axis frequency response – stepped sine (dB SPL 1 watt at 1 meter).

An additional frequency response test was conducted under similar conditions above using the Room EQ Wizard® computer software to generate a sinusoidal sweep and record the frequency response at each frequency (rather than in 1/3-octave bands as previously discussed). A MiniDSP UMIK-1 USB measurement microphone was used as the input source. Figure 39 displays the results of the computer-generated frequency response test.

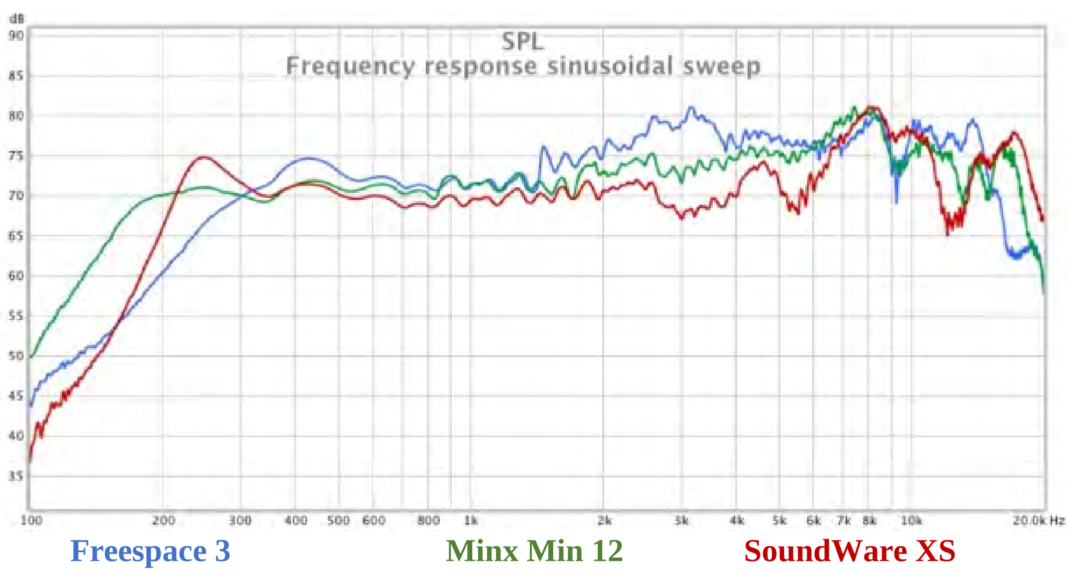


Figure 39. On-Axis frequency response – sine sweep (dB SPL 1 watt at 1 meter).

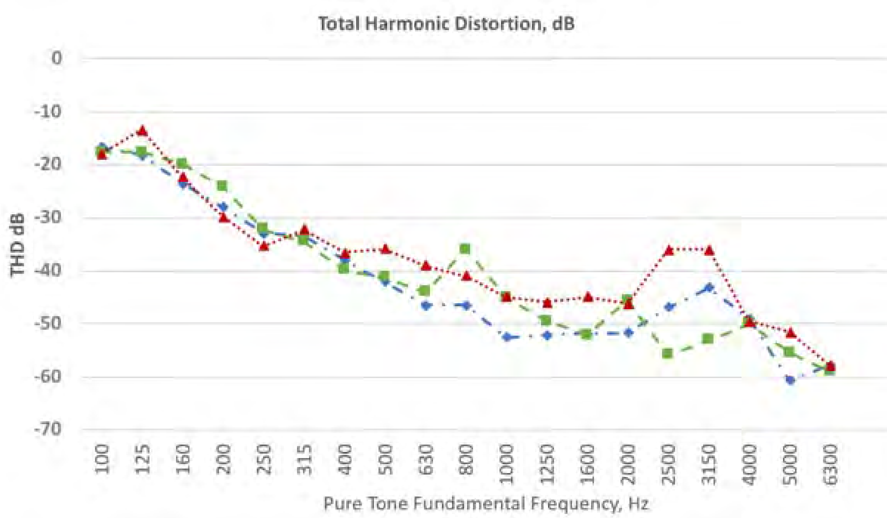
<u>Total Harmonic Distortion</u>

Total harmonic distortion (THD) is the amount of amplitude distortion present in a signal as a result of mechanical or magnetic nonlinearities in the loudspeaker or impurities in the voltages and currents in the power system (Eargle, 2003; Skvarenina, 2002). The distortions occur at integral multiples (i.e., harmonics) of the fundamental frequency of the signal. Calculated relative levels of harmonics compared to the fundamental can be expressed in percentages or decibels (dB) as below (Newell & Holland, 2007):

0 dB	100 %
-10 dB	30 %
-20 dB	10 %
-30 dB	3 %
-40 dB	1 %

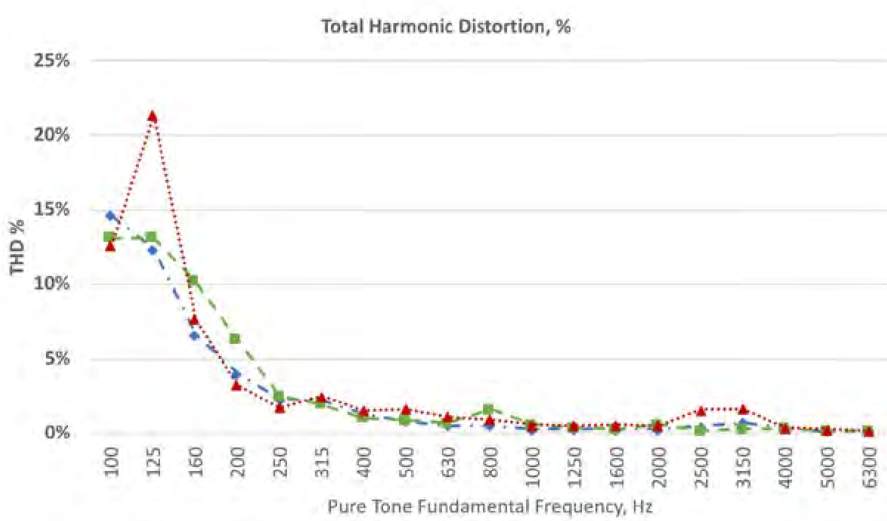
In the PALAT system, the performance effect on localization based on THD was expected to vary depending on the frequency spectrum and the sound pressure level of the signal presented. *<u>As a general rule, Total Harmonic Distortion (THD) should be minimized.</u>*

Harmonic distortion was measured with both a stepped sine and sine sweep measuring the second- and third-harmonic components compared with the fundamental frequency. The first harmonic distortion test was conducted by the investigator at the VT–ASL using a manual-stepped pure tone sinusoidal signal from 100 Hz to 20000 Hz. This allowed for measurements of the 6th-harmonic up to 3150 Hz and 3rd-harmonic at 6300 Hz. The test was conducted in the VT-ASL anechoic chamber using the same measurement set-up as used above in the frequency response test (1 W at 1 m). Again, some of the distortion may be attributable to the computer soundcard or amplifier, but measurements were consistent across all alternatives. Figure 40 displays the total harmonic distortion in dB below the fundamental frequency. Figure 41 displays the total harmonic distortion as a percentage referenced to the fundamental frequency.



Freespace 3 **Minx Min 12** **SoundWare XS**

Figure 40. Total harmonic distortion reference to the fundamental frequency (dB).



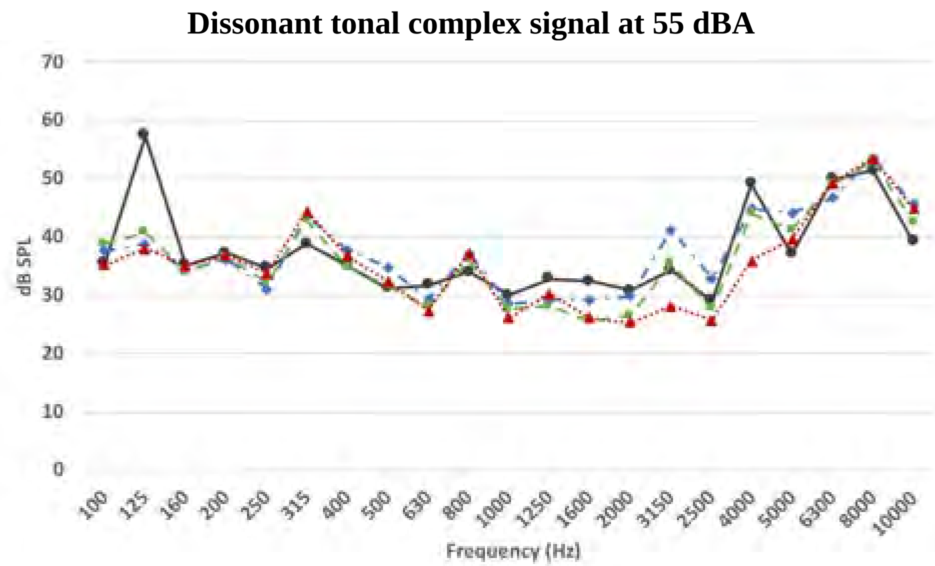
Freespace 3 **Minx Min 12** **SoundWare XS**

Figure 41. Total harmonic distortion reference to the fundamental frequency (%).

<u>*Ability to Reproduce DRILCOM Signals*</u>

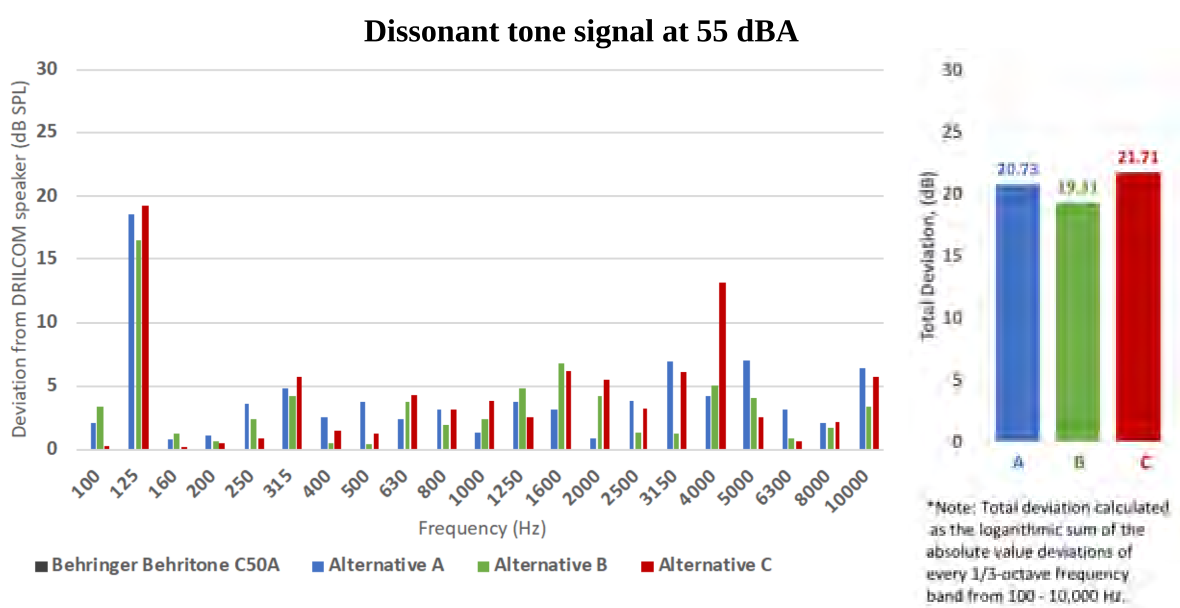
The PALAT system was required to reproduce the auditory signals presented on the full-scale, laboratory grade DRILCOM system consisting of 12 integrally-powered, 5.25-inch loudspeakers. Previous studies using the DRILCOM system demonstrated the ability to improve the open ear's absolute correct performance by over 25% and demonstrated that participants using certain TCAPS can learn and perform at similar ballpark levels to the open ear with relatively little training (Casali & Lee, 2016a; Casali & Robinette, 2014). *The PALAT system loudspeakers were required be able to reproduce the localization training signals with similar fidelity as produced by the DRILCOM loudspeakers in order to achieve comparable localization training effects.*

The investigator conducted a test in the hemi-anechoic DRILCOM laboratory room to measure the sound pressure level across the frequency spectrum from 100 Hz to 10000 Hz for a dissonant tone signal and four military relevant signals, a simulated whistle from an incoming artillery round (Whistle), the rotor sounds of an approaching Apache helicopter (Apache), spoken foreign language (Arabic), and an AK-47 three round burst (AK-47). Each loudspeaker was calibrated at 55 dBA and 80 dBA for the dissonant tonal complex signal. The DRILCOM loudspeaker, Behringer Behritone C50A, was measured at a distance of 1.5 meters from the measurement microphone and the three PALAT alternative loudspeakers were measured at a distance of 1 meter from the measurement microphone. The graphs below display the measured sound pressure level (dB SPL) for each 1/3-octave band frequency and the absolute deviation from the DRILCOM reference loudspeaker. The deviation graph on the far right plots the total absolute deviation, logarithmic sum across all frequencies, in order to show the total absolute delta. Figures 42-43 and 44-45 display the frequency response for the dissonant tonal complex signal at 55 dBA and 80 dBA respectively. Frequency response graphs for the Whistle, Apache, Arabic, and AK-47 can be found in Appendix C.



Behritone C50A **Freespace 3** **Minx Min 12** **SoundWare XS**

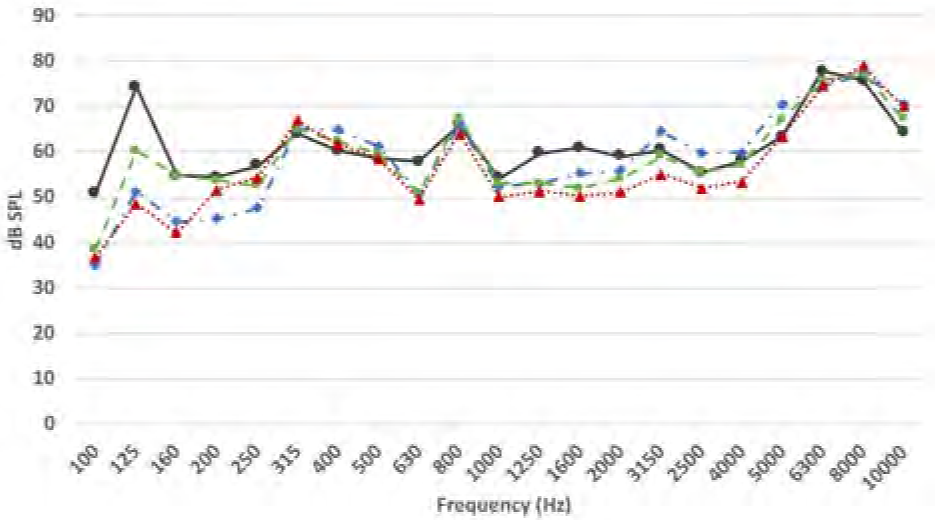
Figure 42. Sound pressure level of dissonant tone signal tone at 55 dBA.



Behritone C50A **Freespace 3** **Minx Min 12** **SoundWare XS**

Figure 43. Sound pressure level deviations from DRILCOM loudspeaker for dissonant tone signal tone at 55 dBA.

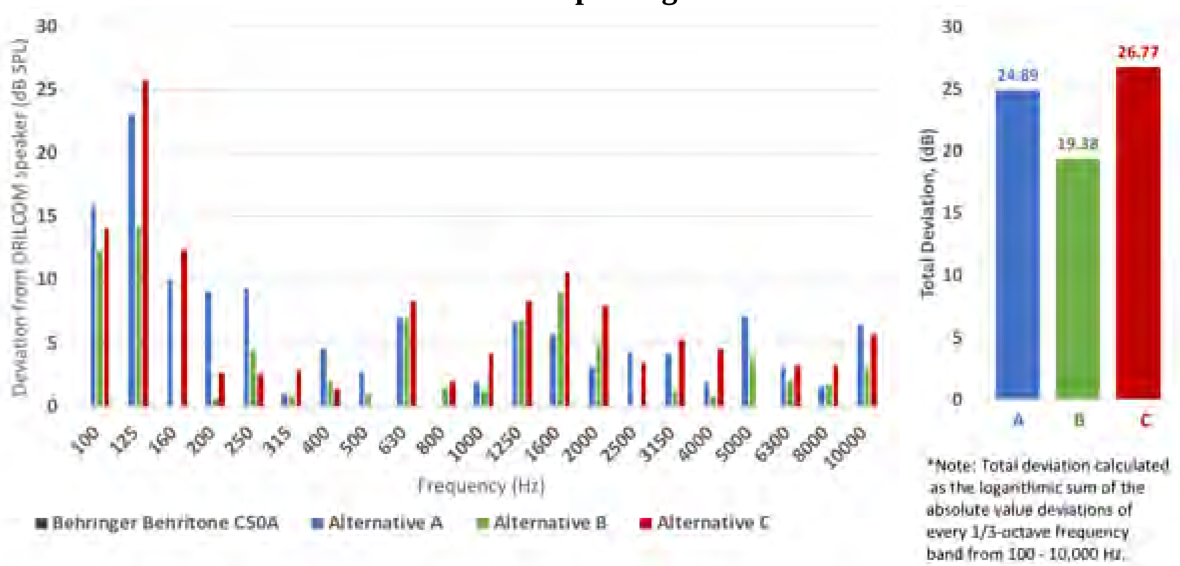
Dissonant tonal complex signal at 80 dBA



Behritone C50A **Freespace 3** **Minx Min 12** **SoundWare XS**

Figure 44. Sound pressure level of dissonant tone signal tone at 80 dBA.

Dissonant tonal complex signal at 80 dBA



Behritone C50A **Freespace 3** **Minx Min 12** **SoundWare XS**

Figure 45. Sound pressure level deviations from DRILCOM loudspeaker for dissonant tone signal tone at 80 dBA.

Performance Measure Criteria Data by Loudspeaker

The seven additional performance measure criteria data were collected from the manufacturers' product descriptions. Each performance measure and desired measurement for the PALAT system was discussed in the SME questionnaire prior to presenting the alternative data. Table 9 summarizes the three loudspeaker alternatives' data for the seven performance measurement criteria. Descriptions of each performance measurement and data can be found in Appendix C.

Table 9. Manufacturer-reported data by loudspeaker alternative.

Performance Measurement Criteria	Freespace 3	Minx Min 12	SoundWare XS
Sensitivity	**84 dB SPL**	**86 dB SPL**	**85 dB SPL**
Power rating	**12 W (48 W peak)**	**25 – 200 W**	**10 – 100 W**
Impedance	**6 Ohms**	**8 Ohms**	**8 Ohms**
Weight	**1.9 lbs**	**0.95 lbs**	**1 lb**
Dimensions	**3.0 x 3.0 x 4.0 inches**	**3.1 x 3.1 x 3.3 inches**	**3.7 x 4.3 x 4.5 inches**
Driver size(s)	**2.5 inch full-range**	**2.25 inch BMR**	**2.5 inch woofer 0.5 inch tweeter**
Driver type	**Full-range cone driver** **Typically consist of a wire voice coil inside a magnetic field attached to a funnel shaped diaphragm, or cone. Pistonic motion is used to create acoustical waves.**	**Balanced Mode Radiator (BMR)** **BMR is a flat loudspeaker that combines the pistonic motion of cone drivers with the vibration motion of flat panel loudspeakers.**	**2-way, coaxial driver** **The tweeter is mounted directly in front of the 2.5 inch woofer via a bridge mount.** **Crossover frequency of 5000 Hz.**
Wire terminal type	**Spring clip terminal**	**4-way Binding post**	**Spring clip terminal**
Compatible connectors	**Bare wire Pin connectors**	**Bare wire Pin connectors Spade connectors Banana plugs**	**Bare wire Pin connectors**

The 10 performance measurement criteria and resulting data for the three loudspeaker alternatives were then evaluated via a Subject Matter Expert (SME) pairwise comparison algorithm developed by Meister (1985) for ing component selection decisions for human-equipment systems.

Meister Analysis Methodology

Meister (1985) developed an algorithm using a series of pairwise comparisons of the evaluation criteria and alternatives to derive a weighted score for each alternative and identify the alternative that best meets the prioritized performance criteria. The first step in the Meister analysis was to conduct a pairwise comparison for every design criterion in order to determine the value or weight of each criterion (Meister, 1985). For the PALAT system loudspeaker analysis, each SME was asked to compare every pair of performance measurement criteria and select the criterion that they believed was most important for the system. This resulted in 45 pairwise comparison between all 10 criteria. The selected criterion from each pairwise comparison was assigned a value of 1 and a value of 0 was assigned to the criterion that was less important (Meister, 1985). The total criterion value, sum of winning pairwise comparisons, for each performance measure was then divided by 45, representing the total number of comparisons, to find a *criterion weighting coefficient.*

The next step in the Meister analysis was to complete a pairwise comparison of each alternative for every design criterion based on objective data presented and the SME's professional experience. This resulted in a total of three pairwise comparisons for of the 10 criteria. The selected alternative from each pairwise comparison was assigned a value of 1 and a value of 0 was assigned to the alternative that was outperformed (Meister, 1985). The total alternative choice tally, sum of winning pairwise comparisons, for each criterion was then divided by 3, representing the total number of comparisons, to find an *alternative choice coefficient.* The final step in the Meister plan analysis was to multiply the criterion weighting coefficient by the alternative choice coefficient for each criterion and calculate the sum of the products resulting in an *alternative score for each alternative* (Meister, 1985). The alternative scores from each SME were then compared using non-parametric statistical analysis to determine optimal loudspeaker for the PALAT system.

Meister Analysis Results

The questionnaire was sent online to 16 hand-selected audio equipment, room acoustics, and hearing experts. Six SMEs from academia, industry, and military responded to the questionnaire evaluating the three loudspeaker alternatives using the Meister analysis pairwise comparisons. Table 10 highlights the professional experience and diversity of the Subject Matter Experts. The investigator used a fillable Adobe™ Acrobat™ pdf file which automatically imported the questionnaire results and exported a comma separate value (CSV) file. Microsoft® Excel was used to calculate weighted criterion coefficients and alternative choice coefficients for each SME. SPSS® Statistics was used to conduct non-parametric statistical analysis.

Table 10. Subject Matter Expert Acoustics and Audiology Team Experience.

SME Professional Title	SME Degree
US Army Audiologist	Ph.D. in Human Factors Engineering
Acoustical Engineer	Ph.D. in Architectural Acoustics
Senior Research Psychologist (Auditory Research)	Ph.D. in Psychology
Media Engineer	MFA in Spatial Audio Composition and Technology
US Army Operations Research Systems Analysis Officer / Army Aviator	MS in Industrial and Systems Engineering
Acoustical Engineer	Ph.D. in Acoustics

The first step in the Meister analysis was to calculate the criterion weighting coefficient for each SME. The sum value of each performance measure criterion was calculated and divided by the total number of pairwise criteria comparisons (45 comparisons) resulting in the criterion weighting coefficient (Meister, 1985). Table 11 shows an example of an SMEs criteria pairwise comparison results and calculated weighting coefficient for each performance measure criterion. Figure 46 displays the mean criterion weighting coefficients for each performance measure. The error bars representing the 95% confidence interval about the means show a large variation in the prioritizing of performance measures. This indicates that the SMEs valued certain different performance measures based on their individual experiences and backgrounds in acoustics and auditory localization.

Table 11. Example SME criteria pairwise comparison and weighting coefficient score.

		Criteria									Choice Tally									Total	Weighting Coefficient
Performance Requirements	1	Frequency Response	1	1	0	1	1	0	0	1	0									5	0.111
	2	Total Harmonic Distortion	0		0	1	1	1	0	0	1	1								5	0.111
	3	Abiity to reproduce DRILCOM signal		0	1		0	0	0	0	0	1	1							3	0.067
	4	Sensitivity			1	0	1		0	1	1	1	1	1						7	0.156
Power Requirements	5	Recommended Power Rating				0	0	1	1		1	0	0	1	1					5	0.111
	6	Impedance					0	0	1	0	0		0	0	1	1				3	0.067
Portability	7	Weight						1	1	1	0	1	1		1	1	1			8	0.178
	8	Physical Dimensions							1	1	1	0	1	1	0		1	1		7	0.156
Durability	9	Driver type								0	0	0	0	0	0	0	0		0	0	0.000
Usability	10	Wire terminal type									1	0	0	0	0	0	0	0	1	2	0.044
																				45	

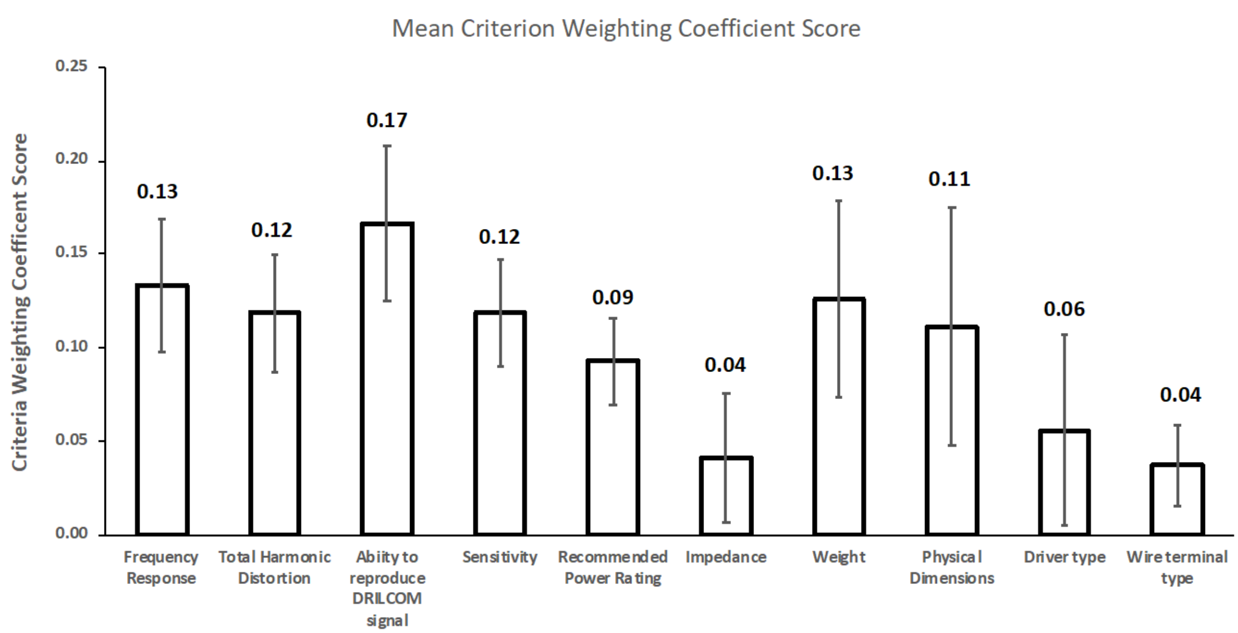


Figure 46. Mean criterion weighting coefficient for each performance measure criterion. Error bars are the 95% confidence interval about the mean. Numbers above the error bars are means.

The next step in the Meister analysis was to calculate the alternative choice coefficient for each alternative on every performance measure. The sum value of each alternative score for each performance measure was calculated and divided by the total number of pairwise criteria comparisons (3 comparisons) resulting in the alternative choice coefficient (Meister, 1985). Table 12 displays an example of a SME alternative pairwise comparison for each performance measure criteria and the resulting alternative choice coefficient score.

Table 12. Example of SME pairwise comparison of alternatives and alternative choice coefficient score.

Performance Measure Criteria	Alternatives	Choice Tally			Total	Choice Coefficient
	Freespace 3	0		1	1	0.33
Frequency Response	Min Minx 12	1	1		2	0.67
	SoundWare XS		0	0	0	0.00
	Freespace 3	1		1	2	0.67
Total Harmonic Distortion	Min Minx 12	0	1		1	0.33
	SoundWare XS		0	0	0	0.00
	Freespace 3	0		1	1	0.33
Abiity to reproduce DRILCOM signal	Min Minx 12	1	1		2	0.67
	SoundWare XS		0	0	0	0.00
	Freespace 3	1		1	2	0.67
Sensitivity	Min Minx 12	0	1		1	0.33
	SoundWare XS		0	0	0	0.00
	Freespace 3	0		0	0	0.00
Recommended Power Rating	Min Minx 12	1	1		2	0.67
	SoundWare XS		0	1	1	0.33
	Freespace 3	1		1	2	1.00
Impedance	Min Minx 12	0	0		0	0.00
	SoundWare XS		0	0	0	0.00
	Freespace 3	0		0	0	0.00
Weight	Min Minx 12	1	1		2	0.67
	SoundWare XS		0	1	1	0.33
	Freespace 3	0		1	1	0.33
Physical Dimensions	Min Minx 12	1	1		2	0.67
	SoundWare XS		0	0	0	0.00
	Freespace 3	1		1	2	0.67
Driver type	Min Minx 12	0	0		0	0.00
	SoundWare XS		1	0	1	0.33
	Freespace 3	0		0	0	0.00
Wire terminal type	Min Minx 12	1	1		2	0.67
	SoundWare XS		0	1	1	0.33

The final step in the Meister plan analysis was to multiply the criterion weighting coefficient by the alternative choice coefficient for each criterion and calculate the sum of the products resulting in an alternative score for each alternative (Meister, 1985). Table 13 displays an example of the alternative scores by criteria and the overall alternative score for the three loudspeaker alternatives.

Table 13. Example of SME alternative scores by performance measure and overall alternative score.

Performance Measure Criteria			Alternatives		
			Freespace 3	Minx Min 12	SoundWare XC
Performance Requirements	1	Frequency Response	0.037	0.074	0.000
	2	Total Harmonic Distortion	0.074	0.037	0.000
	3	Abiity to reproduce DRILCOM signal	0.022	0.044	0.000
	4	Sensitivity	0.104	0.052	0.000
Power Requirements	5	Recommended Power Rating	0.000	0.074	0.037
	6	Impedance	0.067	0.000	0.000
Portability	7	Weight	0.000	0.119	0.059
	8	Physical Dimensions	0.052	0.104	0.000
Durability	9	Driver type	0.000	0.000	0.000
Usability	10	Wire terminal type	0.000	0.030	0.015
		Total	**0.356**	**0.533**	**0.111**
			2nd Choice	1st Choice	3rd Choice

<u>The Meister analysis resulted in a unanimous selection of the Cambridge Audio Minx Min 12 loudspeaker.</u>

A Friedman test was applied to compare SME alternative scores. The non-parametric Friedman test allowed for comparisons between the ordinal rankings of each SME alternative scores. The Friedman test assigned ranks for each SME rating and then ranked the ratings for each alternative (Portney & Watkins, 2009). The null hypothesis was that there were no significant differences between the Bose® FreeSpace® 3 Satellite Speakers, Cambridge Audio Minx Min 12, and Boston Acoustics® SoundWare XS ratings. A significant finding indicated a difference was detected among one of the loudspeaker alternatives. A significant result of the Friedman's test was followed by three Wilcoxon signed-rank test pairwise comparisons between each alternative. The Friedman's test statistic was evaluated using a significance level of α=0.05. Post hoc pairwise comparisons employed a Bonferroni correction of α=0.05/3 = 0.017 to control for the increase risk of Type I errors due to multiple comparisons.

The Friedman non-parametric test for loudspeaker alternative rankings resulted in **a significant difference between mean rankings**, Bose® FreeSpace® 3 Satellite Speakers (*M*=0.30, *SD*=0.07), Cambridge Audio Minx Min 12 (*M*=0.55, *SD*=0.07), and Boston Acoustics® SoundWare XS (*M*=0.15, *SD*=0.4) across both systems ($\chi^2[2]=10.33$, $p=0.006$). Post hoc pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare each alternative. Pairwise comparisons used a Bonferroni adjusted α level of 0.017. The Wilcoxon test results (Table 14) showed **a significant difference in mean rankings of alternative score between Bose® FreeSpace® 3 Satellite Speakers versus Cambridge Audio Minx Min 12** ($Z=-1.17$, $p=0.043$)**, and the Boston Acoustics® SoundWare XS versus Cambridge Audio Minx Min 12** ($Z=1.83$, $p=0.001$). **No significant difference was found between Boston Acoustics® SoundWare XS versus Bose® FreeSpace® 3 Satellite Speakers** ($Z=0.67$, $p=0.248$). In addition, the Cambridge Audio Minx Min 12 was unanimously selected by all six Subject Matter Experts. Given the unanimous selection and significantly higher mean ranking score, the Cambridge Audio Minx Min 12 loudspeaker was chosen for the PALAT system. Figure 47 displays the mean rankings of alternative scores for each loudspeaker with 95% confidence intervals about the means.

Table 14. Wilcoxon results pairwise comparisons between rankings of each loudspeaker alternative on alternative score (bolded text in the table indicates a significant test result at $p<0.05$.)

Loudspeaker Alternatives	Z	*p*
Freespace 3 – Minx Min 12	**-1.17**	**0.043***
SoundWare XS – Minx Min 12	**1.83**	**0.001***
SoundWare XS – Freespace 3	0.67	0.248

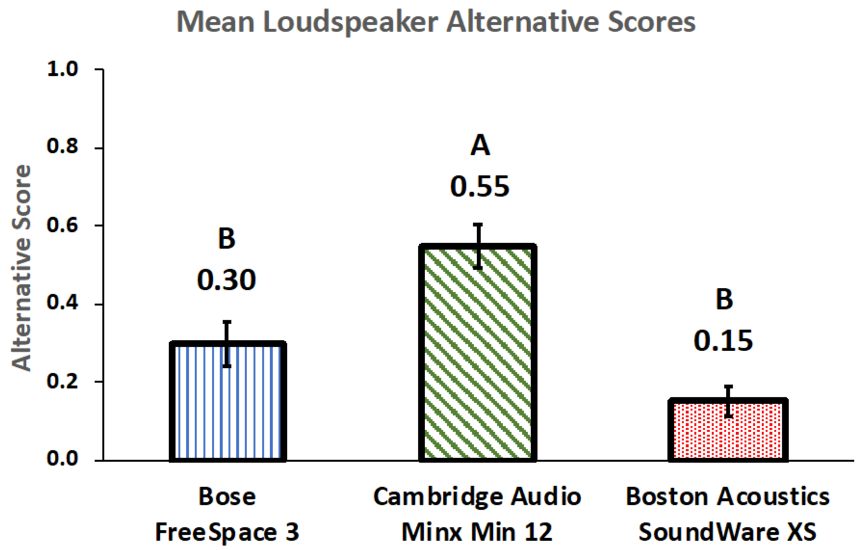


Figure 47. Mean alternative scores for each loudspeaker alternative. Means above 95% confidence interval error bars. Different letters indicate significant differences at $p<0.05$ with Bonferroni adjustment.

2.5 PALAT Auditory Equipment

The PALAT system requirements specified the need for a masking noise source and small amplifier to power the loudspeakers. The DRILCOM system used a compact disc player to play a compact disc pink noise track on loop. The loop function on the compact disc player caused the pink noise to pause and then restart every time the track ended. To resolve this issue in the PALAT system, the investigator researched small, stand-alone pink noise generators and selected the PNG-400 Pink Noise Generator made by Mystic Marvels LLC. The PNG-400 Pink Noise Generator provides continuous pink noise from 15 Hz to 20000 Hz with a pink noise flatness advertised at ±0.5 dB (Mystic Marvels LLC, 2018). The pink noise generator measured 4.2 inches long by 2.8 inches wide by 1.2 inches high and supplied pink noise via a 3.5mm stereo jack.

After testing three small amplifiers, the investigator selected the Stewart Audio AV30MX-2 two channel stereo mixer amplifier to power the PALAT system localization loudspeakers based on the compact size, weight, and low distortion properties. An additional Stewart Audio AV30MX-2 two channel stereo mixer amplifier was used to power the pink noise generator. The Stewart Audio amplifiers were capable of providing 30W amplification from two channels (Stewart Audio, 2018). The small amplifiers measured 4.35 inches long by 3.2 inches wide by 1.25 inches high and weighed less than a pound (Stewart Audio, 2018). Figure 48 displays the PNG-400 Pink Noise Generator attached to one of the Stewart Audio amplifiers as it was mounted in the audio equipment case.

Figure 48. PNG-400 Pink Noise Generator (top) mounted on top of a Stewart Audio AV30MX-2 amplifier (bottom).

A Numato Systems Pvt. Ltd. 32 Channel USB Relay Module was used to receive the LabVIEW™ software output and direct the audio signal from the tablet through one of the amplifiers and out to the proper azimuthal loudspeaker location. An Atlas IED® 15 Amp Half Width Rack Power Conditioner was used to power all of the audio components. The 15A Half

Width Rack Power Conditioner provided five AC outlets located on the back of the case to provide power to the two amplifiers, pink noise generator, and USB relay module and an additional AC outlet on the front of the case to provide power to the tablet computer.

The investigator researched multiple audio case options and selected the Gator Cases™ Laptop and 2-Space Rack Bag to secure and transport the audio equipment. The investigator designed a custom audio rack frame to mount all of the audio equipment inside of the two space audio rack portion of the Gator Cases™ bag leaving room for the tablet computer and wires to be stored when not in use. The audio wires from the relay module were connected into two DB-25 connectors mounted to face of the audio rack along with a USB/3.5mm stereo plug to connect the 32 loudspeakers and tablet computer. Two sets of wires with DB-25 connectors were designed and fabricated to connect the audio components to the 32 loudspeakers. A component list with associated costs can be found in Appendix D. Figure 49 displays the custom audio rack mount with audio components mounted (top left and right) and mounted in the carrying bag (bottom). Figure 50 shows the wiring diagram for the PALAT system. The last item placed in the audio equipment carrying bag was an additional Cambridge Audio Minx Min 12 loudspeaker to provide pink noise during training and testing. The investigator tested several options for the placement of the pink noise loudspeaker to determine the best location to reproduce the pink noise spectral content measured in the DRILCOM room.

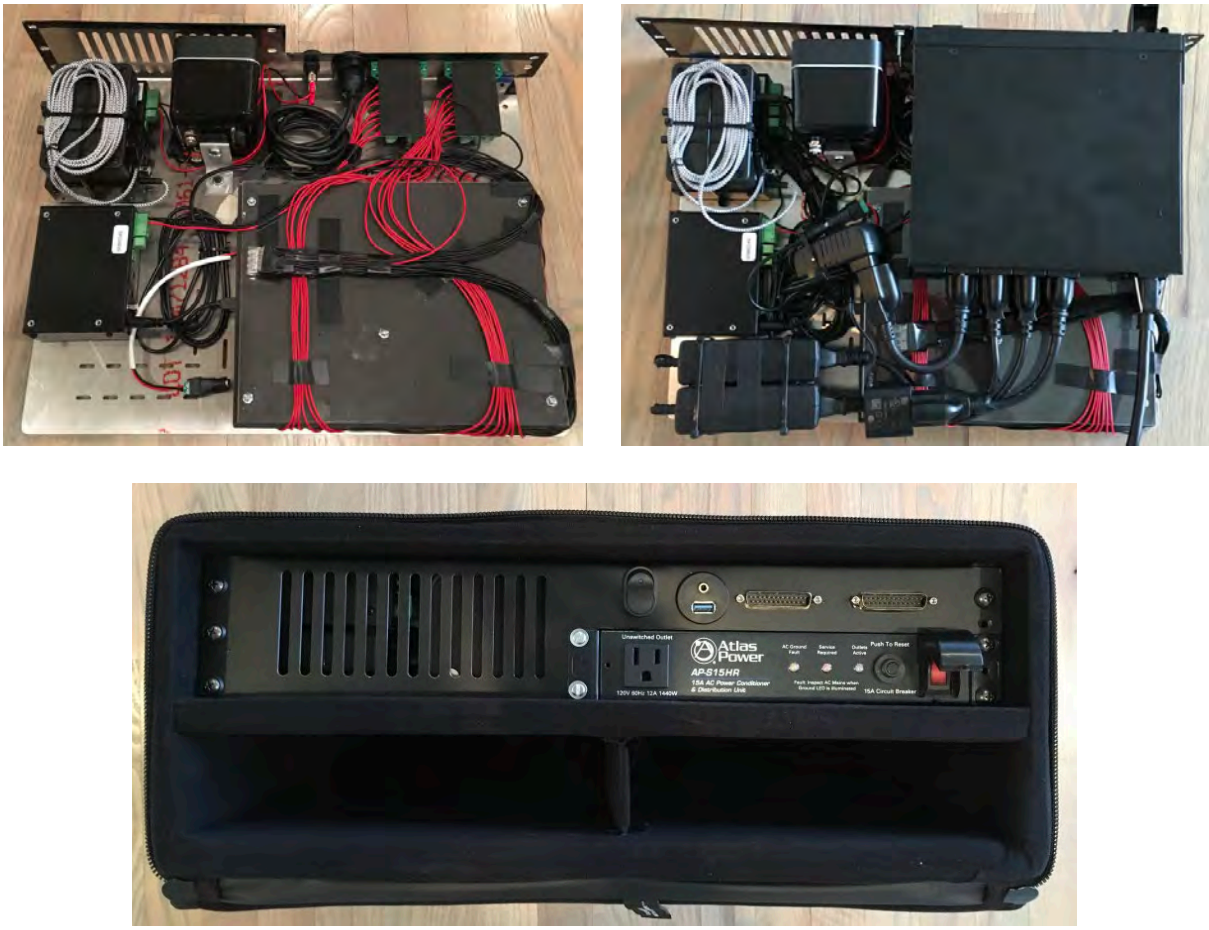


Figure 49. Audio components mounted to a custom audio rack (top left and right) and mounted in audio equipment bag (bottom).

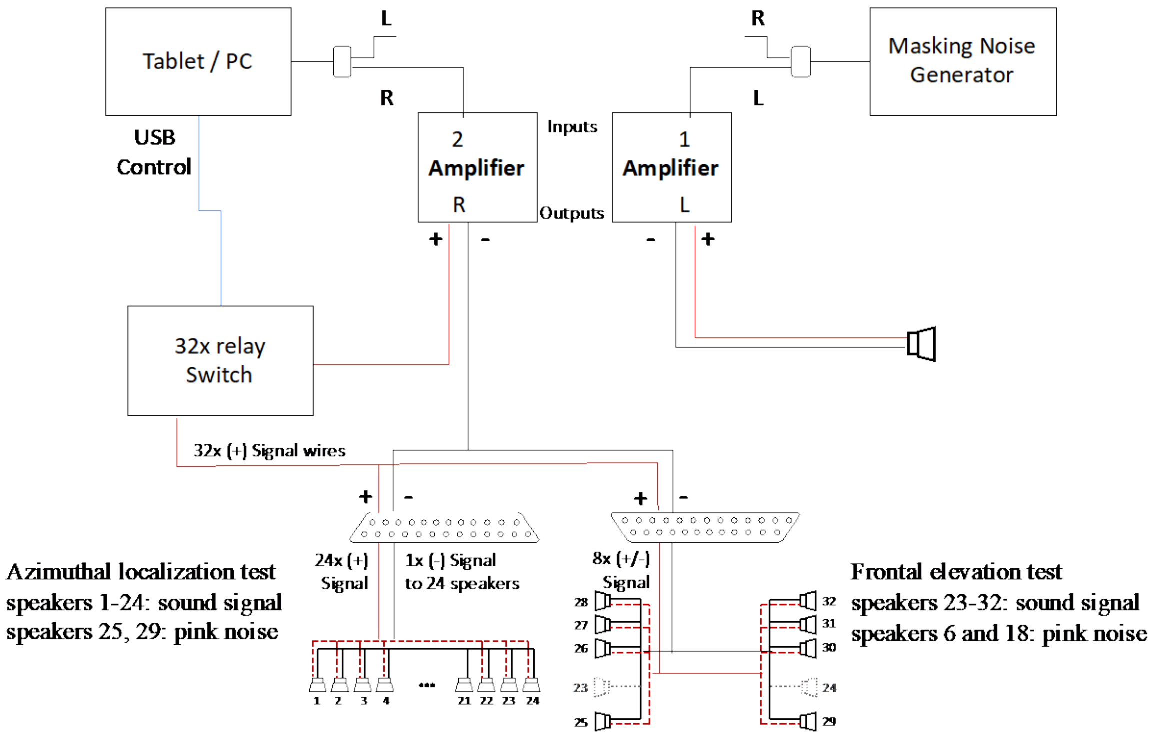


Figure 50. PALAT system wiring diagram.

A spectrum analysis was performed on the masking noise, pink noise, to be used in the PALAT system to determine the optimal placement for the masking noise loudspeaker(s). Four loudspeaker positions were considered but software limitations limited the courses of action to two options. The two loudspeaker location options tested consisted of: 1) using two of the loudspeakers used for elevation testing located 15-degrees below the azimuthal speakers at the 11 o'clock and 1 o'clock position in front of the trainee and 2) an additional, separate loudspeaker located inside the audio bag under the trainee's seat. The investigator tested the sound pressure levels at each 1/3-octave frequency band at the trainee's ear location and compared the results to the DRILCOM system pink noise measurements. Measurements were taken using a with a Larson-Davis Model 2900 spectrum analyzer (SN: A0280) with a ½-inch Larson-Davis 2559 microphone (SN:2575) and a Larson-Davis 9000C Preamp (SN: 0521). The microphone was calibrated at 94.0 dBA (1000 Hz tone) using a Quest QC-20 Calibrator (SN:

QOA070051). Measurements were recorded in small office room that was later used during the Phase II and Phase III investigations for PALAT system training and testing. The pink noise set at 55 dBA at the trainee's ear location was generated from a Mystic Marvels LLC. PNG-400 analog pink noise generator and amplified using a Stewart Audio AV30MX-2 30 Watt amplifier. The pink noise was presented by a single Cambridge Audio Minx Min 12 loudspeaker located directly under the trainee seat positioned facing forward (purple dotted line in Figure 51) and by two Cambridge Audio Minx Min 12 loudspeakers located 15-degrees below the azimuthal speakers at the 11 o'clock and 1 o'clock position in front of the trainee (red dashed line in Figure 51). The identical measurement setup was used in the DRILCOM hemi-anechoic room to measure the DRILCOM system pink noise sound pressure levels at each 1/3-octave frequency band at the trainee's ear location. The DRILCOM system pink noise source set at 55 dBA was played through the two JBL SoundPower SP215-6™ loudspeakers via a compact disc player and a QSC CX1102™ power amplifier (green dash-dot line in Figure 51). Figure 51 shows the single Cambridge Audio Minx Min 12 loudspeaker located inside the audio bag under the trainee's chair resulted in the closest sound pressure levels to the DRILCOM system measurements. As a result of these data, as well as the portability advantage, the decision was made to locate the PALAT system pink noise loudspeaker inside the audio bag.

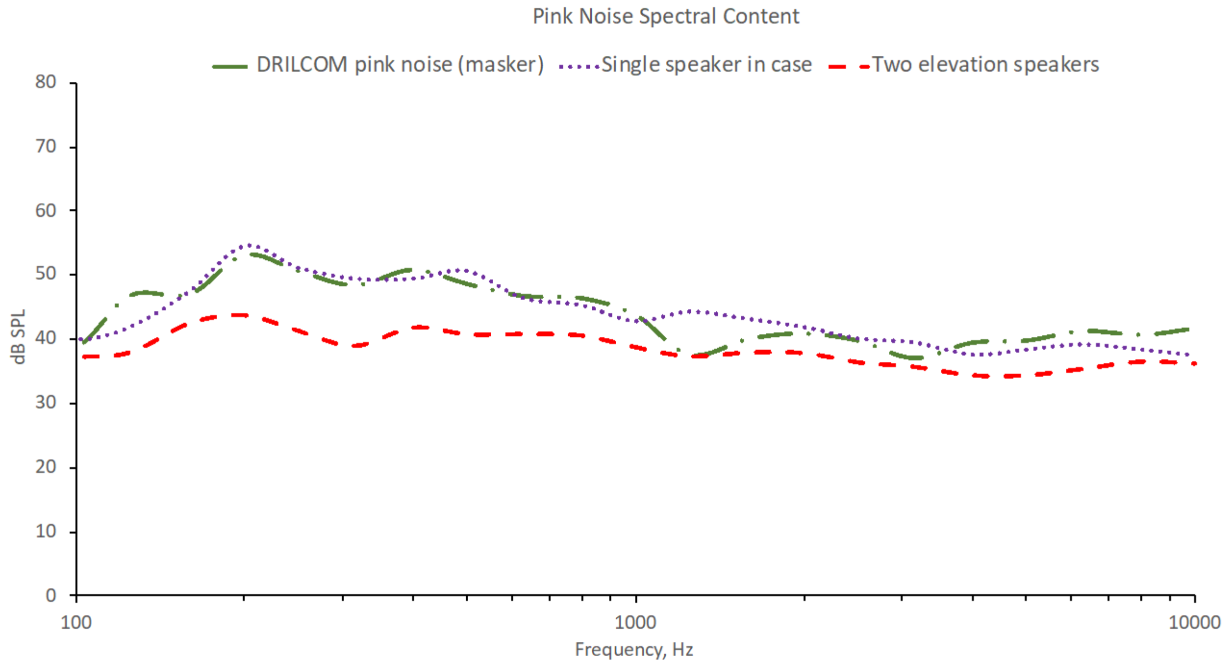


Figure 51. Spectral content of the pink noise masker produced by the DRILCOM system (green dash-dot line) and two alternative pink noise loudspeaker locations for the PALAT system, two Cambridge Audio Minx Min 12 loudspeakers positioned on vertical uprights at 11 o'clock and 1 o'clock (red dashed line) and one Cambridge Audio Minx Min 12 loudspeaker positioned under the trainee chair facing forward (purple dotted line) in 1/3 octave-band frequency. Overall sound pressure level of 55 dBA for the pink noise.

The last major component to be selected for the PALAT system was the laptop or tablet computer user interface. The DRILCOM system used an experimenter-controlled desktop computer equipped with an additional computer monitor and mouse for the participant to control the localization testing screen. The PALAT system required the trainee to operate the system without the assistance of an experimenter. The investigator evaluated several laptop and tablet options before selecting the Microsoft® Surface Pro as the user interface for the PALAT system. The Surface Pro was compatible with the LabVIEW™ software program used for localization testing. It also was lighter than most of the laptop options. The Surface Pro afforded users the ability to use a stylus pen to control the test or use the touch screen interface. Both options provided direct position control allowing the user to physically touch the control buttons and response locations (Wickens, Lee, Liu, & Becker, 2004). The tablet also allowed users to maintain the proper head position by allowing users to hold the tablet up in front of them as opposed to a laptop which would be typically placed on the user's lap requiring them to glance down to initiate the auditory signal and respond with the perceived location. Figure 52 displays the Microsoft® Surface Pro tablet connected to the audio case via a USB cable and 3.5mm audio cable (left) and a trainee operating the PALAT system via tablet user interface (right).

Figure 52. Microsoft® Surface Pro tablet user interface and audio case (left) and a trainee operating the PALAT system via tablet user interface (right).

2.6 PALAT Conclusions and Implications for Phases II and III

As hypothesized, the PALAT system was developed from primarily off-the-shelf components and was able to meet all design requirements established by the Office of Naval Research for size and portability, and to provide acoustically accurate localization cues similar to the full-scale laboratory grade DRILCOM system. However, the Phase II evaluation was still needed to test the capability of the PALAT to provide similar training benefits as the DRILCOM system. The prototype PALAT system structure measured 2-meters in diameter from the inside of the loudspeaker faces and 83 inches wide by 84 inches (7 feet) tall with the elevation loudspeaker poles attached. The expandable/collapsible frame folded down to 20 inches wide by

20 inches deep by 40 inches tall allowing the system to fit into a large shipping trunk that can be hand-carried with an additional audio case bag to store the audio components and tablet computer. One to two trainees were capable of assembling and disassembling the system. The PALAT system allows for azimuthal localization training and testing with 15° accuracy provided by the 24 small loudspeakers and 8 additional elevation loudspeakers (10 with two shared azimuthal loudspeakers) to train and test frontal elevation with 15° accuracy. However, the impacts of the smaller size halo array, smaller size loudspeakers, and 24 visible azimuthal loudspeakers on auditory localization performance were unknown.

The PALAT system placed the loudspeakers slightly over 1-meter from the trainee's ear. This resulted in the trainee possibly being locate in the near field for frequencies below 343 Hz. The dissonant tonal complex contained two tones within this range, 104 Hz and 295 Hz, but six other tones in the complex were clearly in the far field. Previous studies indicated the 1-meter radial distance likely would equate to little impact of the near field effect (Duda & Martens, 1998; Brungart & Rabinowitz, 1999). The small diameter Cambridge Audio Minx Min 12 loudspeakers were able to reproduce the DRILCOM auditory signals at similar sound pressure levels for all frequencies above 160 Hz. While this could reduce ITD localization cues, the dissonant signal contained two additional frequency tones below 1500 Hz capable of providing the requisite ITD cues for localization. The inability to produce higher sound pressure levels at 104 Hz may also reduce the impacts of the near field effects.

Another potential impact of the smaller PALAT system frame is the proximity of the loudspeaker to the trainee and to the nearest loudspeaker. Figure 53 shows the proportional size of the PALAT system (left semicircle) compared to the DRILCOM system (right semicircle). At a radius of 1.5 meters, the DRILCOM loudspeakers are separated by 0.78 meters, or 30.7 inches, whereas the distance between the 12 signal loudspeakers in the PALAT systems is only 0.52 meters, or 20.5 inches. This distance is reduced to approximately 10.5 inches between the very visible 24 azimuthal loudspeakers in the PALAT system. The impacts on auditory localization of both the reduced separation distance and the visibility of the loudspeakers in the PALAT system were unknown.

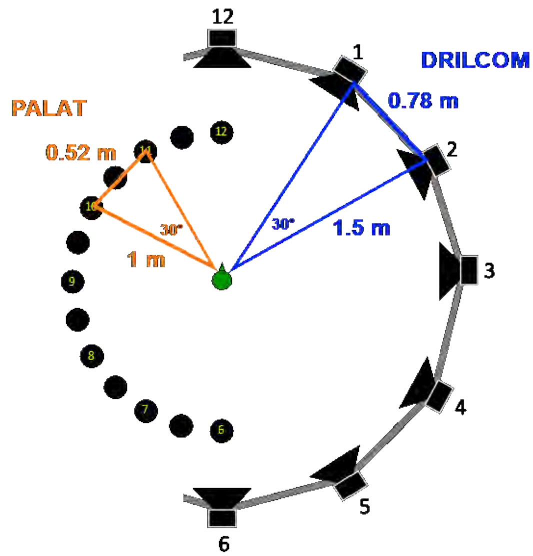


Figure 53. Schematic of the PALAT system (left semicircle) compared to the DRILCOM system (right semicircle) with radial and signal loudspeaker separation distances.

The last major difference between the DRILCOM system and PALAT system was the user interface. The tablet interface of the PALAT system allowed the trainee to physically touch the desired response location on the user interface screen as opposed to having to control a cursor on the screen with a computer mouse located several inches out of the visual frame of the DRILCOM system user interface. The investigator hypothesized that the response times on the PALAT system would be lower as a result of the improved user interface. However, the reduced size of the PALAT system, smaller loudspeakers, near field effects for lowest frequencies, and visibility of the loudspeakers (as compared to DRILCOM) were thought to have a small potential for confounding of response time measurements.

CHAPTER 3. Phase II: Development and In-Laboratory Investigation of a Portable Auditory Localization Acclimation Training System

3.1 Phase II Objectives

The primary objective of the Phase II in-laboratory investigation was to evaluate and validate the effectiveness of the Portable Auditory Localization Acclimation Training (PALAT) system via human subject experimentation in comparison to the full-scale, laboratory-grade DRILCOM system. The experiment also investigated the auditory localization skills acquisition while using the open ear, an in-the-ear TCAPS (3M™ PELTOR™ TEP-100), and over-the-ear TCAPS (3M™ PELTOR™ ComTac™ III). Both TCAPS represented the currently-deployed U.S. military "TCAPS-Lite" devices, which essentially are TCAPS that contain all features except radio communications. Auditory localization skills acquisition was examined over a sequence of five Learning Unit (LU) sessions totaling approximately 1.25 hours, from LU0 (pretest) to LU5 (posttest), per listening condition on each training system. Finally, the investigation determined the TCAPS effects on localization accuracy and response time, in 360-degrees on polar response. The PALAT system was evaluated against the DRILCOM system to determine if the portable system was capable of detecting differences in the effects of TCAPS devices on auditory localization performance. To meet these objectives, localization performance was compared via a within-subjects experiment using a full-factorial design with 12 normal-hearing participants who had no experience in localization testing or prior TCAPS use.

3.2 Phase II Methodology

This investigation aimed to validate the effectiveness of the PALAT system compared to the full-scale, laboratory grade DRILCOM system using the localization training protocol developed in Phase I of the overarching research effort. A series of auditory localization studies conducted at the Virginia Tech Auditory Systems Laboratory (VT-ASL) previously validated the

DRILCOM system's ability to measure localization in terms of accuracy, response time, and subjective rankings of participant-perceived localization ability in a laboratory setting (Casali & Robinette, 2014; Casali & Lee, 2016a; Casali & Lee, 2016b). Unique to this investigation, the localization training imparted by the portable system occurred in a semi-reverberant, office environment. Previous VT-ASL experiments trained and tested localization using the full-scale, laboratory grade DRILCOM system in a hemi-anechoic chamber. In addition, the listening condition order was counterbalanced in this investigation whereas previous VT-ASL experiments trained and tested open ear prior to training or testing localization under TCAPS listening conditions.

The participant first signed a consent form and then was audiometrically and demographically screened. Next, the participant was randomly assigned to a training system order, starting with either the DRILCOM system or PALAT system, and a listening condition order. Participant order was counterbalanced on training system and listening condition. The experiment consisted of six sessions per participant spread out over no more than two weeks. Each session included a full complement of training and testing under one listening condition using one training system as determined in the localization training protocol developed in Phase I of the overarching research effort (Cave, Thompson, Lee, & Casali, 2019). The participant began the experiment on their assigned training system under one listening condition. The participant completed LU0 (pretest) through LU5 with a 20-minute break between LU2 and LU3. Upon completion of the first session, the participant switched training systems and repeated the training and testing under the same listening condition. This process was repeated a total of three iterations until the participant completed training and testing under all three listening conditions on both training systems. Figure 54 displays a participant's progression through the experiment.

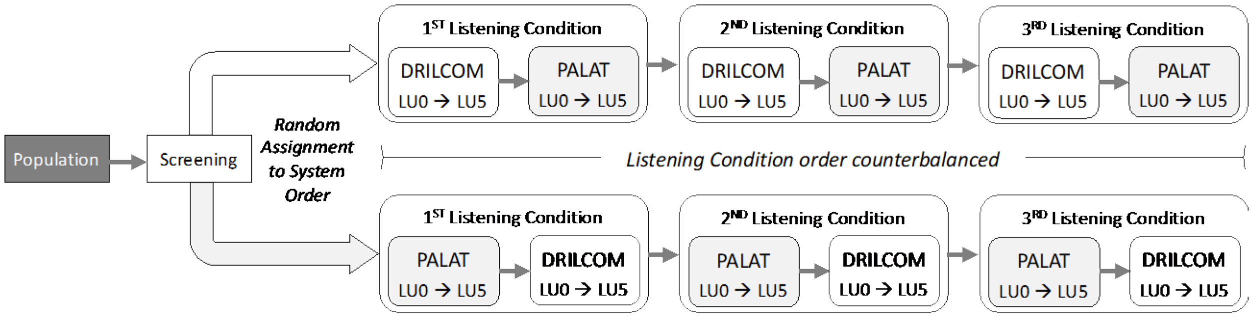
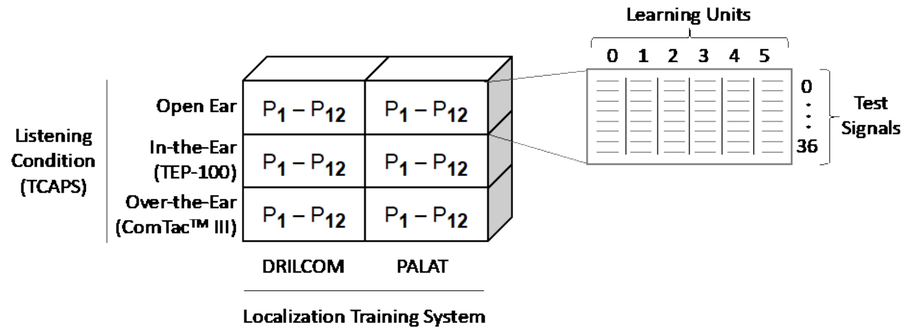


Figure 54. Phase II experimental design order.

3.3 Phase II Experimental Design

Phase II consisted of a full-factorial, repeated-measures experimental design involving 12 normal-hearing participants (Figure 55) who had neither prior experience in auditory localization training or testing and had never used TCAPS devices. The 2 x 3 full-factorial design involved two within-subjects independent variables: training system (DRILCOM and PALAT) and listening condition (open ear, TEP-100, and ComTac™ III). Results were measured using three groups of dependent measures: localization accuracy, response time, and participant subjective responses.



Experimental Design
1. Full factorial, pre-test post-test design
   - 2x Localization training systems (DRILCOM, PALAT)
   - 3x Listening conditions (open ear, In-the-ear, Over-the-ear)
2. Order of:
   - Localization training system – Counterbalanced order
   - Listening condition – Latin square, counterbalanced order
   - Target location (Speaker) – Random (During testing)
3. Participants
   - 18-45 years of age
   - 75-85% male to generalize military population
   - Normal Hearing
   - No experience with DRILCOM/PALAT system, localization training, or TCAPS

Dependent Variables
1. Localization accuracy
   - Percent absolute accurate (response matches signal azimuth)
   - Percent ballpark accurate (response ±15º signal azimuth)
   - Front-back confusion percent
2. Response time (seconds, with msec precision)
3. Perceived localization performance rating
   - Confidence in accuracy for each listening condition
   - Perceived effect of training on localization accuracy
   - Perceived effect on response time for each listening condition

Figure 55. Experimental design with repeated-measures subject assignment for Phase II, with independent variables, experimental order, participant assignment, and dependent measures listed.

The main objective of this investigation was to evaluate the effectiveness of the PALAT system to impart auditory localization skills compared to the DRILCOM system. As a result, it was important that a participant complete each listening condition on both systems before switching to a different listening condition. As an example, if a participant began session one training on the PALAT system using the in-the-ear TCAPS (TEP-100), then in session two they trained on the DRILCOM system using the same in-the-ear TCAPS (TEP-100). A Microsoft® Excel random number generator was used to assign 12 participant numbers to an arrival order. Participants who were assigned numbers 1 to 6 were assigned to begin training and testing using the DRILCOM system, and participants assigned numbers 7 to 12 were assigned to begin training and testing using the PALAT system. Two sets of an identical 3 x 6 Latin square were repeated to counterbalance the listening condition order for each participant, in an effort to guard against practice and order effects. The participant training system order was maintained for each listening condition throughout the study. Table 15 displays the participant order for the Phase II experiment by sex, training system order, and listening condition order.

Table 15. Participant study order by sex, training system order (random assignment based on arrival order), and listening condition by session (counterbalanced using a repeating 3 x 6 Latin square).

Arrival order	Participant Number	Training System order	Listening Condition order by Sessions 1 and 2	3 and 4	5 and 6
M1	P7	PALAT ⇒ DRILCOM	Open	ITE	OTE
M2	P1	DRILCOM ⇒ PALAT	Open	ITE	OTE
F1	P4	DRILCOM ⇒ PALAT	OTE	ITE	Open
F2	P12	PALAT ⇒ DRILCOM	ITE	Open	OTE
M3	P6	DRILCOM ⇒ PALAT	ITE	Open	OTE
M4	P5	DRILCOM ⇒ PALAT	Open	OTE	ITE
F3	P10	PALAT ⇒ DRILCOM	OTE	ITE	Open
M5	P9	PALAT ⇒ DRILCOM	OTE	Open	ITE
M6	P2	DRILCOM ⇒ PALAT	ITE	OTE	Open
M7	P8	PALAT ⇒ DRILCOM	ITE	OTE	Open
M8	P11	PALAT ⇒ DRILCOM	Open	OTE	ITE
M9	P3	DRILCOM ⇒ PALAT	OTE	Open	ITE

3.3.1 Independent Variables (IVs)

Independent Variable – Training system

Two within-subjects training system levels were used in this investigation: the DRILCOM full-scale, laboratory grade system and the PALAT system. Training system order was counterbalanced with half the participants beginning training and testing using each listening condition on the DRILCOM system and the other half beginning training and testing using each listening condition on the PALAT system. Descriptions of each training system are detailed in subsequent sections.

DRILCOM

The DRILCOM test battery and system were designed to test the Auditory Situation Awareness task elements of Detection, Recognition/Identification, Localization, and pass-through COMmunications (DRILCOM) (Casali & Lee, 2016a). The DRILCOM system is comprised of a 3-meter diameter circular, horizontal and vertical (front) localization apparatus consisting of 12 loudspeakers (horizontal) and 3 additional vertical loudspeakers (placed above the 330°, 0°, and 30° horizontal loudspeaker) housed in a large, hemi-anechoic room. Each horizontal loudspeaker is separated by a 30° azimuthal angle with one loudspeaker positioned directly in front of the participant (12 o'clock) so that the loudspeakers are positioned at each hour position on a clock face. The hemi-anechoic room measures 18-feet by 19-feet with 8.5-feet from floor to ceiling. The ceiling is comprised of acoustical drop Celotex™ panels and the walls are lined with two-inch thick Sonex™ eggshell acoustic foam panels (Casali & Lee, 2016a). The system uses 15-Behringer Behritone C50A powered loudspeakers with a 5.25-inch single cone driver, a QSC CX1102™ power amplifier, and two JBL SoundPower SP215-6™ loudspeakers to create background noise generated from a compact disc player. The loudspeakers are mounted on a circular steel pipe located approximately 1.14 meters in height above the floor (Casali & Lee, 2016a). The loudspeakers and metal ring are covered with acoustically transparent black fabric to conceal the location and number of loudspeakers present. An investigator control station is located outside of the loudspeaker ring and consists of a desktop computer to initiate the auditory training and testing. A small participant control station with a computer monitor and mouse are located in the middle of the ring to allow the participant to control the experiment and respond to auditory signals (discussed further in Experimental Procedure). Figure 56 displays a schematic of the DRILCOM system within the hemi-anechoic room.

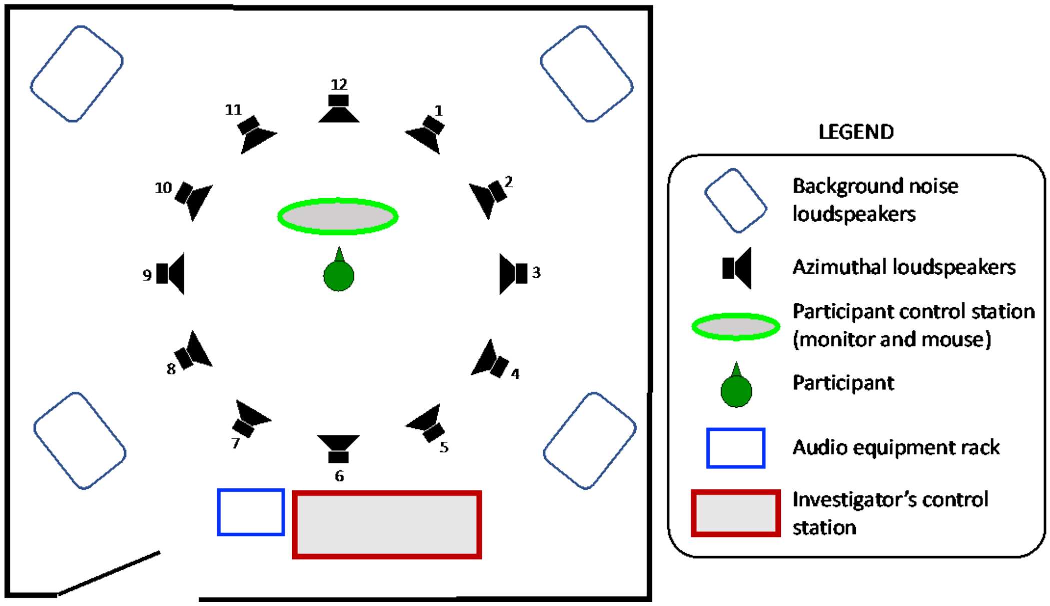


Figure 56. Schematic of DRILCOM system (only azimuthal loudspeakers shown) housed within the hemi-anechoic room. (Actual photograph appears in Figure 72.)

The investigator measured the spectral content of the dissonant tonal complex signal in situ using the DRILCOM system and verified that the stimulus contains frequency content necessary to provide interaural timing differences (ITDs) and interaural level differences (ILDs) (Figure 57). Measurements were taken using a with a Larson-Davis Model 2900 spectrum analyzer (SN: A0280) with a ½-inch Larson-Davis 2559 microphone (SN:2575) and a Larson-Davis 9000C Preamp (SN: 0521). The microphone was calibrated at 94.0 dBA (1000 Hz tone) using a Quest QC-20 Calibrator (SN: QOA070051). The dissonant signal was played at 70 dBA measured at the participant's center head position. A pink noise source set at 55 dBA was played through the two JBL SoundPower SP215-6™ loudspeakers via a compact disc player and a QSC CX1102™ power amplifier. The 55 dBA pink noise served to mask extraneous sounds during the experiment while still allowing for a +15 dBA signal to noise ratio. Figure 57 displays the spectral content of the pink noise in the DRILCOM room and the dissonant tone with the seven slightly shifted pure tone frequencies at 104, 295, 450, 737, 2967, 4959, 7025, and 7880 Hz. The dissonant tonal complex signal was shown to successfully provide both monaural and binaural auditory localization cues in a series of studies at VT-ASL (Casali & Lee, 2016a; Casali & Lee, 2016b; Casali & Robinette, 2014; Cave et al., 2019).

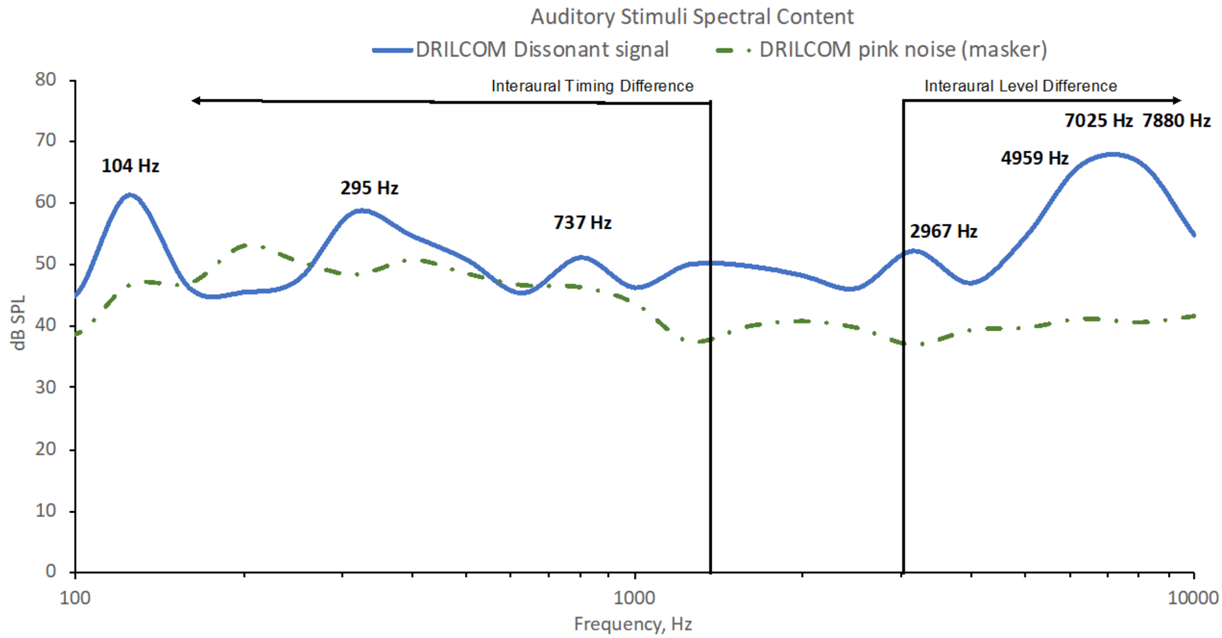


Figure 57. Spectral content of DRILCOM system pink noise (green dash dot line) and dissonant signal (blue solid line) in 1/3 octave-band frequency. Seven pure tones comprising the dissonant tone are labeled above the respective frequency. Overall sound pressure level is 70 dBA for the dissonant signal and 55 dBA for the pink noise masker.

PALAT

The PALAT system was designed to provide a portable version of the DRILCOM system capable of training and testing auditory localization. The PALAT system is a 2-meter diameter circular, horizontal and vertical (front) localization apparatus consisting of 32 loudspeakers with 24 loudspeakers (horizontal) and 8 additional vertical loudspeakers housed in a semi-reverberant room (which is intended to simulate a typical single office environment). Two of the horizontal loudspeakers are used during elevation testing to provide 10 vertical loudspeakers. All loudspeakers are separated by an angle of 15° from the center of the apparatus, or center head position of the participant. The horizontal loudspeakers are located one-meter from the participant. Each horizontal loudspeaker is separated by a 15° azimuthal angle with one loudspeaker positioned directly in front of the participant (12 o'clock) so that every other loudspeaker is positioned at each hour position on a clock face. The loudspeakers are mounted on a portable, collapsible frame consisting of 12 telescopic poles. The telescopic poles allow for horizontal loudspeaker heights of 43.5, 45.5, and 47.5 inches above the floor. The speaker heights are set to 45.5 inches above the floor, as used for the duration of the in-laboratory experiment. The PALAT system is controlled by the participant seated in the middle of the loudspeaker array via a Microsoft® Surface Pro running a LabVIEW™ software program. The system uses Cambridge Audio Minx Min 12 loudspeakers with a 2.25-inch single cone driver, a Stewart Audio AV30MX-2 two channel stereo mixer amplifier, and a Numato 32 channel USB relay module. A pink noise source set at 55 dBA is played through an additional Minx Min 12 loudspeaker via a Mystic Marvels LLC. PNG-400 Pink Noise Generator and additional Stewart Audio AV30MX-2 amplifier mounted inside the audio equipment case located under the participant chair (Figure 58). For this experiment, the two poles housing the elevation speakers were removed during all in-laboratory testing and training in order to present a more uniform apparatus for azimuthal-only testing (all 24 speakers aligned on one horizontal plane).

The PALAT system was located in a small office room on the fifth floor of Whittemore Hall at Virginia Tech for the experiment. The PALAT room is approximately 13.5 feet by 12.5 feet and for the experiment, it contained typical office furniture including a desk, chairs, wooden bookshelf, metal storage cabinets, dry-erase board, metal window blinds, carpeted floor, and dropped panel ceiling. In addition, a metal portable audiometric booth was located in the corner of the room. Figure 58 displays a schematic of the PALAT system within the small office room. The small office space was selected due to its semi-reverberant environment that represents a typical setting where the military (or industry) would employ the PALAT system. Likewise, the investigator decided to leave the acoustically reflective furniture inside the office under the assumption that users of the PALAT system would not be able to move all of the office furniture out of the room during localization training. The PALAT system was positioned in the room so that no speaker was within two feet of any reflective surface but the system was not centered in the room. Centering the PALAT system in the room would be preferred in order produce a more uniform reflective surface and render equidistant sound ray distances for reflections back to the subject's position. Hartmann (1983) found that early reflections off side walls had the largest decremental effect on localization performance because the angle of reflections off of side walls do not agree with the direct sound wave. The investigator decided against centering the portable system assuming that future users of the PALAT system may have similar limitations due to varying room sizes and shapes or furniture that may not be able to remove from the room.

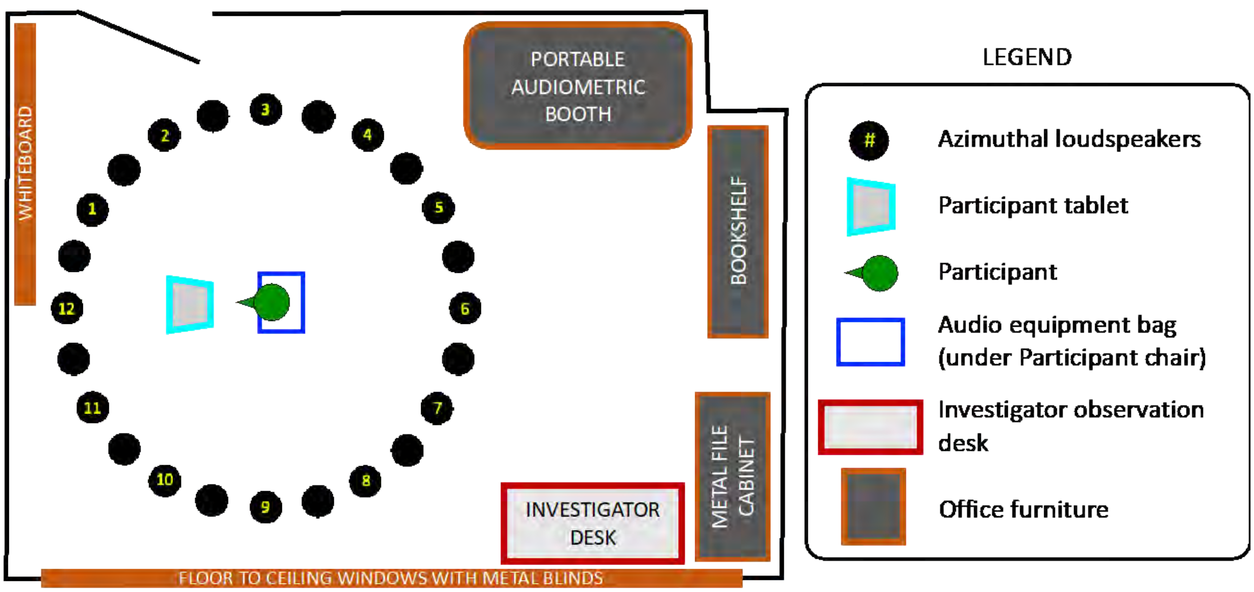


Figure 58. Schematic of the PALAT system (only azimuthal loudspeakers shown) located in a semi-reverberant office room at Virginia Tech. (Actual photograph appears in Figure 72.)

The investigator measured the spectral content of the dissonant tonal complex signal in situ using the PALAT system and verified that the stimulus contains frequency content necessary to provide interaural timing differences (ITDs) and interaural level differences (ILDs) (Figure 59). Measurements were made using the same equipment described above to measure the DRILCOM system. The dissonant signal was played at 70 dBA measured at the participant's center head position. A pink noise source set at 55 dBA was played through one Cambridge Audio Minx Min 12 loudspeaker located under the participant chair in the center of the PALAT system. The pink noise loudspeaker was amplified using a Stewart Audio AV30MX-2 two channel stereo mixer amplifier. The 55 dBA pink noise served to mask extraneous sounds during the experiment while still allowing for a +15 dBA signal to noise ratio. Figure 59 displays the spectral content of the pink noise in the PALAT room and the dissonant tone with the seven slightly shifted pure tone frequencies at 104, 295, 450, 737, 2967, 4959, 7025, and 7880 Hz.

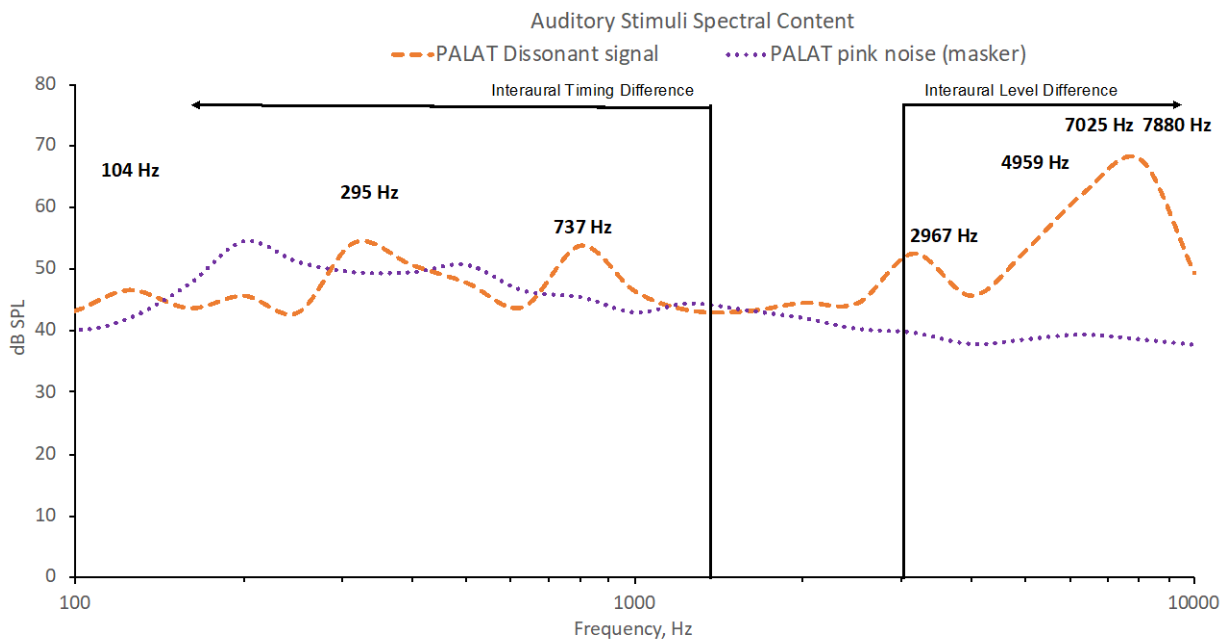


Figure 59. Spectral content of PALAT system pink noise (purple dotted line) and dissonant signal (orange dashed line) in 1/3 octave-band frequency. Seven pure tones comprising the dissonant tone are labeled above the respective frequency. Overall sound pressure level of 70 dBA for the dissonant signal and 55 dBA for the pink noise masker.

The investigator compared the spectral content of the dissonant signal and pink noise between both the DRILCOM and PALAT systems to ensure that experimental results were not confounded by differences between auditory cues. The PALAT system was able to produce similar sound pressure levels across six of the seven slightly off-octave-shifted pure tone frequencies (Figure 60). The smaller-sized loudspeakers of the PALAT system were unable to produce the same sound pressure level of the dissonant signal at 104 Hz. However, the spectral content of the dissonant signal from the PALAT system includes two low frequencies at 295 Hz and 737 Hz in order to produce interaural timing differences need for auditory localization. The pink noise masker sound pressure levels produced by both systems were within 3 dB SPL across the entire frequency spectrum from 80 Hz to 10000 Hz with the exception of the 1250 Hz octave band where the difference was 6.8 dB SPL (Figure 61). Thus, based on these measurements, it was concluded that for both the dissonant signal and the pink noise masker, the spectral levels were sufficiently close in value between DRILCOM and PALAT.

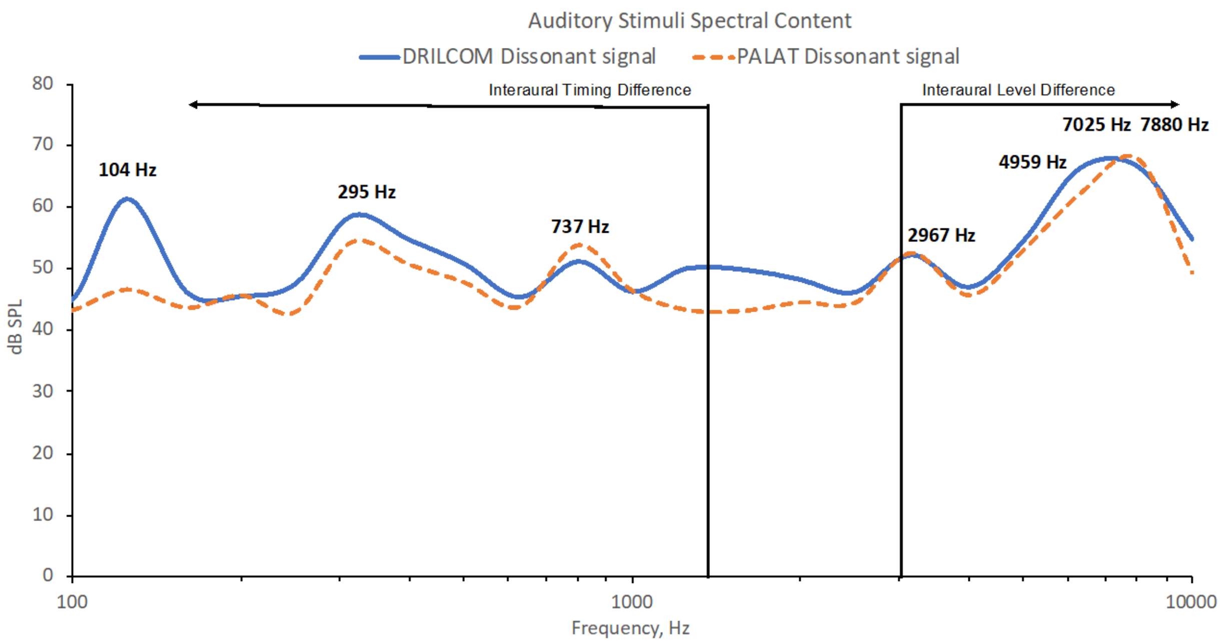


Figure 60. Spectral content of dissonant signal produced by the DRILCOM system (blue solid line) and the PALAT system (orange dashed line) in 1/3 octave-band frequency. Seven pure

tones comprising the dissonant tone are labeled above the respective frequency. Overall sound pressure level of 70 dBA for the dissonant signal.

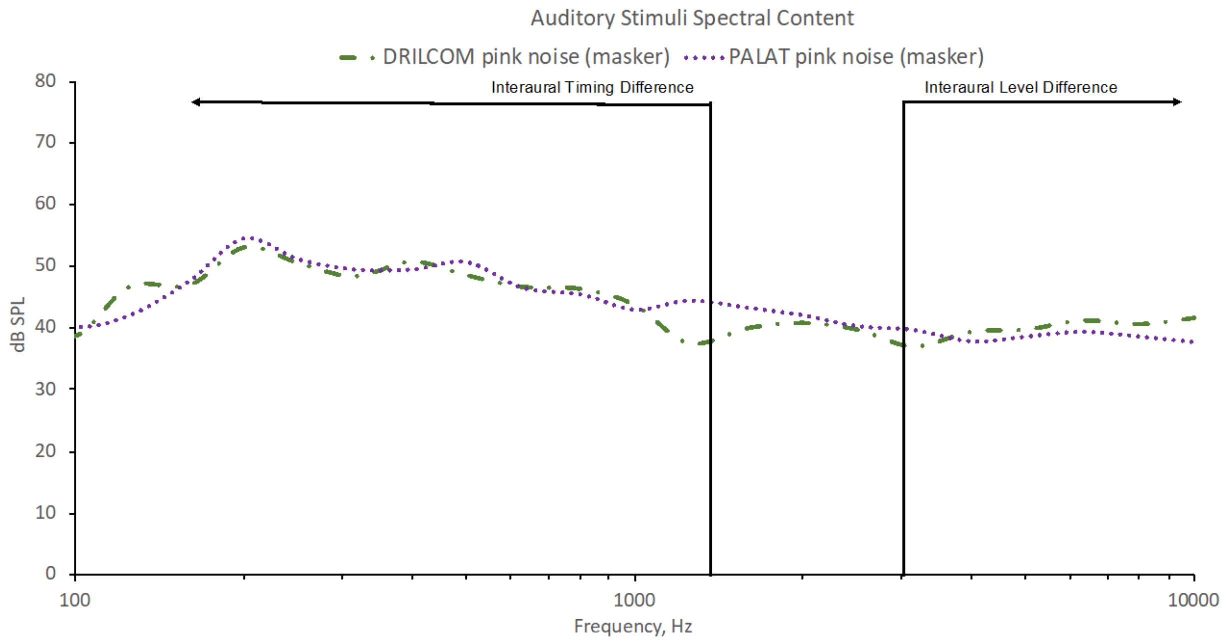


Figure 61. Spectral content of the pink noise masker produced by the DRILCOM system (green dash dot line) and the PALAT system (purple dotted line) in 1/3 octave-band frequency. Overall sound pressure level of 55 dBA for the pink noise.

The small office housing the PALAT system and the large, hemi-anechoic laboratory room housing the DRILCOM system were also tested by the investigator to find the ambient noise floor and reverberation time (RT60). Measurements were made with a Larson-Davis Model 831 Sound Level Meter (SN: 0002486) with a ½-inch Larson-Davis 2575 measurement microphone (SN: LW131180) and Larson-Davis PRM831 Preamp (SN: 017153). The microphone was calibrated at 94.0 dBA (1000 Hz tone) using a Quest QC-20 Calibrator (SN: QOA070051). The investigator performed three measurements from five locations within each room, one in the center of the room and approximately one-meter from each room corner. Reverberation Time 60 (RT60) measurements were taken using an impulse noise at approximately 120 dBA produced by hitting together two-wooden 2-inch x 4-inch blocks. The RT60 measurements were calculated using a 30 dB decrease in level to avoid interference with the noise floor. The calculation to extrapolate the RT30 values to RT60 values was performed automatically by the sound level meter. The mean noise floor and RT60 values of three measurements at five locations are shown in Table 16 for both the DRILCOM and PALAT rooms. The mean noise floor measurements between the two systems was nearly identical for frequencies over 1000 Hz. However, the noise floor in the semi-reverberant PALAT room measured 8.1 db SPL and 2.6 db SPL higher than the hemi-anechoic DRILCOM room at 250 Hz and 500 Hz octave-band frequencies respectively. The reverberation time was much higher within the PALAT room than the DRILCOM room.

Table 16. Mean noise floor and reverberation time (RT60) measurements of the DRILCOM and PALAT rooms, as measured in 1/3 octave-band frequency.

		Frequency (Hz)					
Room		250	500	1000	2000	4000	8000
Noise Floor (dB SPL)	DRILCOM	34.0	30.5	31.3	33.9	36.6	40.0
	PALAT	42.1	33.1	31.5	33.5	36.6	40.0
RT60 (ms)	DRILCOM	408	272	182	144	119	110
	PALAT	407	402	348	339	410	396

Independent Variable – Listening Condition

Three within-subjects listening conditions (open ear, in-the-ear TCAPS, and over-the-ear TCAPS) were used in this investigation to encompass the type of hearing protectors currently used by ground-combat service members in the U.S. Armed Forces. Listening conditions were counterbalanced using a repeating 3 x 6 Latin square resulting in two sets of every combination of order, one identical set for each beginning training system. Depending on the mission requirements, most service members serving in a combat role are issued, or equipped with, an in-the-ear or over-the-ear TCAPS device. The 3M™ PELTOR™ TEP-100 Tactical Earplug and the 3M™ PELTOR™ ComTac™ III headset were chosen as the in-the-ear and over-the-ear TCAPS-Lite, respectively. The TEP-100 and ComTac™ III were selected because they represent the two program of record products under the U.S. Army TCAPS-Lite Program, or fielding program for TCAPS without external device connections. Testing the TEP-100 and ComTac™ III aligned with specific military applications anticipated by the Office of Naval Research sponsor. Additionally, the results obtained using these devices can generalize to both in-the-ear and over-the-ear products used by law enforcement and emergency personnel. Descriptions of each listening condition are detailed in subsequent sections.

Open ear

The open ear listening condition was included in this investigation for several reasons. First, testing the open ear condition established a baseline performance, enabling a within-subjects comparison of training effect for each TCAPS device. Secondly, the open ear condition is the most commonly-encountered listening condition for service members in training and combat environments where hazardous noise exposure is not imminent or expected, but threat or hazard localization remains paramount. Lastly, as previously covered in the literature review, several studies have identified barriers to HPDs and TCAPS compliance. Abel (2008) and Bevis et al. (2014) specifically described discomfort and a perceived loss of auditory situation awareness as reasons for non-compliance by service members. In Bevis et al. (2014), all 16 focus groups mentioned that auditory localization was negatively affected by hearing protection devices. One British Army Soldier stated, “If you can’t locate that position then you’re redundant” (Bevis et al., 2019, p131). Therefore, by examining localization performance with the open ear, the influence of device-imposed changes to environmental cues and comfort could be eliminated.

In-the-ear TCAPS (TEP-100)

The earplug-style 3M™ PELTOR™ TEP-100 Tactical Earplug is an active, or powered electronic sound transmission, in-the-ear hearing protection device, shown in Figure 62. The

TEP-100 Tactical Earplugs are issued as a set of two identical, rechargeable electronic earplugs with a recharging case. For testing purposes, the investigator designated a right and left ear device in each set according to serial numbers. The right and left device designations were maintained throughout the study to reduce confounding effects of differences between earplugs. The 3M™ PELTOR™ level-dependent technology is advertised by its manufacturer to "provide hearing protection, and helps improve situational awareness and communication" (3M, 2016a, p1). As a passive earplug, the TEP-100 is advertised by its manufacturer to provide a mean attenuation of 23 NRR according to the EPA-required labeling on the device (3M, 2016a). The TEP-100 is compatible with several styles of eartips, including the 3M™ PELTOR™ Ultrafit eartips shown in Figure 62 which are the standard issue version for the military. As a result, each participant in this experiment was fitted with one of the three sizes of Ultrafit eartips with the TEP-100. The investigator, who was trained by U.S. Army audiological personnel in the earplug fitting process, conducted a visual inspection of each participant's ear canal and ensured the participant was fitted with the proper Ultrafit eartip size.

Figure 62. 3M™ PELTOR™ TEP-100 electronic earplug-style TCAPS device.

The TEP-100 tactical earplug is equipped with two volume settings, "normal" and "high," that is operated by a single button. The investigator tested the TEP-100 volume settings to identify the unity gain setting. Unity gain was previously defined by Casali and Lee (2016a) as the state where the electronic gain control is set to overcome or offset the passive attenuation of the earplug and provide as close to natural hearing as possible. Four TEP-100 devices loaned to the Virginia Tech Auditory Systems Laboratory, comprising two devices from U.S. Army PEO Soldier and two devices from 3M™, were tested in a reverberation chamber to identify the unity gain setting during Phase II of the overarching experiment. Three TEP-100 devices were evenly assigned between the participants for the experiment.

The following steps were performed to identify the unity gain setting for the TEP-100. A ½ inch Larson-Davis 2575 measurement microphone (SN: 2559) and Larson-Davis 9000C Preamp (SN: 0521) were placed in the center of the reverberation chamber and connected to a Larson-Davis 2900 Model Spectrum Analyzer (SN: A0280) at an investigator table located outside of the chamber. The microphone was calibrated at 94.0 dBA (1000 Hz tone) using a Quest QC-20 Calibrator (SN: QOA070051). A pink noise signal was generated via a MATLAB® program and measured at 70 dBA, 10 second Leq, fast time constant. Next, an acoustical test manikin, known as KEMAR (Knowles Electronics Manikin for Acoustic Research by GRAS), was positioned in the center of the reverberation chamber and the measurement microphone was fitted inside the left ear canal of the KEMAR and the right ear canal was occluded with tightly-packed putty for maximal attenuation. The pink noise signal was measured in the open ear listening condition at 77.6 dBA which served as the reference level for unity gain. Each TEP-100 earplug was then fitted in the left ear of the KEMAR and the sound pressure level of the pink noise signal was measured three times at each volume setting: off, or passive, setting, normal volume, and high volume.

The “normal” volume setting provided the closest unity gain for the TEP-100 and thus was the setting used for this experiment (Table 17). Figure 63 displays the sound pressure level measurements of the 70 dBA pink noise signal at each 1/3 octave-band frequency for the open ear and TEP-100 at normal volume setting on the KEMAR manikin. The sound pressure levels measured under the TEP-100 are noticeably lower from 100 Hz to 315 Hz than the open ear levels. The TEP-100 also did not transmit the pink noise at the 10000 Hz 1/3 octave-band frequency.

Table 17. Sound pressure level (SPL) of 70 dBA pink noise measured using KEMAR manikin for three sets of TEP-100 devices and mean (by left and right ear designation). Levels compared with the open ear 77.6 dBA reference level displayed as the delta (Δ), open ear – TEP-100.

		Device 1		**Device 2**		**Device 3**		**Mean**	
Listening Condition	**Gain level**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**
Open ear (Reference Level)		77.6		77.6		77.6		77.6	
TEP-100		Left (SN: 326)		Left (SN: 61389)		Left (SN: 36576)		Left ear	
	Off (passive)	30.1	(-47.4)	31.8	(-45.8)	29.1	(-48.5)	30.3	(-47.3)
	Normal	**73.5**	**(-4.0)**	**76.9**	**(-0.7)**	**77.2**	**(-0.4)**	**77.2**	**(-1.7)**
	High	84.4	(6.9)	87.4	(9.8)	88.0	(10.4)	87.8	(9.0)
		Right (SN: 292)		Right (SN: 64343)		Right (SN: 36173)		Right ear	
	Off (passive)	29.2	(-48.3)	32.0	(-45.6)	29.8	(-47.8)	31.1	(-47.3)
	Normal	**72.6**	**(-4.9)**	**78.2**	**(0.6)**	**75.2**	**(-2.4)**	**77.7**	**(-2.3)**
	High	83.4	(5.9)	89.1	(11.5)	86.0	(8.4)	88.5	(8.6)

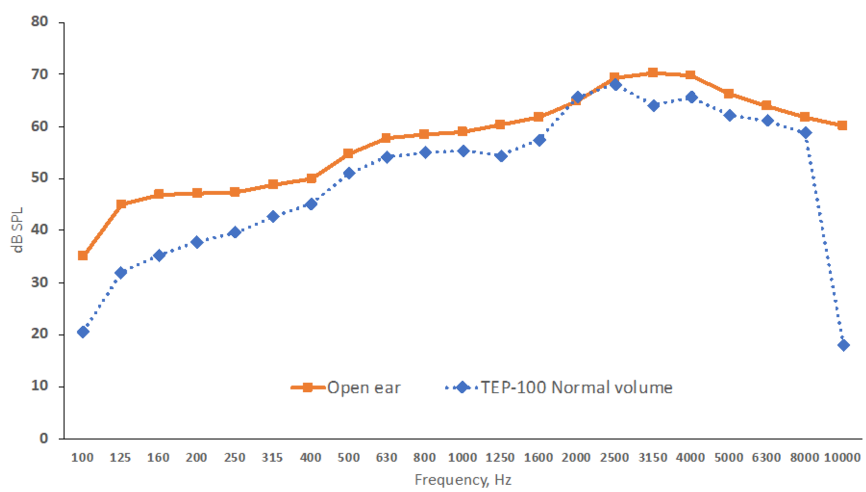


Figure 63. Mean frequency response of 70 dBA pink noise measured using KEMAR under open ear and TEP-100 devices on normal volume setting by 1/3 octave-band frequencies.

<u>Over-the-ear TCAPS (ComTac™ III)</u>

The earmuff-style 3M™ PELTOR™ ComTac™ III headset is an active, or electronic sound transmission, over-the-ear hearing protection device, shown in Figure 64. This battery-powered TCAPS is equipped with four volume settings and an additional boost mode to amplify low level external sounds to audible, but not hazardous levels, and pass them through the muff. According to the manufacturer's literature, the 3M™ PELTOR™ ComTac™ III utilizes a proprietary digital audio circuit to compress hazardous noise to a permissible safe exposure level of less than 82 dBA (3M, 2016b). As a passive headset, i.e., when the electronic sound transmission circuit is off, the 3M™ PELTOR™ ComTac™ III is advertised to provide a NRR of 23 (3M, 2016b).

Figure 64. 3M™ PELTOR™ ComTac™ III electronic earmuff-style TCAPS device.

Three ComTac™ III headsets were loaned to the Virginia Tech Auditory Systems Laboratory, two headsets from U.S. Army PEO Soldier and one headset from 3M™, for the study. All three headsets were tested to identify the unity gain setting using the same procedure described above. The highest volume setting, or fourth increase from default, provides the closest unity gain for the ComTac™ III and was used for this experiment (Table 18). Figure 65 displays the sound pressure level measurements of the 70 dBA pink noise signal at each 1/3 octave-band frequency for the open ear and ComTac™ III at the highest volume setting on the KEMAR manikin. The sound pressure levels measured under the ComTac™ III are noticeably lower from 4000 Hz to 10000 Hz than the open ear levels, which is indicative that the sound transmission circuit has gone into signal compression at 70 dBA for a pink noise input.

Table 18. Sound pressure level (SPL) of 70 dBA pink noise measured using KEMAR manikin for three sets of ComTac™ III devices and mean. Levels compared with the open ear 77.6 dBA reference level displayed as the delta (Δ), open ear – ComTac™ III.

		Device 1 (SN: 7607)		Device 2 (SN: 1098)		Device 3 (SN: 1099)		Mean	
Listening Condition	**Gain level**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**
Open ear (Reference Level)		77.6		77.6		77.6		77.6	
ComTac™ III	Off (passive)	38.0	(-39.6)	40.7	(-36.8)	40.1	(-37.5)	39.6	(-38.0)
	1 (Low)	57.5	(-20.1)	56.8	(-20.7)	57.4	(-20.2)	57.2	(-20.4)
	2	63.4	(-14.2)	62.8	(-14.7)	63.3	(-14.3)	63.2	(-14.4)
	3	69.4	(-8.2)	68.8	(-8.7)	69.3	(-8.3)	69.2	(-8.4)
	4 (High)	**75.4**	**(-2.2)**	**74.8**	**(-2.7)**	**75.2**	**(-2.4)**	**75.1**	**(-2.5)**

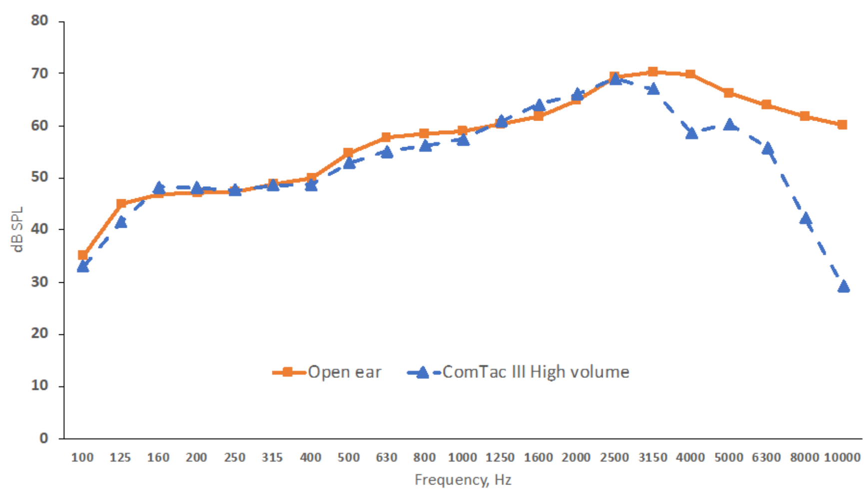


Figure 65. Mean frequency response of 70 dBA pink noise measured using KEMAR under open ear and ComTac™ III devices at high volume setting by 1/3 octave-band frequency.

3.3.2 Dependent Measures

Three classes of dependent measures were used during Phase II to test localization performance: 1) localization accuracy, 2) response time, and 3) subjective ratings. The following sections describe each dependent measure in detail.

Localization accuracy

Two measures of localization accuracy were recorded and analyzed: 1) absolute correct response scores and 2) number of front-back errors. Each test in this investigation presented three dissonant tonal signals from each of the 12 loudspeakers locations in random order for a possible maximum score of 36 correct responses on each test. The 12 signal locations were separated azimuthally by 30° resembling the 12-hour positions on an analog clock face. Military service members are trained to identify and communicate threat direction or points-of-interest using the 12 clock face number positions with 12 o'clock serving as the frontal midline reference (Department of the Army, 2017). For example, if a military unit were on a patrol walking through the woods in a northerly direction and heard gunshots from an enemy located directly to the east, the members of the unit would yell, "contact, enemy 3 o'clock." Thus, the investigator decided to present signals from all 12 clock face azimuthal locations, as they would be used for directional location by service members. A series of previous auditory localization studies conducted using the DRILCOM system presented the same 12 azimuthal locations during training but allowed for 24 response locations during testing, rendering one "dummy" position between each real sound source (Casali & Lee, 2016a; Casali & Lee, 2016b; Casali & Robinette, 2014; Cave et al., 2019). The same procedure of training with 12 azimuthal locations and providing 24 response options during testing was followed for the DRILCOM system in this investigation. However, the investigator designed the PALAT system to present 24 azimuthal locations during both training and testing to allow the participant to select a direction between two adjacent clock face positions if they were unsure of the exact signal location. Due to the differences in user interface training screens, the participant was informed before the experiment that only the 12 loudspeakers representing the 12 clock face positions would present signals during the study but was still given the option to choose any of the 24 response locations. As a result, participants very rarely selected the dummy loudspeaker locations in between the 12 actual signal locations. This response behavior resulted in the absolute and ballpark, or within ±15° of signal location, accuracy scores to be redundant. Figure 66 shows the training screen of the DRILCOM system participant screen (left) that the participant used during training displaying 12 signal locations (grey circles arranged like a clock face). Figure 66 shows the training screen of the PALAT system Surface Pro computer tablet (right) that the participant used during training displaying 24 response options (black circles) and 12 signal locations (black circles marked with yellow numbers).

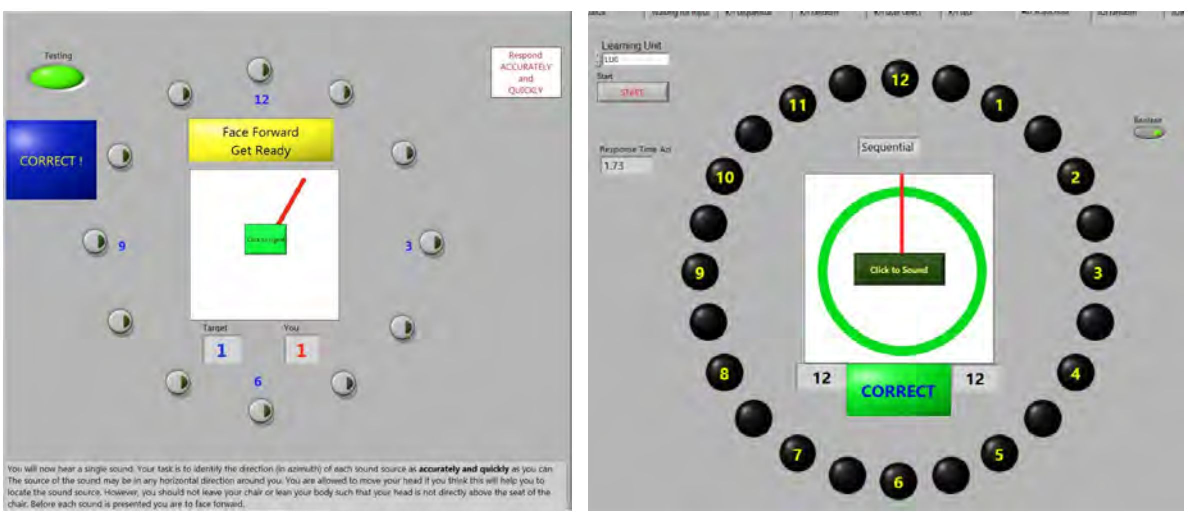


Figure 66. DRILCOM system participant training screen (left) displaying 12-signal locations (grey circles) and PALAT system participant training screen (right) displaying 24-response options (black circles) and 12-signal locations (black circles with yellow numbers).

Figure 67 shows the test screen of the DRILCOM system participant screen (left) displaying 24 response options (grey circles) with four loudspeaker locations marked (12, 3, 6, and 9 o'clock positions) to orient the participant. Figure 67 shows the PALAT system Surface Pro computer tablet (right) that the participant used during the training and testing displaying 24 response options (black circles) and 12 signal locations (black circles marked with yellow numbers).

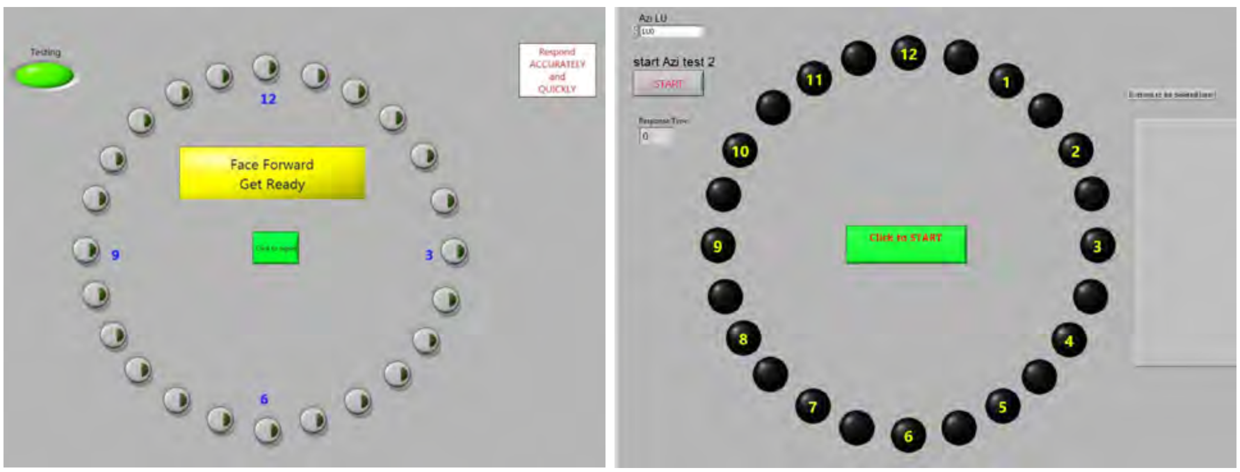


Figure 67. DRILCOM system participant testing screen (left) displaying 24-response locations (grey circles) and PALAT system participant testing screen (right) displaying 24-response options (black circles) and 12-signal locations (black circles with yellow numbers).

1. ***Absolute correct response scores** (also referred to as **absolute score**):* the total number of occurrences in which the participant responded with the exact azimuthal location of the signal location. Figure 68 displays an example of an absolute correct response indicated by the arrow if the signal originated from the 1 o'clock position.

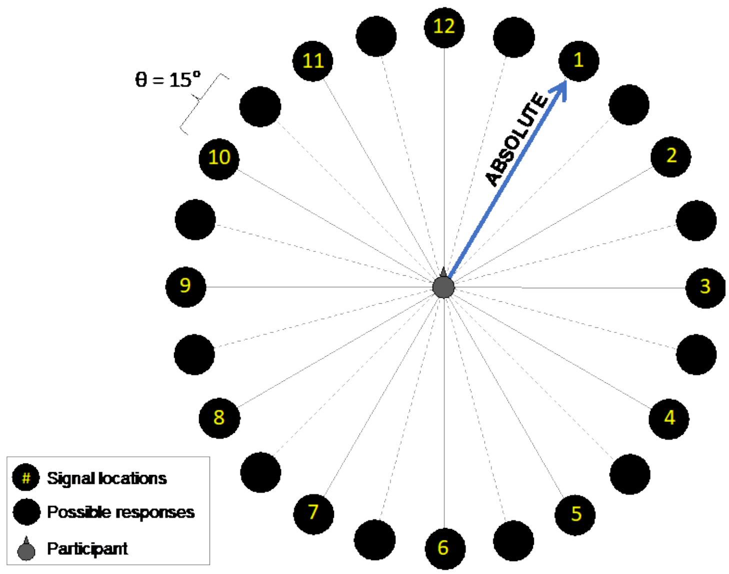


Figure 68. Absolute correct response (arrow) when the signal emanates from the 1 o'clock position.

2. *Front-back reversal errors:* the total number of occurrences in which the participant responded with an azimuthal location in the back (to the rear of participant) from 4 o'clock to 8 o'clock (120-degees to 240-degrees) when the signal was presented in the front from 10 o'clock to 2 o'clock (300-degrees to 60-degrees) and vice-versa. This window for front and back reversals is consistent with the new ANSI S3.71 standard window from 290-degrees to 70-degrees in front of the participant and 110-degrees to 250-degrees behind the participant (American National Standards Institute (ANSI), 2019). However, this experiment's operational definition of front-back reversal differs from the ANSI standard by allowing front-back reversals to occur if the difference between the source and response crosses the median plane. For example, a front-back reversal occurs in this experiment if a sound originates from the 7 o'clock position and the participant responds with the 1 o'clock position. The investigator felt this offered a more realistic operational definition of front-back reversals for auditory situation awareness in military operations. If a U.S. service member perceived a gunshot from the 1 o'clock position (in front of them) that actually originated from 7 o'clock (behind them), then the

service member would have made a front-back reversal that could be detrimental to survivability, but that is not considered a reversal in the language of the ANSI S3.71 standard (ANSI, 2019). Figure 69 displays the front and back regions where either the signal originated and the response was selected to constitute a front-back reversal error if the signal and response were in opposite regions.

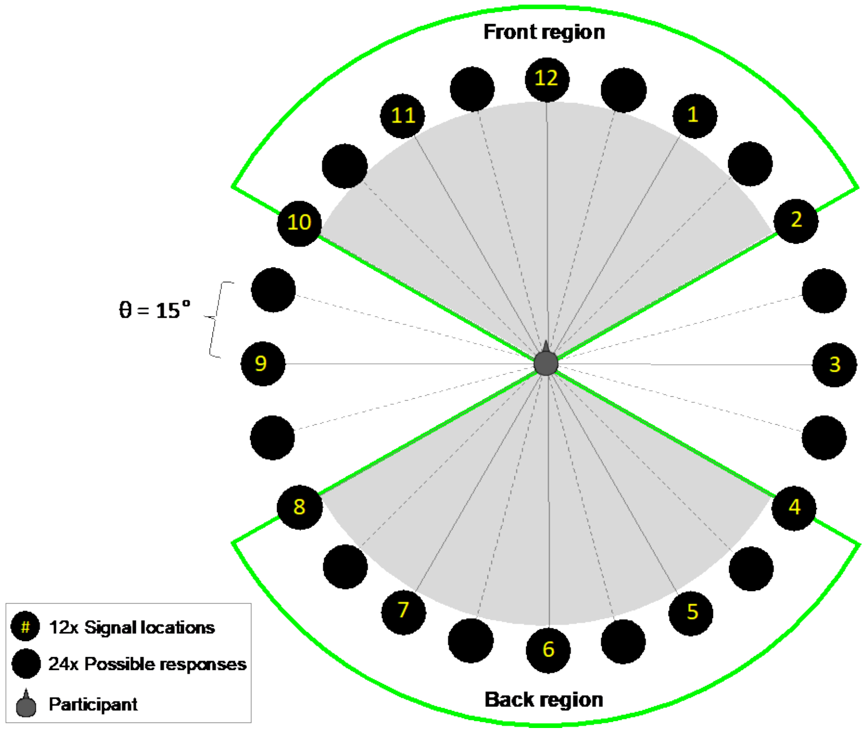


Figure 69. Front and back regions (shaded regions) depicting the range of signal locations and response locations for possible front-back reversal errors.

Response time

Response time was measured as the duration of time occurring from signal onset, dissonant tone, to the participant response selection on the computer (DRILCOM) or tablet (PALAT). Response time was automatically calculated via the LabVIEW^{TM} computer program on the DRILCOM system desktop computer and the PALAT system Surface Pro computer tablet. The response time clock onset was triggered by the participant selecting the green “Click to Signal” icon (DRILCOM) or “Click to START” icon (PALAT) located in the center of the test screen (Figure 66). Selecting the “Click to Signal” icon or “Click to START” icon simultaneously presented the dissonant tone. The response time clock offset occurred when the participant selected a speaker icon on the response display. A window located on the left side of the test screen displayed the running clock. After response selection, the display showed the most recent response time, allowing the participant to view their response time. Response times were recorded in 100 millisecond resolution. The maximum allowable response time was set at 10 seconds. Mean response times were calculated for each LU and used as the dependent measure score.

Subjective ratings

The participant completed a questionnaire at the conclusion of every session, LU0 through LU5 for each listening condition on each training system (Appendix F). Upon completion of the first, third, and fifth sessions, or first training system under each listening condition, the questionnaire included 10 questions focused on evaluating the training effectiveness and usability of the training system. After completion of the second, fourth, and sixth sessions, or second training system under each listening condition, the questionnaire included the same 10 questions so that comparisons could be made between training systems. An additional 10 questions were included to compare the effectiveness and usability between training systems. Questions 11 through 14 asked the participant to compare the second or most recent system to the first system under each listening condition based on their confidence in ability to localize, and how the training, system user interface, and system room environment impacted ability to localize. Question 15 through 20 then asked the participant rate their preferred training system on the aspects of confidence in accuracy and quick decision in localizing the signal, and preference of room environment, loudspeaker configuration, and user interface. All questions used a semantic differential, bipolar rating scale with seven discrete choices (example shown in Figure 70). The common 10 questions included on all questionnaires are listed below as they were presented following the DRILCOM system. Questionnaire verbiage following the PALAT system are shown in parenthesis.

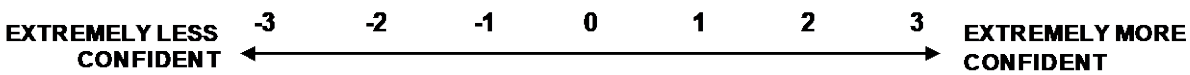


Figure 70. Example of semantic differential rating scale with seven anchors and bipolar descriptors at end points.

1. *Training impact on confidence in ability to localize*:
   "Rate how **training** using the DRILCOM (PALAT) system **impacted your confidence** in your ability to localize sounds, from **before to after** all the training you received using this system," from -3 (extremely less confident) to 3 (extremely more confident).
2. *Impact of the proximity of the loudspeakers on ability to train to localize sound*:
   "Rate the **impact** you felt the **proximity (distance) of the loudspeakers** of the DRILCOM (PALAT) system contributed to your **ability to train to localize** sounds," from -3 (extremely negative impact) to 3 (extremely positive impact).
3. *Ease of use of the system*:
   "Rate how **easy it was to operate** the DRILCOM (PALAT) system hardware and software during your localization training," from -3 (extremely difficult) to 3 (extremely easy).
4. *Impact of the room environment on ability to train to localize sound*:
   "Rate the **impact** you felt the **room environment** of the DRILCOM (PALAT) system contributed to your **ability to train to localize** sounds," from -3 (extremely negative impact) to 3 (extremely positive impact).
5. *Training impacts on ability to localize sound:*
   "Rate how much you feel your **ability** to determine sound location improved as a **result of training** with this system," from -3 (extremely less capable) to 3 (extremely more capable).

6. *Difficulty in judging the signal location*:
   "Rate how **difficult** it was to judge the **location** of the sounds **using this system**," from -3 (extremely difficult) to 3 (extremely easy).
7. *Impact on reaction time before to after training:*
   "Rate how **training** using the DRILCOM (PALAT) system **impacted your reaction time** in determining sound location, from **before to after** all the training you received using this system," from -3 (extremely slower reaction time) to 3 (extremely faster reaction time).
8. *Impact of the user interface on ability to train to localize sound:*
   "Rate how much of an **impact** the DRILCOM (PALAT) system **user interface** (monitor, software, loudspeakers, wires, etc.) had on your **ability to train your sound localization skills**," from -3 (extremely negative impact) to 3 (extremely positive impact).
9. *Impact of room environment on reaction time*:
   "Rate how training in the **room environment** of the DRILCOM (PALAT) system **impacted your reaction time** in determining sound location," from -3 (extremely slower reaction time) to 3 (extremely faster reaction time).
10. *Impact of loudspeaker visibility on ability to train to localize:*
   "Rate the **impact** you felt the **hidden loudspeakers** of the DRILCOM system (visible loudspeakers of the PALAT system) contributed to your **ability to train to localize** sounds," from -3 (extremely negative impact) to 3 (extremely positive impact).

3.3.3 Participants

The Phase II human-subjects experiment was approved by the Virginia Tech Institutional Review Board (protocol number VT-IRB 11-047, Appendix E). In order to generalize to the U.S. Military population, participants were required to be between the ages of 18 to 45 years with up to 25% females (Defense Manpower Data Center (DMDC), 2018). The study sample consisted of 12 participants: 9 males and 3 females, age 20 to 33 years with a mean age of 26.5 years (*SD*=4.3). Participants were recruited from Virginia Tech and the surrounding communities.

Each participant was compensated $10 per hour and received a $25 bonus upon completion of the study.

Participants were required to have normal hearing and no previous experience with auditory localization studies or auditory skills training. All participants were screened for hearing thresholds not to exceed 25 dB HL at 250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz, with no threshold difference between each ear to exceed 15 dB (bilaterally-symmetrical). Following the participant's informed consent, they were otoscopically inspected to check for ear canal obstructions, irritation, or infections that could affect localization performance. Each participant received an otoscopic inspection and hearing test administered by an Active Duty Army Audiologist. If the participant passed otoscopic inspection, a manual pure-tone audiogram using a standard Hughson-Westlake procedure was conducted using a Beltone Electronics Corporation Model 119 Audiometer (SN: 10B0561, calibrated 26 December 2019). The test was performed in the VT-ASL portable test booth located in the same room as the PALAT system (Figure 71). Table 19 displays the mean pure-tone hearing level thresholds (dBHL) for all participants and by group. Following the audiogram, participants were screened to ensure no prior experience with localization training or TCAPS devices (Appendix G).

Table 19. Mean pure-tone hearing level thresholds (dBHL) for all participants.

		Frequency (Hz)							
	Ear	250	500	1000	2000	3000	4000	6000	8000
Participants	Right	3.3	2.9	1.7	0.8	1.3	0.0	5.8	4.2
	Left	6.7	4.2	3.3	0.0	0.8	0.0	2.1	4.2

Figure 71. Portable audiometric booth (left) co-located with the PALAT system in the office environment used for training and pretesting.

3.4 Phase II Apparatus

The Phase II investigation was conducted using two localization training systems located within two rooms on the fifth floor of Whittemore Hall on the campus of Virginia Tech. As described previously in the Independent Variables section, the DRILCOM system was housed in a large hemi-anechoic room that was part of the Virginia Tech Auditory Systems Laboratory space. The investigator utilized a small office room located a few doors down from the DRILCOM room to set up and operate the PALAT system (Figure 72). Both the DRILCOM and PALAT systems are equipped with loudspeakers mounted vertically in front of the participant to train and test elevation localization. However, the Phase II investigation only included azimuthal localization training and testing, as explained previously. Figure 72 evidences several substantial differences between the DRILCOM and PALAT systems, while schematics of each system appear in Figures 56 and 58, respectively.

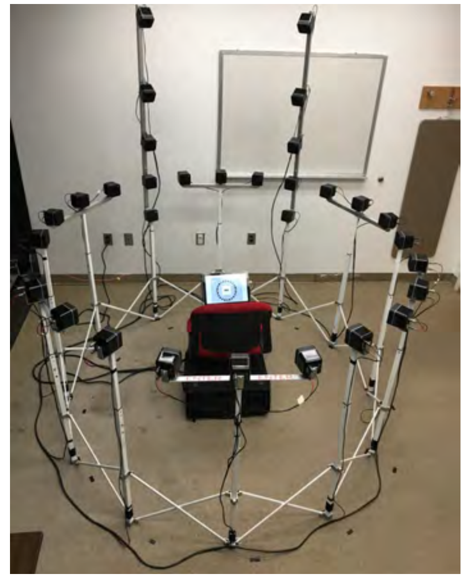


Figure 72. DRILCOM system (left) in large hemi-anechoic room and PALAT system (right) in small semi-reverberant office room on the fifth floor of Whittemore Hall on the campus of Virginia Tech.

The DRILCOM system training and test program is initiated and monitored via an investigator control station consisting of a Dell desktop computer, monitor, keyboard, and mouse. The investigator control station is situated on a small desk just outside of the loudspeaker array behind the participant, behind the 6 o'clock loudspeaker position (Figure 56). An additional computer monitor and mouse connected to the desktop computer is located directly in front of the participant on a small table inside of the loudspeaker array. A LabVIEW™ software program containing the training protocol developed in Phase I of the overarching investigation is used to train and test auditory localization. Once the LabVIEW™ program is started, the investigator hands off control of the mouse and screen to the participant seated in the middle of the loudspeaker array. The LabVIEW™ software output from the desktop computer is sent through a 3.5mm audio cable and USB cable to an automated 15-position switch and patchbay and then out to the desired Behritone C50A 5.25-inch powered loudspeaker. The automated switch and patchbay is located in an audio rack next to the investigator control station outside of the loudspeaker array. In addition, an Optimus 1850 compact disc player located on top of the audio rack delivers background pink-noise through a QSC CX1102™ power amplifier to four JBL SoundPower SP215-6 loudspeakers (Casali & Lee, 2016a).

The PALAT system training and test program is initiated and controlled by the participant via a Microsoft® Surface Pro. The investigator observes the participant from a desk located outside the loudspeaker array and is able to remotely monitor results written to file shared by the Surface Pro. The DRILCOM system LabVIEW™ software program was slightly modified to accommodate the touch screen interface of the tablet computer and eliminate the need for the investigator controls. However, the same auditory localization training and testing protocol developed in Phase I was used for both the DRILCOM and PALAT system. The LabVIEW™ software program output from the tablet is delivered via a 3.5mm audio cable and USB cable through a Numato 32 channel USB relay module and a Stewart Audio AV30MX-2 two channel stereo mixer amplifier and sent to the designated Cambridge Audio Minx Min 12 loudspeakers with a 2.25-inch single cone driver. The switch and amplifier are located within a small audio rack bag located under the participant's chair in the middle of the loudspeaker array (Figure 58). The small audio rack bag also contains a PNG-400 Pink Noise Generator, additional Stewart Audio AV30MX-2 amplifier, and a Cambridge Audio Minx Min 12 loudspeakers that is used to provide background pink-noise. Table 20 further compares features of the two training systems as used in Phase II of this investigation.

Table 20. DRILCOM (left column) and PALAT (right column) apparatus comparison.

DRILCOM		PALAT
- Permanent lab setup - 3-meter diameter - Stationary, rigid frame - 12-azimuthal loudspeakers - 6-elevation loudspeakers - 5.25-inch powered loudspeakers - Hidden loudspeakers - Pink-noise from compact disc player - Four pink-noise background loudspeakers located outside array - LabVIEW™ software program	Apparatus	- Transportable field setup - 2-meter diameter - Portable, expandable frame - 24-azimuthal loudspeakers - 10-elevation loudspeakers - 2.25-inch loudspeakers - Visible loudspeakers - Pink-noise generator - One pink-noise loudspeaker under participant chair - LabVIEW™ software program
- Hemi-anechoic room - Acoustically precise environment	Environment	- Semi-reverberant room - Standard office environment
- Investigator-controlled - Computer mouse interface - Desktop computer with participant monitor and mouse	Controls	- Participant-controlled - Stylus touch screen interface - Tablet computer
- ≤ $32000	System Approximate Setup Cost	- ≤ $16000

3.5 Phase II Experimental Procedure

The Phase II experimental procedure involved a recruitment and screening phase followed by six training and testing sessions for each participant. Each of the six sessions followed an identical training and testing procedure from Learning Unit 0 (LU0), familiarization and pretest, to LU5. Six sessions were necessary for each participant to train and test auditory localization under all three listening conditions (open ear, in-the-ear TCAPS, and over-the-ear TCAPS) on both the DRILCOM and PALAT system. The order of listening condition and training system were counterbalanced using two identical 3 x 6 Latin squares resulting in two participants for every order. The participant training system order was maintained for each listening condition throughout the study, meaning half the participants always started the

DRILCOM system for each listening condition and half the participants started on the PALAT system for each listening condition. In addition, each participant completed a listening condition on both training systems prior to switching listening conditions. The following sections detail the experimental procedures for Phase II.

3.5.1 Recruitment and screening

The investigator advertised the Phase II study via posted flyers on the Virginia Tech campus (Appendix H), emails to Virginia Tech graduate listserv, and word of mouth. Participants were asked to contact the investigator through email and a screening session was scheduled. Prior to the screening session, potential participants were emailed a copy of the Phase II informed consent and notified they would receive a hard copy at their screening session (Appendix E). Upon arrival at the screening session, the participant was provided two copies of the informed consent, one to keep and one to review and sign if willing to participate. The investigator reviewed the informed consent with the participant and briefed them on the details of the study. The participant was then provided as much time as needed to review the informed consent and to decide to participate in the study. After agreeing to participate and signing the informed consent, the participant was administered an otoscopic inspection and an audiogram by an U.S. Army Audiologist (discussed in section 3.3.3). The participant was then asked two questions to ensure the participant had no previous experience using military, law enforcement, or industrial Hearing Protection Devices or TCAPS which have a pass-through communication feature and no prior experience with auditory localization training or testing (Appendix G). Upon successful screening, the participant was scheduled to begin the first of six training and testing sessions.

3.5.2 Calibration and setup

The investigator calibrated both the DRILCOM and PALAT systems on a daily basis during the Phase II investigation. Calibration sound pressure levels were measured using a Larson-Davis Model 2900 spectrum analyzer (SN: A0280) with a ½-inch Larson-Davis 2559 microphone (SN:2575) and a Larson-Davis 9000C Preamp (SN: 0521). The measurement microphone was calibrated at 94.0 dBA (1000 Hz tone) using a Quest QC-20 Calibrator (SN: QOA070051). The microphone was then placed in the center of the DRILCOM and PALAT system array. All 12 loudspeakers used during training and testing were then calibrated to 70 dBA within a 1.5 dBA range of accuracy by adjusting the rotary dial on the LabVIEW™ program calibration screen (Figure 73). The DRILCOM system's active loudspeakers allowed for individual, manual volume adjustments due to each Behritone C50A loudspeaker containing a Class D 30-Watt amplifier (Behringer, 2012). The PALAT system's Cambridge Audio Minx Min 12 loudspeakers volume levels were centrally controlled by the Surface Pro laptop volume through a single Stewart Audio AV30MX-2 two channel stereo mixer amplifier. Figure 73 displays the LabVIEW™ calibration control screens used to calibrate the 12 loudspeakers used during training and testing for both the DRILCOM system with 12 loudspeakers positions and the PALAT systems with 24 loudspeakers positions.

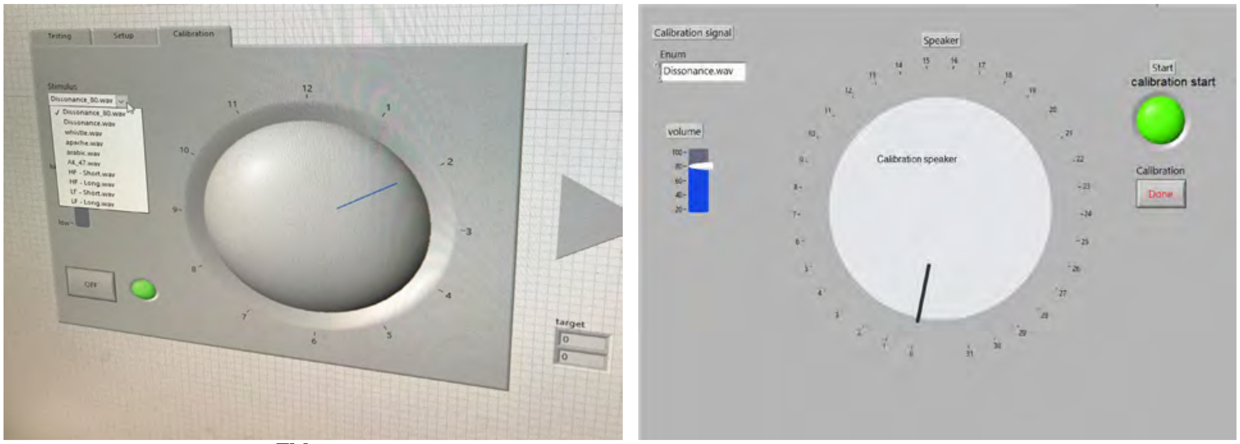


Figure 73. LabVIEW™ calibration screen for DRILCOM system (left) and PALAT system (right).

Prior to the participant's arrival, the investigator entered the participant identification number, learning unit number, auditory signal sound file, and listening condition into the LabVIEW™ software program on the assigned training system (Figure 74). The LabVIEW™ software program saved and recorded all learning unit test data to a comma separated value file stored on the DRILCOM system desktop computer or a file folder on the PALAT system that was shared via Dropbox™. Results stored included participant number, training system, listening condition, learning unit, date and time, signal source location, response location, and response time. The LabVIEW™ software program also calculated and summed the number of absolute correct responses (response location exactly matched the signal source location) and ballpark correct responses (response location was within ±15-degrees of the signal source location). As a backup data source, the investigator manually recorded the number of absolute correct responses and number of ballpark absolute correct responses for every test on a participant score sheet.

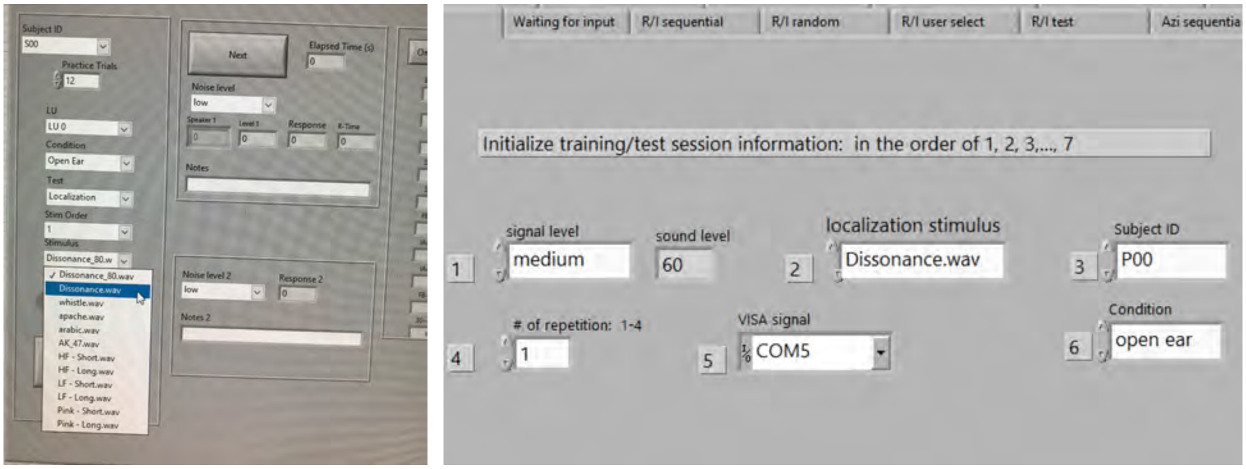


Figure 74. DRILCOM system (left) and PALAT system (right) initialization screens.

3.5.3 Training and testing sessions

The following sections detail the experimental procedures for the DRILCOM and PALAT systems. The PALAT system was designed to perform the same functions as the DRILCOM system and as a result the procedures are very similar. However, a few updates were made to the LabVIEW™ software program to accommodate for the tablet touch screen interface and to allow for a participant or user-controlled training and testing program as opposed to the investigator-controlled DRILCOM system. One experimental procedure is described below with differences between the DRILCOM and PALAT system highlighted.

Upon arrival for the first session, the participant was given a tour and overview of both training systems. The purpose of the study was read aloud to the participant. Next the investigator provided a more thorough familiarization to the training session that the participant would be using during that session using a script that appears in Appendix I. The investigator demonstrated how to use the DRILCOM system participant monitor screen and mouse and the PALAT system tablet computer interface using both visual aids and the equipment. The participant was then given a chance to operate the user controls. When the participant was comfortable with operating the training system user interface and software program, the investigator set the participant up with the correct listening condition. For the two TCAPS listening conditions, the investigator ensured the TCAPS were turned on and set to the unity gain prior to fitting the participant. Then, the investigator fitted the participant with their assigned TCAPS devise to ensure consistency of proper fit. The investigator verified with the participant that the TCAPS device was working and that fit was comfortable. The participant was instructed to notify the investigator of any discomfort or change in the TCAPS device fit or function. Next, the investigator aligned the participant so that their head was centered within the loudspeaker array. The investigator visually inspected to ensure the participant's head was in line with the 12 and 6 o'clock loudspeakers and their ears were in line with the 3 and 9 o'clock loudspeakers. The participant was informed that they were free to turn their head to aid in locating the auditory signal during the training and testing but were instructed to return their head to the forward facing position prior to each signal presentation. Studies have found that head movements improve auditory localization by creating momentary changes in interaural level differences and interaural timing differences (Thurlow & Mergener, 1970; Muller & Bovet, 1999). The investigator sat outside of the loudspeaker array and monitored the participant's head location during each learning unit.

Each of the six sessions began with Learning Unit 0 which consisted of a familiarization sequence followed by the first test, or pretest, to establish a baseline auditory localization score. For the DRILCOM system, the investigator selected LU0 on the software program and then gave control of the mouse to the participant on the monitor within the center of the array. For the PALAT system, the participant was instructed to select LU0 under the horizontal sequential training menu. The investigator then briefed the participant on the sequential order of the familiarization unit. The familiarization sequence consisted of four presentations of the dissonant signal from the 12, 3, 6, and 9 o'clock loudspeaker positions. The participant was informed to press the "Click to Signal" (DRILCOM) or "Click to Sound" (PALAT) green button located in the center of the screen when ready and to listen for the dissonant signal presented from the 12 o'clock position (Figure 75). Following the dissonant signal, the investigator informed the participant to select the response button representing the 12 o'clock loudspeaker location (Figure 75). A response was entered by using the mouse to direct the pointer over the button and left clicking the mouse on the DRILCOM system and by touching the button with the stylus pen on the PALAT system. Localization performance feedback was displayed in the white box in the center of the screen. A correct answer was indicated by two overlapping lines, designating the signal location and response location, pointing to the signal loudspeaker location (Figure 76). Two text boxes were used to display the signal location corresponding clock face number (left text box) and response location corresponding clock face number (right text box). An absolute correct response resulted in the same clock face number in both signal and response location. In addition to the overlapping lines, a square colored text box with the "CORRECT" was displayed on the screen (Figure 76). The investigator then informed the participant to repeat the signal and response procedure for the 3 o'clock and 6 o'clock dissonant signal presentations. Finally, the investigator directed the participant to initiate the dissonant signal presentation from the loudspeaker at the 9 o'clock position but to respond with an incorrect response. This served to demonstrate the system feedback for an incorrect response. Figure 77 displays the incorrect feedback indicated by a dotted blue line pointing to the signal presentation location and a solid red line pointing to the response location. Two text boxes were used to display the signal location corresponding clock face number (left text box) and response location corresponding clock face number (right text box). An incorrect response resulted in different clock face numbers in the signal and response location. In addition, a square text box is displayed with the work "WRONG!" for the DRILCOM system and "MISSED" for the PALAT system (Figure 77).

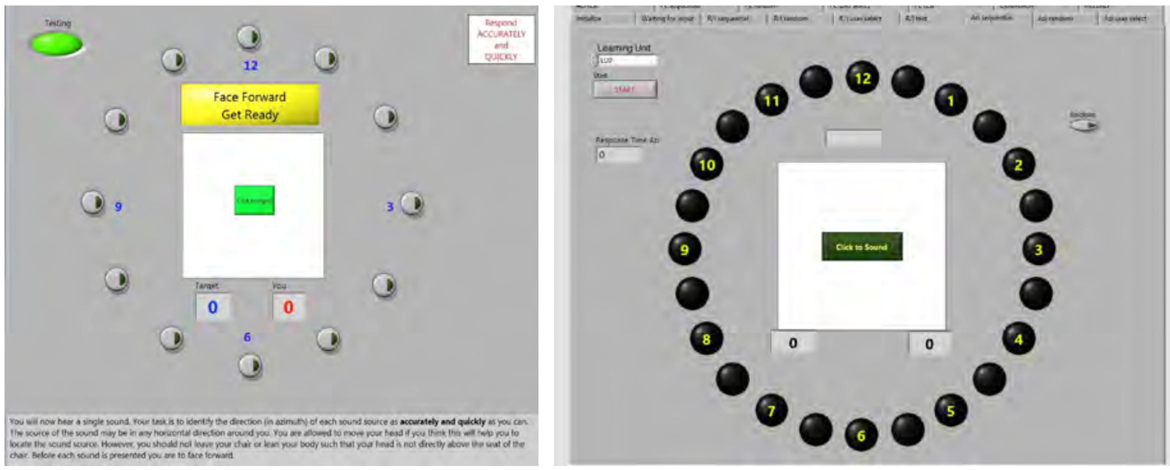


Figure 75. DRILCOM system (left) and PALAT system (right) training screen displaying initiation of training trial.

Figure 76. DRILCOM system (left) and PALAT system (right) training screen displaying feedback for absolute correct response.

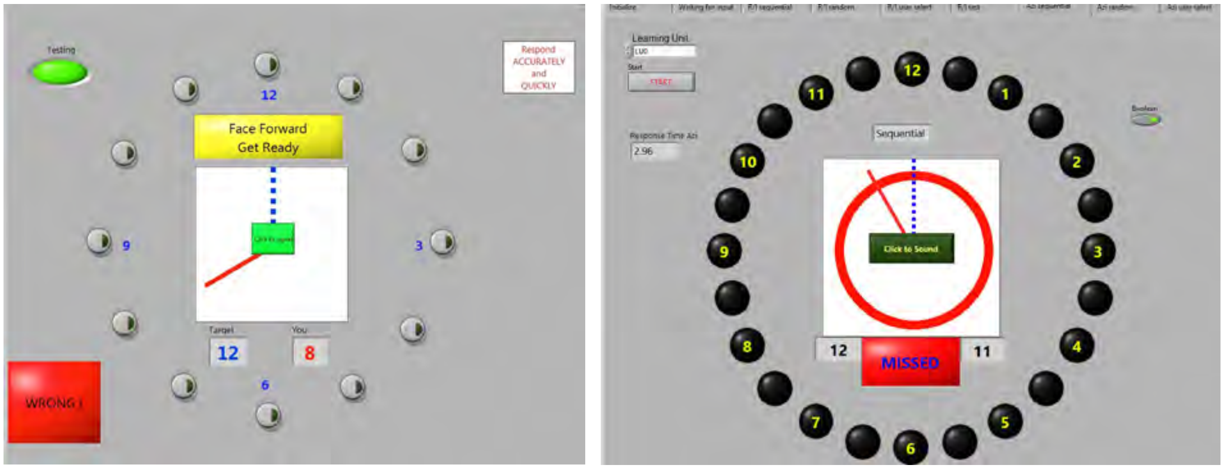


Figure 77. DRILCOM system (left) and PALAT system (right) training screen displaying feedback for incorrect response.

Upon completion of the familiarization sequence, the investigator ensured the participant was comfortable operating the training system software program and understood the auditory localization task. If there were any questions, the investigator demonstrated how to use the equipment until the participant was comfortable. However, the participant was not allowed to repeat the familiarization sequence or listen to additional signal presentations in order to maintain consistency for the baseline localization test.

Following familiarization, the participant was administered the LU0 localization test, which comprised the pretest for the training experiment. For the DRILCOM system, the investigator started the localization test software program and entered the participant number, listening condition, and learning unit into the system before giving control of the mouse to the participant. For the PALAT system, the investigator informed the participant to select the “H test” button, representing a horizontal test, from the main menu and set the learning unit to “LU0” which would initiate the PALAT system localization testing screen. Figure 78 displays the auditory localization testing screens for both the DRILCOM and PALAT systems. For consistency and ease of use, the testing screen used the same layout as the training screen with the exception of the white feedback window and text boxes. The participant was reminded from the system overview that a localization test consisted of 36 random presentations, or trials, with three signal presentations being played from each of the 12 numbered loudspeakers, or 12 corresponding clock face locations. The investigator instructed the participant to “respond as accurately and as quickly as possible,” by selecting one of the 24 buttons representing the 12 active loudspeaker locations or 12 inactive (dummy) loudspeaker locations. The participant was instructed to press the “Click to Signal” (DRILCOM) or “Click to START” (PALAT) green button located in the center of the screen when ready to begin the localization test. Each time the participant selected the “Click to Signal” (DRILCOM) or “Click to START” (PALAT) button, a dissonant signal was emitted from one of the 12 loudspeakers located at the 12 clock face positions. Upon selecting one of the 24 response buttons, the system recorded the signal location, response location, and response time and reset to allow for the next signal presentation. At the completion of the 36 presentations and participant responses, the training system software informed the participant that the test was completed and the system returned to the main screen.

Figure 78. DRILCOM system (left) and PALAT system (right) localization testing screen displaying initiation of testing trial.

Following the completion of the LU0 localization test, the participant began the auditory localization training protocol originally designed by Lee and Casali (2017) and modified during the Phase I investigation of the overarching experiment by K. Cave (Cave et al., 2019). The auditory training protocol consisted of five Learning Units (LUs) given in sequential order over the course of about 1.5 hours per participant. A participant was allowed to take breaks when needed but was encouraged to wait until completing a Learning Unit. The localization training protocol LU consisted of four distinct subunits.

1) Sequential – The sequential subunit consisted of dissonant signal presentations in a circular pattern. Every time the participant initiated the dissonant signal, the software program simultaneously presented the auditory signal and indicated which loudspeaker position the sound originated. Figure 76 and 77 display the sequential training screen with feedback for an absolute correct response and an incorrect response, respectively. LU1 consisted of four “laps” of sequential presentations of the dissonant tone around the 12-loudspeaker locations for a total of 48 presentations. The sequential order was as follows:

a) started at 12 o’clock and moved clockwise through 11 o’clock,
b) started at 9 o’clock and moved counterclockwise through 10 o’clock
c) started at 3 o’clock and moved clockwise through 2 o’clock, and
d) started at 6 o’clock and moved counterclockwise through 5 o’clock

LUs 2-5 consisted of only one “lap” around the 12 loudspeaker locations with a randomly assigned starting location.

2) Random – Following the sequential subunit, the investigator opened the random training software file for the DRILCOM system. The PALAT system automatically transitioned from sequential to random indicated by the textbox above the feedback loop. During the random training subunit, the participant was not informed of the location of the signal but was provided feedback after their response was entered. The random training subunit consisted of three signal presentations from each of the 12 loudspeaker positions at random for a total of 36 presentations.

3) User-select – The third subunit, named user-select, allowed the participant to choose 18 signal presentations from any of the 12 loudspeaker locations based on where the participant needed to practice. The DRILCOM system user select software program screen looked identical to the sequential and random screens. The PALAT system user select screen only displayed 12 numbered buttons representing the 12 clock face loudspeaker locations (Figure 79).

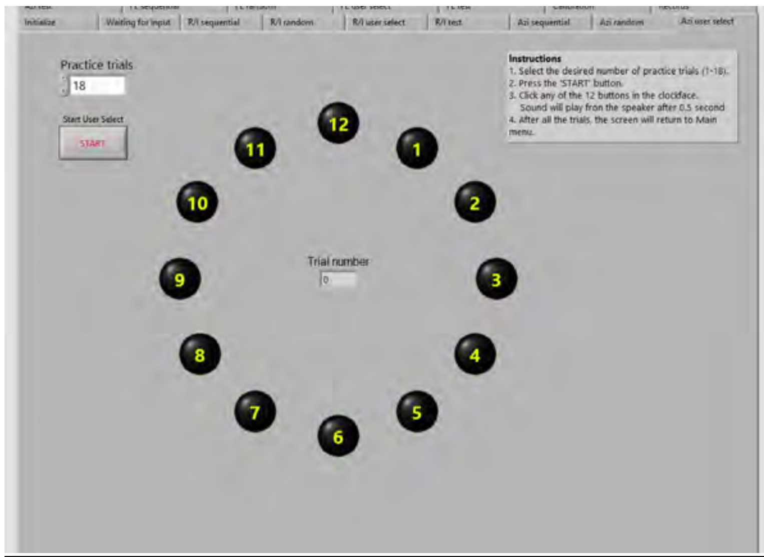


Figure 79. PALAT system user-select training screen displaying loudspeaker locations used for practice trials.

4) Localization Test – Following the user-select subunit, an auditory localization test was administered. The localization test was identical to the test taken in LU0. The investigator opened the localization test software program on the DRILCOM system and gave control of the mouse to the participant. The PALAT system returned to the main menu following the user-select training and the participant had to select the “H test” button and input the LU number. The investigator was present to ensure the participant selected the correct program and parameters. Each localization test consisted of three dissonant signal presentations from all 12 clock face loudspeakers in random order for a total of 36 presentations. During the localization test, the participant was not provided with the location of the signal and was not informed if their answer was accurate or incorrect.

At the end of each LU, the participant was given the option to take a break or keep training. Breaks were not controlled for time but the participant was asked to limit a break to around 10 minutes. Upon completion of LU5, the participant was asked to complete a questionnaire (discussed in Dependent Measures and included in Appendix F). At the end of each session, the investigator scheduled the next session. The participant was required to wait a minimum of three hours between each session and had to complete all six sessions within a two-week period.

3.6 Phase II Results

Phase II data reduction and analyses were performed using Microsoft® Excel, IBM® SPSS® Statistics and JMP® 14 software. The PALAT and DRILCOM systems' LabVIEW™ software automatically recorded the auditory signal loudspeaker location, the participant response loudspeaker location, and the response time for each trial into a comma separated value (CSV) file. The data contained in the CSV file were exported to an Excel file. As an oversight check, the sum of the absolute correct responses for each Learning Unit (LU) test by training system and listening condition recorded by the LabVIEW™ software was verified by the investigator. The investigator calculated the mean absolute correct response score, termed ***absolute score***, by training system and listening condition for each LU test using Excel. SPSS® Statistics was used to fit the absolute scores from LU0, pretest, through LU5, posttest, for each participant to a simple linear regression line to predict future training effects. The slopes of each participant's learning rate were calculated in Excel and compared using SPSS® Statistics to analyze training effects between training systems under each listening condition. The mean slopes and regression equations by training system and listening condition were analyzed and reported to compare the training effects on the dependent measures. Individual participant slopes and regression lines were not compared or reported since this was a within-subjects design and performance differences between participants was not evaluated. A percent accuracy score was calculated for each loudspeaker location by training system and listening condition for each LU test. SPSS® Statistics was used to conduct a correlation analysis between percent accuracy by loudspeaker location. In similar fashion to the dependent measure ***absolute score***, the investigator calculated the sum of ***front-back reversal errors*** and mean ***response times*** for each LU test by training system and listening condition using Excel and SPSS® Statistics was used for analyses. Subjective data from participant questionnaires were manually entered into an Excel file and imported to IBM® SPSS® Statistics and JMP® 14 software for analyses. Prior to conducting statistical analysis, the investigator verified there was no missing data and checked for outliers.

3.6.1 Outlier Analysis

A Dixon *Q*-test was performed on all dependent measures to identify outliers. The outlier analysis was performed separately for both objective, quantitative dependent measures (two accuracy measures and response time) for each listening condition for LU0, pretest, and LU5, posttest. The objective of Phase II was to compare the training effects of the DRILCOM system and PALAT system. Each of the 12 participants trained and tested on both training systems. The resulting sample size for each Dixon *Q*-test was n=24, 12 tests on the DRILCOM system and 12 tests on the PALAT system. To perform the Dixon *Q*-test, the subset of data for each test was arranged sequentially from lowest to highest value. A Dixon *Q*-test was then performed manually using one of the following formulae:

(2) $$Q = \frac{|x_n - x_{n-1}|}{|x_n - x_1|} \quad \text{or} \quad Q = \frac{|x_2 - x_1|}{|x_n - x_1|}$$

where n is the sample size and the x represents the ordered values from lowest to highest, $x_1 < x_2 < \ldots < x_n$ (Dixon, 1951). The numerator in equation (2) represents the gap between the two highest values, $|x_n - x_{n-1}|$, or the gap between the two lowest values, $|x_2 - x_1|$. The denominator in equation (2) represents the range of the data, $|x_n - x_1|$. The equation that resulted in the largest gap was used to identify the existence of a single outlier for each data subset. The calculated Q-value for each data subset was compared to Dixon's r_{10} table for n=24 using a 95% confidence interval (Dixon, 1951). If $Q \geq 0.34$, then an outlier was deemed present. No outliers were found using the Dixon Q-test on any of the three objective measures (Table 21).

Table 21. Sample statistic Q for dependent measures at LU0 and LU5 by listening condition using 95% confidence interval. Dixon's r_{10} critical statistic for n=24 is Q=0.34.

	Open ear		**TEP-100**		**ComTac™ III**	
Dependent Measure	LU0 (pretest)	LU5 (posttest)	LU0 (pretest)	LU5 (posttest)	LU0 (pretest)	LU5 (posttest)
Absolute correct score	0.33	0.14	0.19	0.04	0.12	0.11
Front-back errors	0.25	0.00	0.14	0.11	0.18	0.09
Response time	0.20	0.03	0.10	0.22	0.16	0.19

3.6.2 Phase II Objective Measures Overview and Data Graphs for Initial Visual Inspection

The primary objective of the Phase II in-laboratory investigation was to evaluate and validate the effectiveness of the PALAT system compared to the DRILCOM system in auditory localization skills assessment and training. Phase II objective data analyses focused on three dependent measures, *absolute score*, *front-back reversal errors*, and *response time*. Prior to performing statistical analysis, to enable visual inspection, the investigator plotted the mean *absolute score* of both the DRILCOM system and PALAT system under each listening condition. Figure 80 displays several similar trends under both the DRILCOM and PALAT system. The participants' pretest absolute scores at LU0 for each listening condition were similar while using the DRILCOM and PALAT system. In addition, the posttest absolute scores at LU5 ended in very close values under both the DRILCOM and PALAT systems. While the learning rate lagged very slightly under the PALAT system in LU1 and LU2, participants tended to learn at a higher rate on the PALAT system at LU3 and LU4. Overall, the learning rates from LU0 to LU5 seemed similar from this visual inspection. Participants experienced an 11% improvement of mean absolute score from LU0 to LU5 when using both the DRILCOM and PALAT training systems in the open ear listening condition, 11% and 19% improvement of mean absolute score when using DRILCOM and PALAT respectively in the ComTac™ III listening condition, and 20% and 19% improvement of mean absolute correct response when using DRILCOM and PALAT respectively in the TEP-100 listening condition. From these percentage improvements in absolute accuracy across all LU's, it was evident that the PALAT's training benefit was approximately equal to, or in one instance (ComTac™ III) much better, than that of the DRILCOM system.

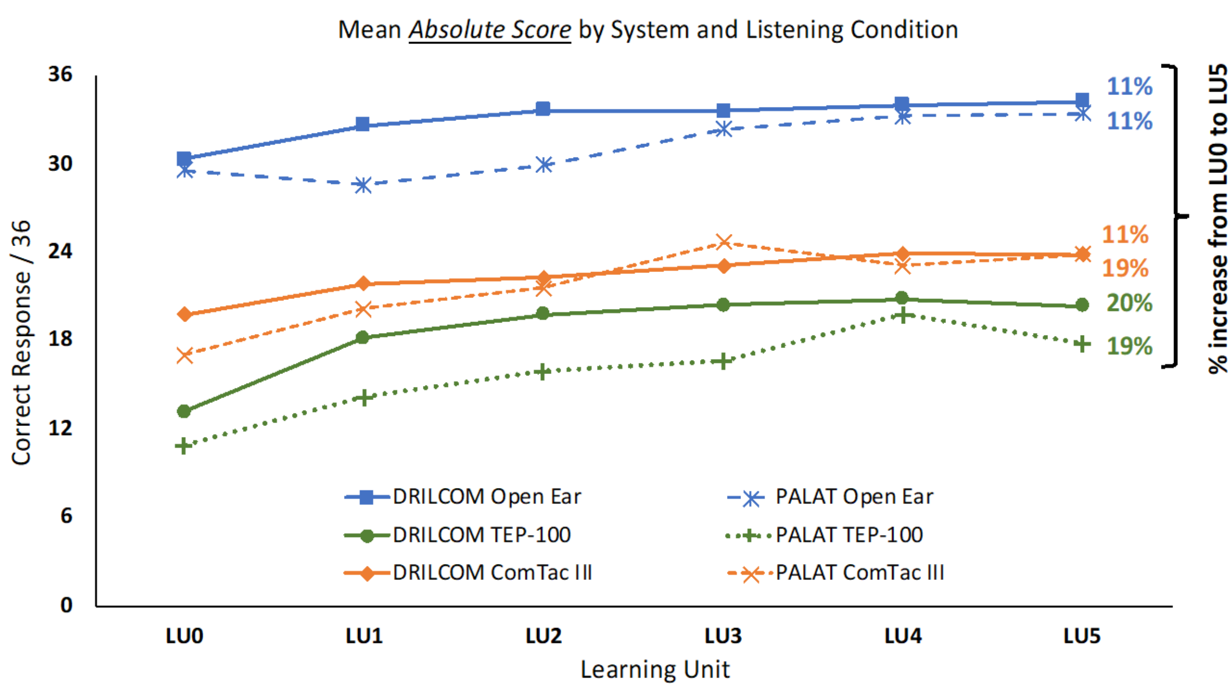


Figure 80. Mean *absolute scores* across all participants by training system and listening condition. Percent improvement from LU0 to LU5 displayed (right).

Another trend evident in the Figure 80 is the ability of both the DRILCOM and PALAT system to assess localization performance under different listening conditions. One of the objectives of the PALAT system was to demonstrate the capability of accurately differentiating localization performance under different listening conditions. As expected, participant performance was highest on both training systems under the open ear listening condition. Surprisingly, under both the DRILCOM and PALAT systems, the mean absolute score under the ComTac™ III earmuff style TCAPS was higher than the TEP-100 earplug style TCAPS. This was deemed atypical because earmuff-style (over-the-ear) devices often suffer in localization performance due to the loss of pinna effects and monaural localization cues, cues which are indeed physically present with in-the-ear devices.

Similar trends discovered for *absolute score* were also observed in the mean *front-back reversal error* plots for the overall number of errors and errors under each listening condition. Figure 81 shows a slightly higher mean number of front-back reversal errors at LU0 between the two training systems but the same number of mean front-back reversal errors at LU5. Participants made the fewest front-back reversal errors under the open ear condition and the highest front-back reversal errors under the TEP-100 condition (Figures 82, 83, and 84). Front-back reversal errors were slightly higher while using the PALAT system under the TEP-100 condition at LU0 but were slightly lower under the ComTac™ III at LU5. Figures 81 through 84 display the general trends of front-back reversal errors by listening condition for both training systems. Statistical analyses were performed to identify if the difference in means were significant and will be presented and discussed later in the results section.

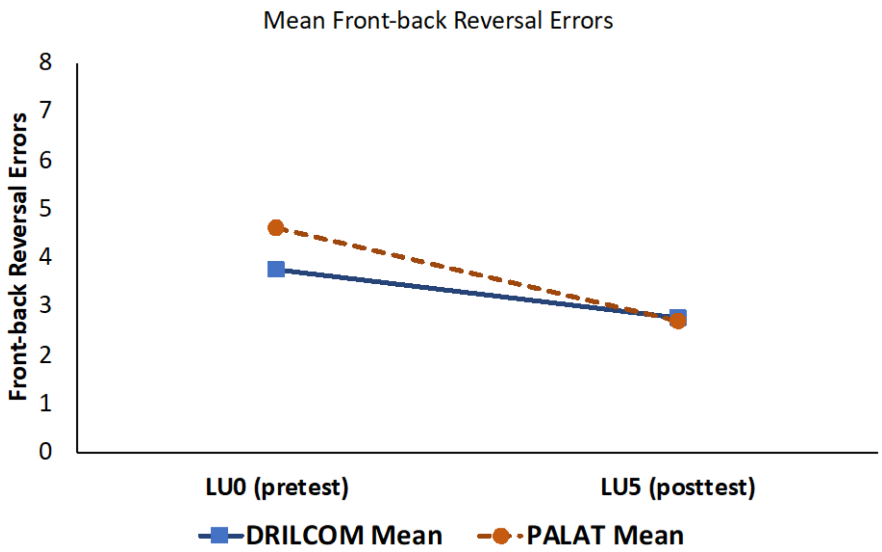


Figure 81. Mean overall *front-back reversal errors* for all listening conditions at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line).

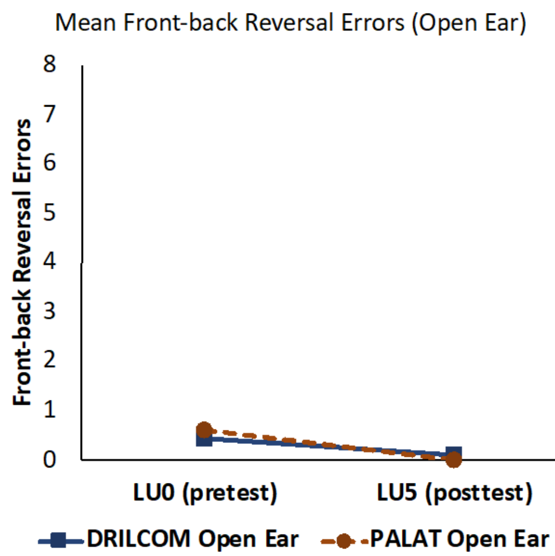


Figure 82. Mean *front-back reversal errors* for the open ear listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line).

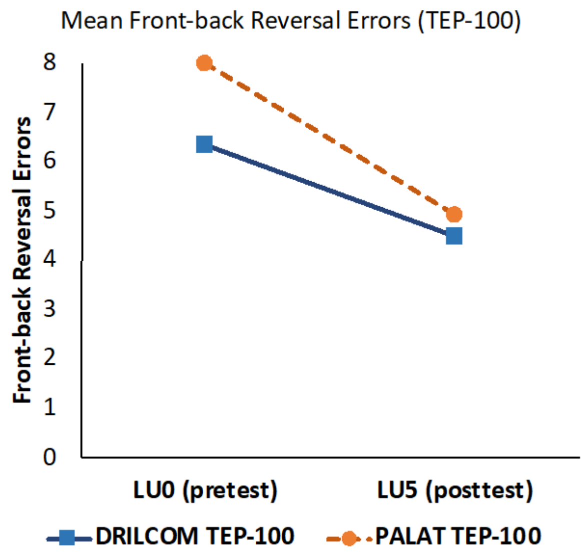


Figure 83. Mean *front-back reversal errors* for the TEP-100 listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line).

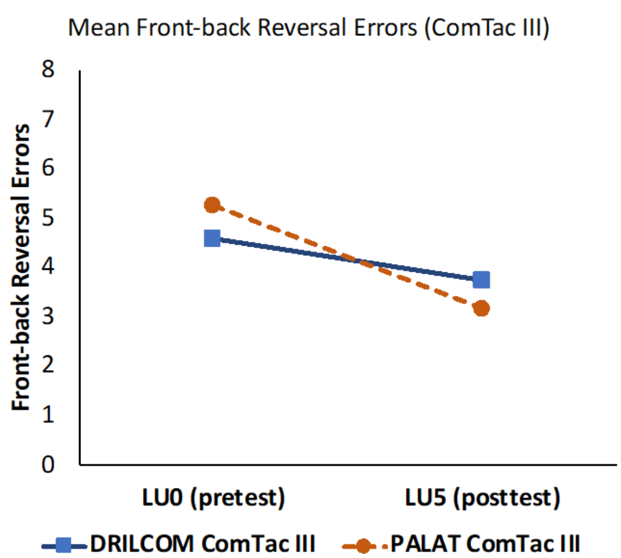


Figure 84. Mean *front-back reversal errors* for the ComTac™ III listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line).

Unlike *absolute score* and *front-back reversal errors*, the PALAT system was specifically designed to improve upon the user interface by reducing the DRILCOM equipment consisting of a desktop computer, two monitors, keyboard, and mouse to a tablet and stylus.

Based on the improvements made to the user interface, response times while using the PALAT system were expected to be quicker compared to the same task while using the DRILCOM system. The expected difference in response times would then be more attributable to an easier user interface on the PALAT system and not a result of training effects of the PALAT system compared to the DRILCOM system. One of the Human Factors design principles of response selection is the location compatibility which states that the control location should be close to the item being controlled or the display of the item being controlled (Wickens, Lee, Liu, & Becker, 2004). The *direct position control* of the tablet and stylus allow the user to physically touch the desired response location without having to locate a cursor on the screen or possibly shift their visual attention to the *indirect position control* of the computer mouse (Wickens, Lee, Liu, & Becker, 2004). Based on these principles, the investigator expected a quicker response time while using the PALAT system because the touchscreen controls were placed at the location of the item being controlled and the stylus pen user interface (direct position control) reduced potential lag time of the computer mouse interface. Response time was measured automatically by the LabVIEW™ software beginning at the onset of dissonant tone signal and ending when the participant selected the response loudspeaker location. The mean response time plots in Figures 85, 86, 87, and 88 support the expected results for overall response time as well as response time by listening condition. Statistical analyses were performed to confirm observed results seen in graphs and to compare training effects using each training system and will be presented and discussed further in the results section.

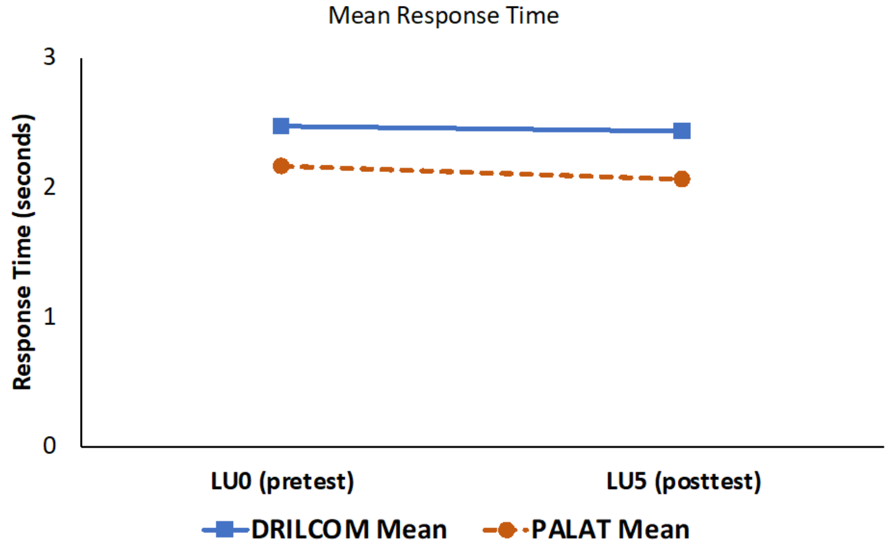


Figure 85. Mean overall *response time* for all listening conditions at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line).

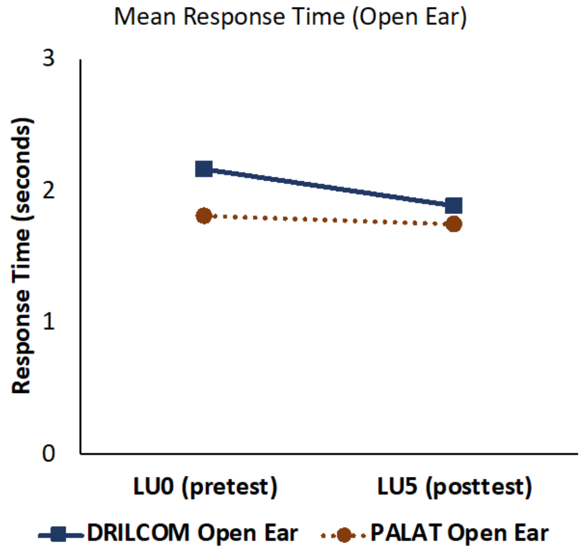


Figure 86. Mean *response time* for the open ear listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line).

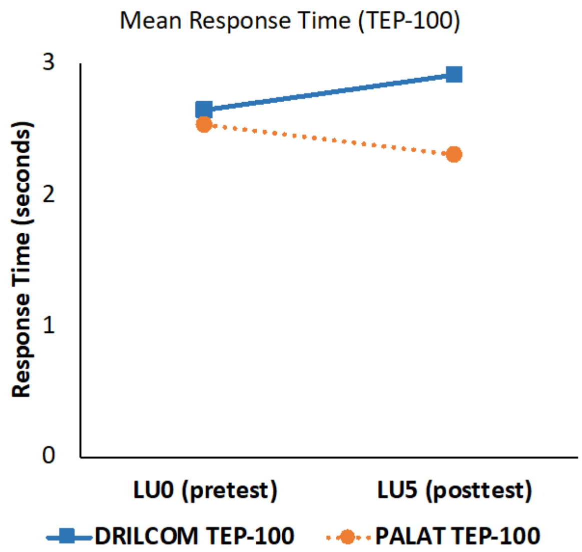


Figure 87. Mean *response time* for the TEP-100 listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line).

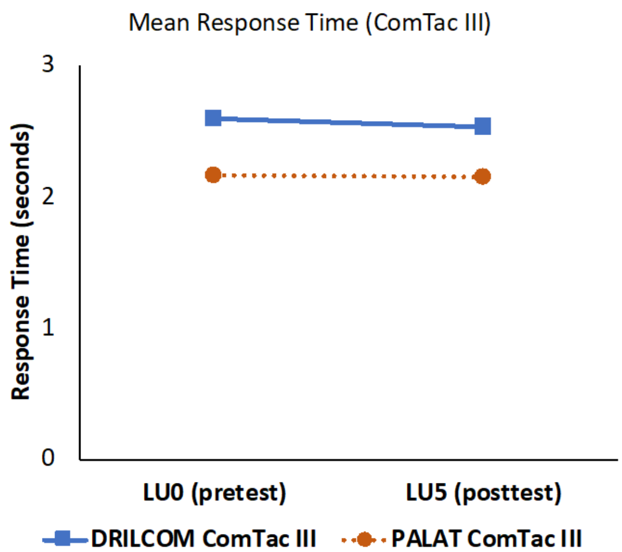


Figure 88. Mean *response time* for the ComTac™ III listening condition at LU0 and LU5 by training system, DRILCOM (blue square markers and solid line) and PALAT (orange circle markers and dashed line).

3.6.3 Phase II Objective Measures Statistical Analyses

To fully evaluate the PALAT system, the investigator employed statistical analyses to compare localization performance at various stages of training to detect differences between the training systems. Mean ***absolute score***, ***front-back reversal errors***, and ***response time*** were compared for each training system and listening condition at LU0 (pretest), LU5 (posttest), and for a training effect from LU0 to LU5. All three dependent measures at various stages of training were evaluated using a full factorial, within-subjects ANOVA, or repeated-measures ANOVA, to test for differences among mean values of the dependent measure. Statistical analyses were performed using Excel v16.16.10, JMP® 14 software, and IBM® SPSS® Statistics. All statistical analyses used a significance level of $\alpha=0.05$ to control for Type I errors and a power of $1-\beta=0.8$ to control for Type II errors. For each ANOVA, a Mauchly's Test of Sphericity was performed on all factors containing more than two levels to evaluate the assumption of homogeneity of variances (Keppel & Wickens, 2004). A violation of Mauchly's Test of Sphericity, denoted by a significant *p*-value, could result in an increase in Type I error rate (Keppel & Wickens, 2004). When Mauchly's test was significant, reductions were made to the degrees of freedom using the Greenhouse-Geisser estimate or Huynh-Feldt estimate to obtain a more conservative *F*-ratio (Dunn & Clark, 1987).

All significant ANOVA main effect findings were followed by post hoc testing using pairwise comparisons for each dependent measure factor level. Given the full-factorial, or completely within-subjects, investigation, paired-samples *t*-tests with Bonferroni correction were performed to compare means at each level of the significant main effect factor. The Bonferroni correction for multiple comparisons reduced the significance level by dividing the alpha by the number of comparisons to reduce the risk of making a Type I error (Keppel & Wickens, 2004). The listening condition factor had three levels so a Bonferroni adjusted $\alpha=0.05/3$, or $\alpha=0.017$, was used to test a significant difference between listening condition means. SPSS® Statistics software adjusts the *p*-value by multiplying by the number of comparisons so that the original α=0.05 significant level can be used to check for significant differences. All tabled results requiring Bonferroni adjustments for Phase II are displayed with a Bonferroni adjusted *p*-value.

In addition to ANOVAs, simple linear regression fitting, correlation, and graphical analyses were used to evaluate localization performance between independent measures. The absolute scores from LU0, pretest, through LU5, posttest, for each participant were fitted to a simple linear regression line to predict future training effects by training system and listening condition. The slopes of each participant's learning rate were compared to analyze training effects between training systems under each listening condition. A percent accuracy score was calculated for each loudspeaker location by training system and listening condition for each LU test followed by a Spearman's rank correlation to identify if there was a relationship between the percent accuracy at each loudspeaker location between the DRILCOM and PALAT systems. Lastly, mean absolute score by loudspeaker location were graphed on radial plots to allow for comparison between the DRILCOM and PALAT systems at LU0 and LU5.

The following sections detail the results of statistical analyses for ***absolute score***, ***front-back reversal errors***, and ***response time*** by training system and listening condition at various stages of training. Analyses for each dependent measure are presented in a stand-alone section in the order listed above.

Absolute Correct Response Score Analyses

Performance at LU0 (pretest) on Absolute Score

A full factorial repeated-measures Analysis of Variance (ANOVA) was used to analyze the effect of training system (DRILCOM and PALAT) and listening condition (open ear, TEP-100, and ComTac™ III) on the dependent measure *absolute score* at LU0 (pretest). A Mauchly's Test of Sphericity was performed for all independent variables with more than two levels since this was a full factorial within-subjects investigation. The Mauchly's Test of Sphericity resulted in no violations of homogeneity of variances of the differences between all pairs of within-subjects conditions (Table 22). The **main effect for training system was not significant**, $F(1,11)=1.68$, $p=0.22$, $\eta_p^2=0.13$. The **interaction between training system and listening condition was also not significant**, $F(2,22)=0.50$, $p=0.62$, $\eta_p^2=0.04$. For the **main effect of listening condition on absolute score at LU0, statistically significant differences existed between means**, $F(2,22)=29.08$, $p<0.001$, $\eta_p^2=0.73$ (Table 23).

Table 22. Mauchly's test of sphericity for full factorial ANOVA for the effect of training system and listening condition on ***absolute score*** at LU0 (pretest).

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.86	1.56	2	0.46	0.87	1
Training System x Listening Condition	0.94	0.63	2	0.73	0.94	1

Table 23. Full factorial ANOVA table evaluating **differences in *absolute score* at LU0 (pretest)** according to training system and listening condition (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	53.00			
Within Subjects					
Training System (A)	1	66.13	1.68	0.221	0.133
Error (A x S)	11	39.34			
Listening Condition (C)	**2**	**1981.17**	**29.08**	**<0.001***	**0.726**
Error (C x S)	22	68.14			
A x C	2	6.50	0.50	0.615	0.043
Error (A x C x S)	22	13.08			
Total	71	2227.36			

Listening Condition Main Effect: Post hoc test on Absolute Score at LU0

Pairwise comparisons were conducted for each listening condition (within the main effect of listening condition) using the measure of ***absolute score***. Post-hoc comparisons, using a paired-samples *t*-test with Bonferroni correction ($\alpha = 0.05/3$) for each listening condition, showed a **significant difference between all three listening conditions**. The SPSS® Statistics software adjusted the *p*-values allowing comparison of tabled results at α=0.05 (Table 24). The mean absolute score at LU0 for the open ear condition (*M*=29.96, *SD*=1.44) differed significantly from the TEP-100 condition (*M*=12.04, *SD*=1.44) and ComTac™ III condition (*M*=18.38, *SD*=1.95). In addition, the TEP-100 condition differed significantly from the ComTac™ III condition (Table 24). Figure 89 displays the mean absolute scores for each listening condition and error bars representing the 95% confidence intervals about the means. Mean performance for absolute score was highest under the open ear condition followed by the ComTac™ III condition and lowest under the TEP-100 condition.

Table 24. Paired-samples *t*-test pairwise **comparisons between listening conditions on *absolute score* at LU0 (pretest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**17.92**	**2.58**	**<0.001***
Open ear	**ComTac™ III**	**11.58**	**2.62**	**0.003***
TEP-100	**ComTac™ III**	**-6.33**	**1.88**	**0.019***

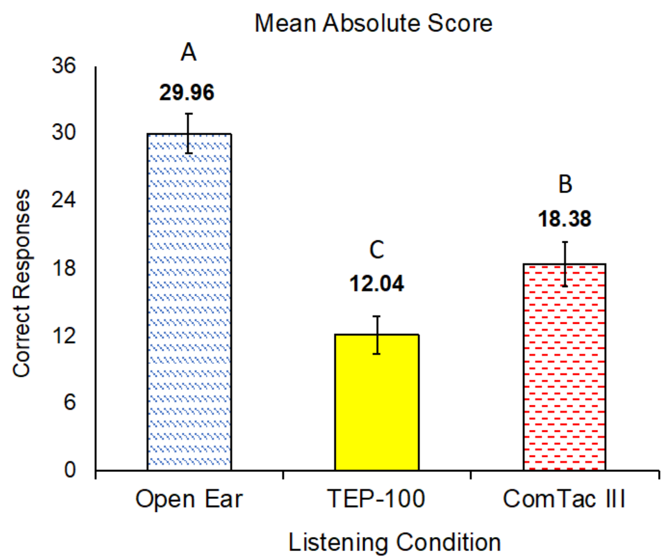


Figure 89. **Mean *absolute score* for each listening condition at LU0 (pretest)** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples t-test.

Results supported group equivalence at pretest collapsed across training system. No significant differences were found for the main effect of training system or the interaction between training system and listening condition. This indicated that the PALAT system was able to measure the participants' pretest localization ability just as well as the DRILCOM system under each listening condition. Figure 90 shows the similarity of localization performance at pretest on the two training systems by listening condition. The chart displays the mean pretest *absolute score* for each listening condition by training system with mean values given above the 95% confidence interval error bars about the mean. Similar analyses were performed on each individual training system comparing means of the three listening conditions and results are discussed later.

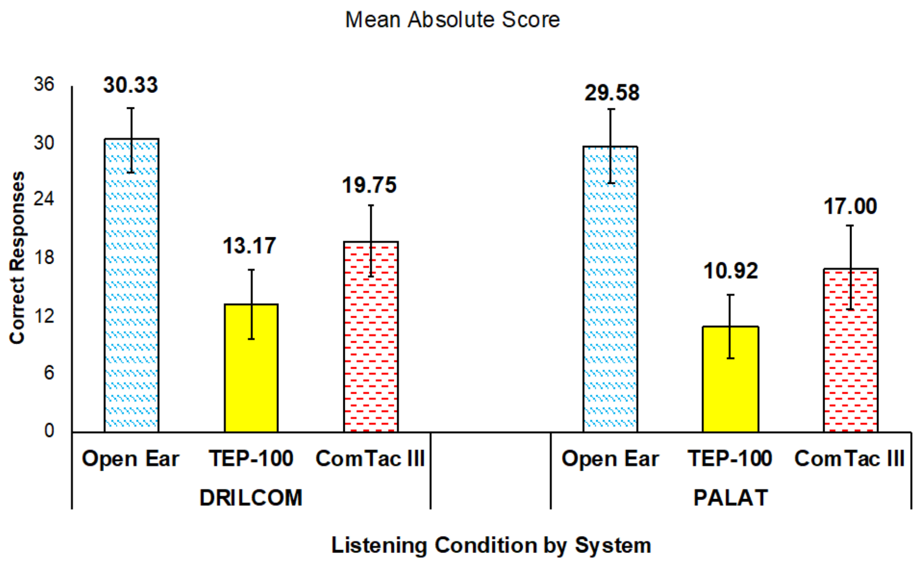


Figure 90. **Mean _absolute score_ for each listening condition by training system at LU0 (pretest)** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means.

*Listening Condition differences using the **DRILCOM** system on Absolute Score at LU0*

Figure 90 indicated that both the DRILCOM and PALAT systems were sensitive to auditory localization performance differences between listening conditions. To test this theory, a one-way repeated-measures ANOVA was conducted to analyze the effect of listening condition (open ear, TEP-100, and ComTac™ III) while using the DRILCOM system on the dependent measure *absolute score* at LU0 (pretest). A Mauchly's Test of Sphericity was performed since the listening condition factor contained three levels. The Mauchly's Test of Sphericity resulted in no violations of homogeneity of variances of the differences for the listening condition variable (Table 25). **The ANOVA resulted in a significant difference between the mean absolute score for listening conditions**, $F(2,22)=19.78$, $p<0.001$, $\eta_p^2=0.64$. (Table 26).

Post-hoc pairwise comparisons using a paired-samples *t*-test with Bonferroni correction ($\alpha = 0.05/3$) for each listening condition comparison were conducted using the measure of *absolute score*. Post-hoc comparisons showed **significant differences between the open ear**

condition (*M*=30.33, *SD*=5.88) **and the TEP-100 condition** (*M*=13.17, *SD*=6.41) **and between the open ear condition and the ComTac™ III condition** (*M*=19.75, *SD*=6.50). **No significant difference was found between TEP-100 and ComTac™ III** (Table 27). Figure 91 displays the mean absolute scores for each listening condition while using the DRILCOM system and error bars representing the 95% confidence intervals about the means. Mean performance for absolute score was highest under the open ear condition followed by the ComTac™ III condition and lowest under the TEP-100 condition.

Table 25. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the DRILCOM system on ***absolute score*** at LU0 (pretest).

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.94	0.64	2	0.73	0.94	1

Table 26. One-way ANOVA table evaluating **differences in *absolute score* at LU0 (pretest)** according to listening condition while using the DRILCOM system (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	26.86			
Within Subjects					
Listening Condition (C)	**2**	**900.08**	**19.78**	**<0.001***	**0.643**
Error (C x S)	22	45.51			
Total	35	972.45			

Table 27. Paired-samples *t*-test pairwise **comparisons between listening conditions while using the DRILCOM system on *absolute score* at LU0 (pretest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**17.17**	**2.99**	**<0.001***
Open ear	**ComTac™ III**	**10.58**	**2.84**	**0.010***
TEP-100	ComTac™ III	-6.58	2.40	0.057

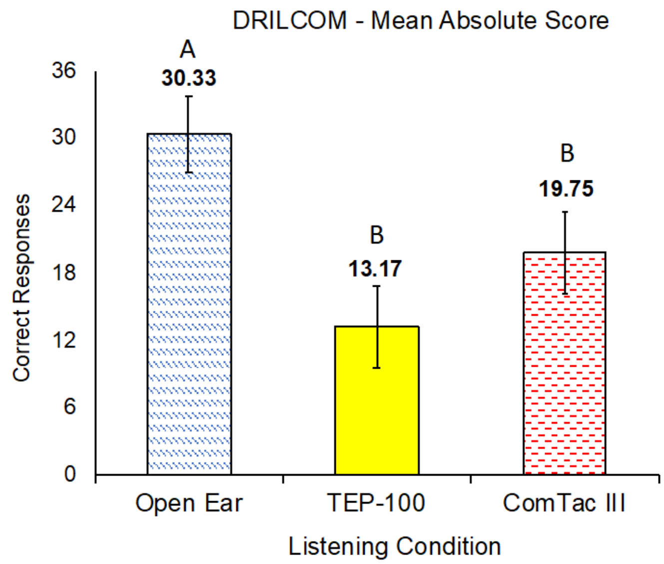


Figure 91. **Mean *absolute score* for each listening condition at LU0 (pretest) while using the DRILCOM system** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at p<0.05 per a paired-samples t-test.

*Listening Condition differences using the **PALAT** system on Absolute Score at LU0*

An identical one-way repeated-measures ANOVA was conducted to analyze the effect of listening condition (open ear, TEP-100, and ComTac™ III) while using the PALAT system on the dependent measure *absolute score* at LU0 (pretest). A Mauchly's Test of Sphericity was performed since the listening condition factor contained three levels. The Mauchly's Test of Sphericity resulted in no violations of homogeneity of variances of the differences for the listening condition variable (Table 28). **The ANOVA resulted in a significant difference between the mean absolute score for listening conditions**, $F(2,22)=30.46$, $p<0.001$, $\eta_p^2=0.74$. (Table 29).

Post-hoc pairwise comparisons using a paired-samples t-test with Bonferroni correction ($\alpha = 0.05/3$) for each listening condition comparison were conducted using the measure of *absolute score*. Post-hoc comparisons showed **significant differences between all three conditions, open ear** (M=29.58, SD=6.79) **and TEP-100** (M=10.92, SD=5.73), **open ear and**

ComTac^{TM} III (*M*=17.00, *SD*=7.62), **and TEP-100 and ComTac^{TM} III** (Table 30). Figure 92 displays the mean absolute scores for each listening condition while using the PALAT system and error bars representing the 95% confidence intervals about the means. Mean performance for absolute score was highest under the open ear condition followed by the ComTac™ III condition and lowest under the TEP-100 condition.

Table 28. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the PALAT system on ***absolute score*** at LU0 (pretest).

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.82	1.97	2	0.37	0.85	0.99

Table 29. One-way ANOVA table evaluating **differences in *absolute score* at LU0 (pretest) according to listening condition while using the PALAT system** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	65.49			
Within Subjects					
Listening Condition (C)	**2**	**1087.58**	**30.46**	**<0.001***	**0.735**
Error (C x S)	22	35.71			
Total	35	1188.78			

Table 30. Paired-samples *t*-test pairwise **comparisons between listening conditions while using the PALAT system on *absolute score* at LU0 (pretest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**18.67**	**2.64**	**<0.001***
Open ear	**ComTac™ III**	**12.58**	**2.72**	**0.002***
TEP-100	**ComTac™ III**	**-6.08**	**1.86**	**0.022***

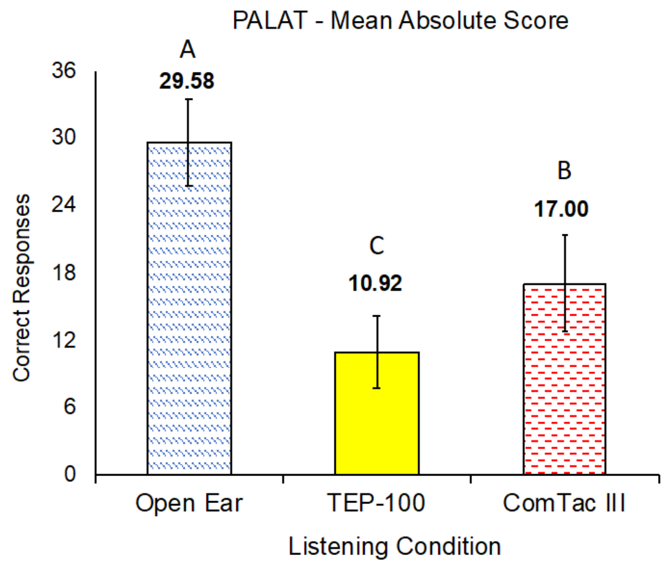


Figure 92. **Mean *absolute score* for each listening condition at LU0 (pretest) while using the PALAT system** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples t-test.

Performance at LU5 (posttest) on Absolute Score

In order to compare localization performance after training, a full factorial repeated-measures ANOVA was used to analyze the effect of training system (DRILCOM and PALAT) and listening condition (open ear, TEP-100, and ComTac™ III) on the dependent measure *absolute score* at LU5 (posttest). A Mauchly's Test of Sphericity performed for all independent variables resulted in no violations of homogeneity of variances between all pairs of within-subjects conditions (Table 31). The **main effect for training system was not significant**, $F(1,11)=1.50$, $p=0.25$, $\eta_p^2=0.12$. The **interaction between training system and listening condition was also not significant**, $F(2,22)=0.98$, $p=0.39$, $\eta_p^2=0.08$. **Significant differences were found for the main effect of listening condition on *absolute score* at LU5**, $F(2,22)=17.22$, $p<0.001$, $\eta_p^2=0.61$ (Table 32).

Table 31. Mauchly's test of sphericity for full factorial ANOVA for the effect of training system and listening condition on <u>*absolute score*</u> at LU5 (posttest).

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.98	0.243	2	0.89	0.98	1
Training System x Listening Condition	0.72	3.303	2	0.19	0.78	1

Table 32. Full factorial ANOVA table evaluating **differences in <u>*absolute score*</u> at LU5 (posttest)** according to training system and listening condition (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	89.35			
Within Subjects					
Training System (A)	1	23.35	1.50	0.246	0.120
Error (A x S)	11	15.53			
Listening Condition (C)	**2**	**1367.01**	**17.22**	**<0.001***	**0.610**
Error (C x S)	22	79.38			
A x C	2	10.43	0.98	0.390	0.082
Error (A x C x S)	22	10.61			
Total	71	1595.66			

<u>*Listening Condition Main Effect: Post hoc test for Absolute Score at LU5*</u>

A paired-samples *t*-test pairwise comparison was conducted for each listening condition (within the main effect of listening condition) using the measure of <u>*absolute score*</u>. Post-hoc comparisons, using a paired-samples *t*-test with Bonferroni correction (α = 0.05/3) for each listening condition, showed a **significant difference between the open ear condition** (*M*=33.83, *SD*=1.44) **and both TCAPS listening conditions**, TEP-100 condition (*M*=12.04, *SD*=1.44) and ComTac™ III condition (*M*=18.38, *SD*=1.95). **No significant difference in absolute score existed between the TEP-100 condition and ComTac™ III condition** (Table 33). Figure 93 displays the mean absolute score for each listening condition and 95% confidence interval about the means.

Table 33. Paired-samples *t*-test pairwise **comparisons between listening conditions on *absolute score* at LU5 (posttest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**14.79**	**2.54**	**<0.001***
Open ear	**ComTac™ III**	**10.00**	**2.41**	**0.005***
TEP-100	ComTac™ III	-4.79	2.75	0.328

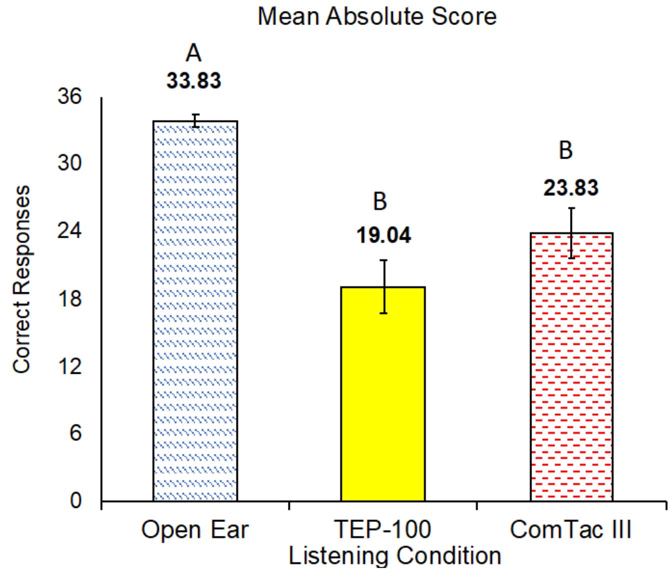


Figure 93. **Mean *absolute score* for each listening condition at LU5 (posttest)** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test.

As with the results from the pretest, no significant differences were found for the main effect of training system or the interaction between training system and listening condition on *absolute score* at LU5 (posttest). For comparison purposes, Figure 94 displays the mean *absolute score* for each listening condition by training system with mean values given above the 95% confidence interval error bars about the mean. Once again, both systems displayed the same trend in localization performance as seen during the pretest.

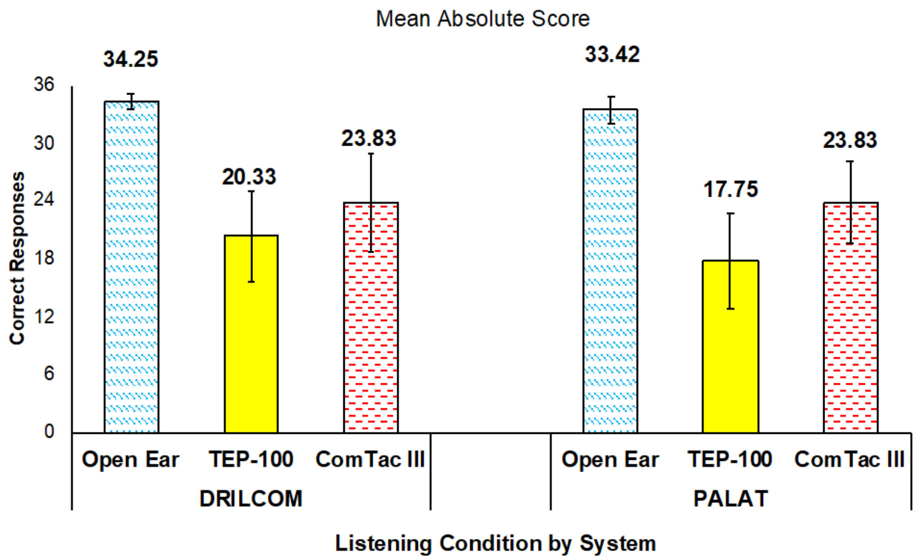


Figure 94. **Mean *absolute score* for each listening condition by training system at LU5 (posttest)** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means.

*Listening Condition differences using the **DRILCOM** system on Absolute Score at LU5*

Figure 94 indicated that both the DRILCOM and PALAT systems were sensitive to auditory localization performance differences between listening conditions. To test this theory, a one-way repeated-measures ANOVA was conducted to analyze the effect of listening condition (open ear, TEP-100, and ComTac™ III) while using the DRILCOM system on the dependent measure *absolute score* at LU5 (posttest). A Mauchly's Test of Sphericity was performed since the listening condition factor contained three levels. The Mauchly's Test of Sphericity resulted in no violations of homogeneity of variances of the differences for the listening condition variable (Table 34). **The ANOVA resulted in a significant difference between the mean absolute score for listening conditions**, $F(2,22)=19.78$, $p<0.001$, $\eta_p^2=0.64$. (Table 35).

Post-hoc pairwise comparisons using a paired-samples *t*-test with Bonferroni correction ($\alpha = 0.05/3$) for each listening condition comparison were conducted using the measure of *absolute score*. Post-hoc comparisons showed **significant differences between open ear**

(*M*=34.25, *SD*=1.49) **and TEP-100** (*M*=20.33, *SD*=8.18) **and open ear and ComTac™ III** (*M*=23.83, *SD*=8.91). **No significant difference was found between TEP-100 and ComTac™ III** (Table 36). Figure 95 displays the mean absolute scores for each listening condition while using the DRILCOM system and error bars representing the 95% confidence intervals about the means. Mean performance for absolute score was highest under the open ear condition followed by the ComTac™ III condition and lowest under the TEP-100 condition.

Table 34. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the DRILCOM system on ***absolute score*** at LU5 (posttest).

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.93	0.77	2	0.68	0.93	1

Table 35. One-way ANOVA table evaluating **differences in *absolute score* at LU5 (posttest) according to listening condition while using the DRILCOM system** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	53.36			
Within Subjects					
Listening Condition (C)	**2**	**628.86**	**13.206**	**<0.001***	**0.546**
Error (C x S)	22	47.62			
Total	35	729.84			

Table 36. Paired-samples *t*-test pairwise **comparisons between listening conditions while using the DRILCOM system on *absolute score* at LU5 (posttest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**13.92**	**2.57**	**0.001***
Open ear	**ComTac™ III**	**10.42**	**2.67**	**0.007***
TEP-100	ComTac™ III	-3.50	3.18	0.882

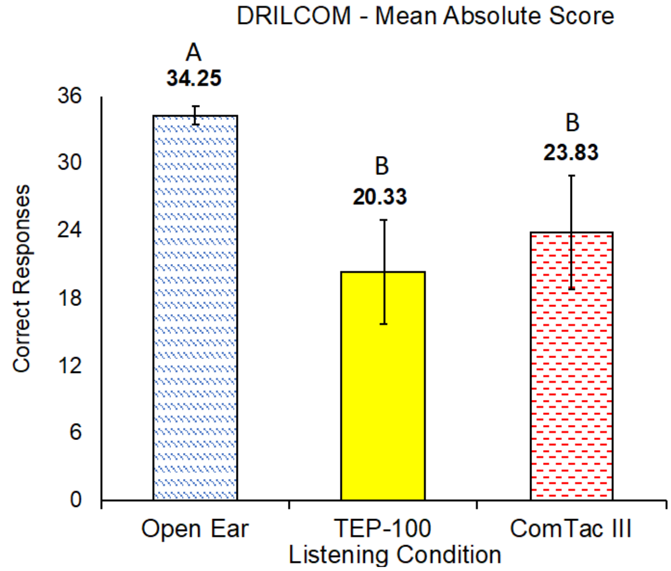


Figure 95. **Mean *absolute score* for each listening condition at LU5 (posttest) while using the DRILCOM system** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test.

Listening Condition differences using the ***PALAT*** *system on Absolute Score at LU5*

An identical one-way repeated-measures ANOVA was conducted to analyze the effect of listening condition (open ear, TEP-100, and ComTac™ III) while using the PALAT system on the dependent measure *absolute score* at LU5 (posttest). A Mauchly's Test of Sphericity was performed since the listening condition factor contained three levels. The Mauchly's Test of Sphericity resulted in no violations of homogeneity of variances of the differences for the listening condition variable (Table 37). **The ANOVA resulted in a significant difference between the mean absolute score for listening conditions**, $F(2,22)=17.67$, $p<0.001$, $\eta_p^2=0.62$. (Table 38).

Post-hoc pairwise comparisons using a paired-samples *t*-test with Bonferroni correction ($\alpha = 0.05/3$) for each listening condition comparison were conducted using the measure of *absolute score*. Post-hoc comparisons showed **significant differences between open ear** ($M=33.42$, $SD=2.47$) **and TEP-100** ($M=17.75$, $SD=8.18$) **and open ear and ComTac™ III**

(*M*=23.83, *SD*=7.41). **No significant difference was found between TEP-100 and ComTac™ III** (Table 39). Figure 96 displays the mean absolute scores for each listening condition while using the PALAT system and error bars representing the 95% confidence intervals about the means. Mean performance for absolute score was highest under the open ear condition followed by the ComTac™ III condition and lowest under the TEP-100 condition.

Table 37. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the PALAT system on ***absolute score*** at LU5 (posttest).

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.98	0.24	2	0.89	0.98	1

Table 38. One-way ANOVA table evaluating **differences in *absolute score* at LU5 (posttest) according to listening condition while using the PALAT system** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	51.52			
Within Subjects					
Listening Condition (C)	**2**	**748.58**	**17.67**	**<0.001***	**0.616**
Error (C x S)	22	42.37			
Total	35	842.47			

Table 39. Paired-samples *t*-test pairwise **comparisons between listening conditions while using the PALAT system on *absolute score* at LU5 (posttest)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**15.67**	**2.73**	**<0.001***
Open ear	**ComTac™ III**	**9.58**	**2.45**	**0.007***
TEP-100	ComTac™ III	-6.08	2.78	0.154

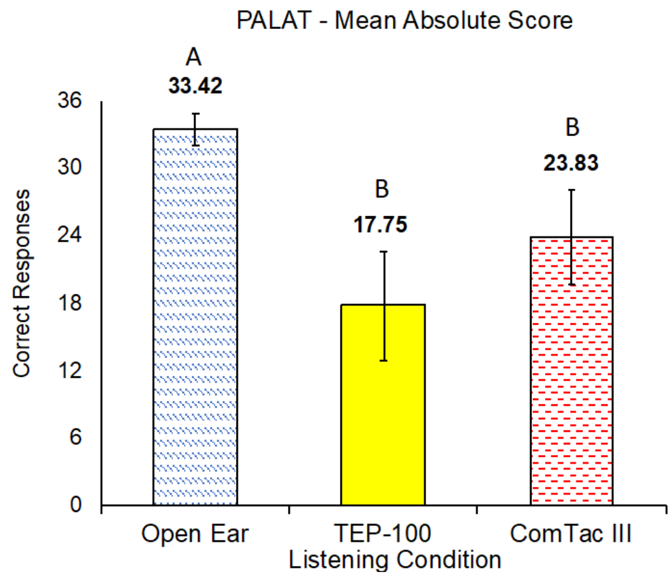


Figure 96. **Mean *absolute score* for each listening condition at LU5 (posttest) while using the PALAT system** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test.

Performance by System, Listening Condition, and Stage of Training on Absolute Score

The previous two statistical analyses demonstrated a similar capability of the DRILCOM and PALAT systems to measure localization performance prior to and following training, that is, at LU0 and LU5. To test whether the two systems achieved similar training results across training, a full factorial repeated-measures ANOVA was conducted to analyze the training effect of training system (DRILCOM and PALAT), listening condition (open ear, TEP-100, and ComTac™ III), and stage of training (LU0, pretest, and LU5, posttest) on the measure of *absolute score*. Results were considered significant at $\alpha=0.05$. The Mauchly's Test of Sphericity performed for all independent variables resulted in no violations of homogeneity of variances (Table 40). The **main effect for training system was not significant**, $F(1,11)=2.48$, $p=0.14$, $\eta_p^2=0.18$ (Table 41). **No significant differences were found in the interactions between training system and listening condition**, $F(2,22)=0.72$, $p=0.50$, $\eta_p^2=0.06$, **between training system and stage of training**, $F(1,11)=0.26$, $p=0.62$, $\eta_p^2=0.02$, **between listening conditions**

and stage of training, $F(2,22)=1.28$, $p=0.30$, $\eta_p^2=0.10$, and **between the three-way interaction between training system, listening condition, and stage of training**, $F(2,22)=0.72$, $p=0.50$, $\eta_p^2=0.06$. **Significant differences were found for the main effect of listening condition**, $F(2,22)=26.63$, $p<0.001$, $\eta_p^2=0.71$, **and for the main effect of stage of training on *absolute score***, $F(1,11)=65.19$, $p<0.001$, $\eta_p^2=0.86$.

Table 40. Mauchly's test of sphericity for full factorial repeated-measures ANOVA for the effect of training system, listening condition, and stage of training on *absolute score*.

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.99	0.06	2	0.97	0.99	1
Training System x Listening Condition	0.79	2.33	2	0.31	0.83	0.96
Listening Condition x Stage of training	0.89	1.23	2	0.54	0.90	1
Training System x Listening Condition x Stage of training	0.97	0.26	2	0.88	0.98	1

Table 41. Full-factorial ANOVA table evaluating **differences in *absolute score* according to training system, listening condition, and stage of training (LU0 and LU5)** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	125.98			
Within Subjects					
Training System (A)	1	84.03	2.48	0.144	0.184
Error (A x S)	11	33.95			
Listening Condition (C)	**2**	**3318.88**	**26.63**	**<0.001***	**0.708**
Error (C x S)	22	124.64			
Stage of Training (T)	**1**	**1067.11**	**65.19**	**<0.001***	**0.856**
Error (T x S)	11	16.37			
A x C	2	8.13	0.72	0.500	0.061
Error(A x C x S)	22	11.37			
A x T	1	5.44	0.26	0.620	0.023
Error(A x T x S)	11	20.91			
C x T	2	29.30	1.28	0.298	0.104
Error(C x T x S)	22	22.87			
A x C x T	2	8.80	0.72	0.500	0.061
Error (A x C x T x S)	22	12.31			
Total	143	4890.09			

Listening Condition Main Effect: Post hoc test for training effect on Absolute Score

Pairwise comparisons were conducted using a paired-samples *t*-test with Bonferroni correction (α = 0.05/3) for each listening condition (within the main effect of listening condition) using the measure of *absolute score*. Post-hoc comparisons showed a **significant difference between the open ear condition** (*M*=31.90, *SD*=4.98) **and both TCAPS listening conditions**, TEP-100 condition (*M*=15.54, *SD*=8.04) and ComTac™ III condition (*M*=21.10, *SD*=7.97). **No significant difference in absolute score existed between the TEP-100 condition and ComTac™ III condition** (Table 42). Figure 97 displays the mean absolute score for each listening condition and 95% confidence intervals about the means. The highest mean localization performance was achieved under the open ear condition followed under the ComTac™ III condition and lastly under the TEP-100 condition.

Table 42. Paired-samples *t*-test pairwise **comparisons between listening conditions on *absolute score* across both training systems and stages of training (LU0 and LU5)** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**16.36**	**2.36**	**<0.001***
Open ear	**ComTac™ III**	**10.79**	**2.27**	**0.002***
TEP-100	ComTac™ III	-5.56	2.21	0.086

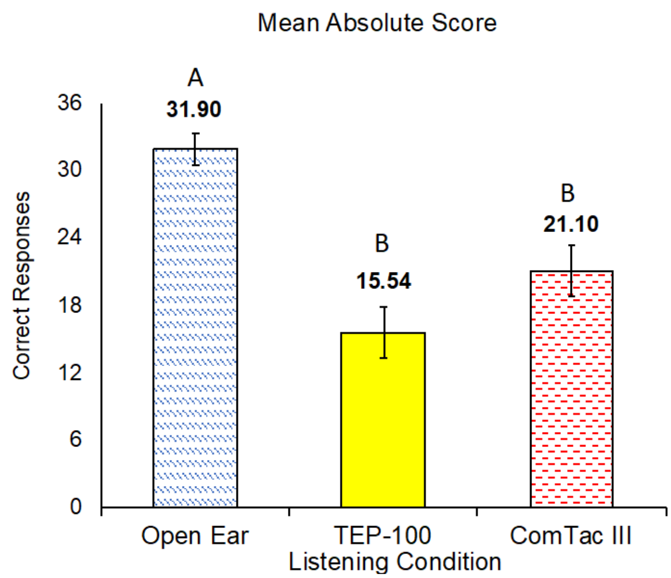


Figure 97. **Mean *absolute score* for each listening condition across both training systems and stages of training (LU0 and LU5)** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test.

Again, no significant differences were found for the main effect of training system or for the interaction between training system and listening condition on *absolute score* from LU0 (pretest) and LU5 (posttest). For comparison purposes, Figure 98 displays the mean *absolute score* for each listening condition by training system with mean values given above the 95% confidence interval error bars around the mean values.

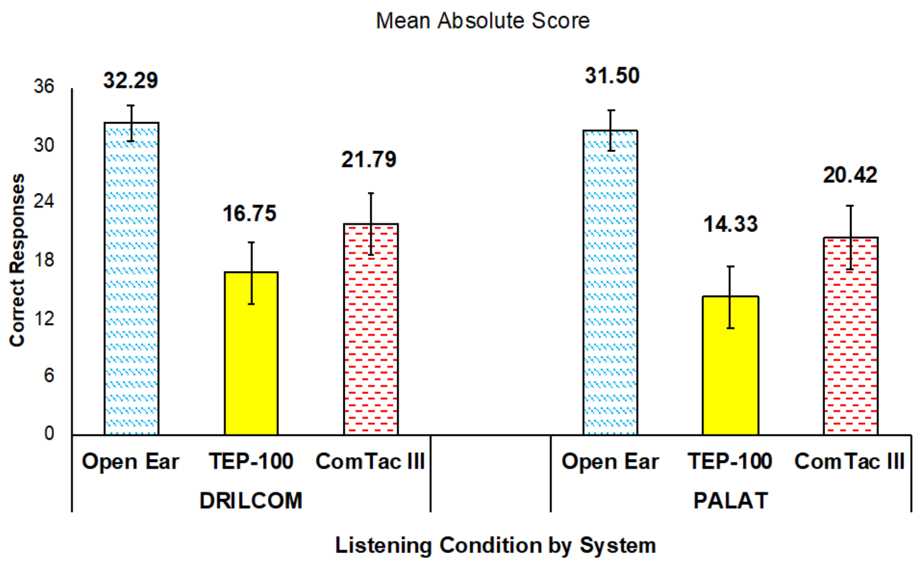


Figure 98. **Mean *absolute score* by training systems and listening condition across stages of training (LU0 and LU5)** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means.

Stage of Training Main Effect: Descriptive statistics for training effect on Absolute Score

The repeated-measures ANOVA found a statistically significant difference between the two level independent variable stage of training, LU0 and LU5, on *absolute score*. Examining the means showed significantly higher localization performance after training at LU5 (posttest, *M*=25.57, *SD*=9.12) compared to pre-training at LU0 (pretest, *M*=25.57, *SD*=9.12). Figure 99 displays the mean absolute score at LU0 (pretest) and LU5 (posttest) with 95% confidence intervals about the means, for the data collapsed across training system and listening condition.

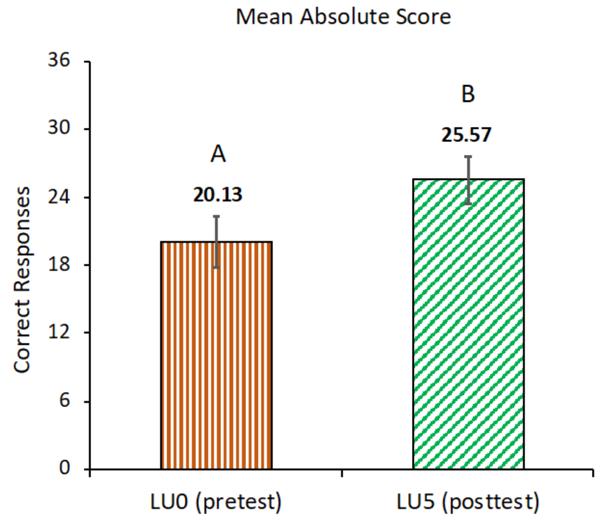


Figure 99. **Mean *absolute score* at LU0 (pretest) and LU5 (posttest)** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per repeated-measures ANOVA.

No significant differences were found for the main effect of training system or the interaction between training system and stage of training on *absolute score*. To show the similar trends between the two training systems, Figure 100 displays the mean *absolute score* at LU0 (pretest) and LU5 (posttest) by training system with mean values given above the 95% confidence interval error bars around the mean values.

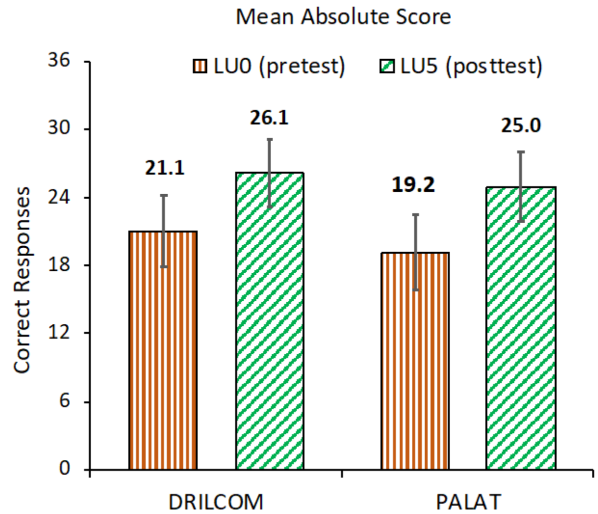


Figure 100. **Mean *absolute score* by training systems at LU0 (pretest) and LU5 (posttest)** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means.

Training effect by System across all Listening Condition on Absolute Score

The mean *absolute score* graph in Figure 100 showed similarities between the DRILCOM and PALAT systems across all listening conditions at LU0 and LU5. The investigator used a series of one-way repeated-measures ANOVAs to compare the *absolute scores* at LU0, pretest, and LU5, posttest, between the DRILCOM and PALAT systems. First, an ANOVA was performed to determine if there was a difference in mean *absolute scores* at LU0 and LU5 while using the DRILCOM system and then the PALAT system. **A significant difference was found between the mean *absolute scores* at LU0 and LU5 while using the DRILCOM system**, $F(1,35)=35.32$, $p<0.001$, $\eta_p^2=0.50$ (Table 43). Likewise, a **significant difference was also found between the mean *absolute scores* at LU0 and LU5 while using the PALAT system**, $F(1,35)=35.00$, $p<0.001$, $\eta_p^2=0.43$ (Table 44). Given that the investigation was a full-factorial investigation, the investigator compared the *absolute scores* of the DRILCOM system at LU0 with the *absolute scores* of the PALAT system at LU5, and vice versa. **A significant difference was found between the mean *absolute scores* at LU0 on the DRILCOM system and LU5 on the PALAT system**, $F(1,35)=16.83$, $p<0.001$, $\eta_p^2=0.33$ (Table 45). Likewise, a **significant difference was also found between the mean *absolute scores* at LU0 on the PALAT system and LU5 on the DRILCOM system**, $F(1,35)=37.93$, $p<0.001$, $\eta_p^2=0.52$ (Table 46). Figure 101 displays the mean *absolute score* at LU0 (pretest) and LU5 (posttest) by training system with mean values given above the 95% confidence interval error bars about the mean.

Table 43. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and LU5** across all listening conditions (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	35	158.09			
Within Subjects					
Stage of Training (T)	**1**	**460.06**	**35.32**	**<0.001***	**0.502**
Error (T x S)	35	13.03			
Total	71	631.18			

Table 44. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and LU5** across all listening conditions (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	35	167.79			
Within Subjects					
Stage of Training (T)	**1**	**612.50**	**26.65**	**<0.001***	**0.432**
Error (T x S)	35	22.99			
Total	71	803.28			

Table 45. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and the PALAT system at LU5** across all listening conditions (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	35	157.67			
Within Subjects					
Stage of Training (T)	**1**	**276.13**	**16.83**	**<0.001***	**0.325**
Error (T x S)	35	16.41			
Total	71	450.21			

Table 46. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and the DRILCOM system at LU5** across all listening conditions (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	35	164.74			
Within Subjects					
Stage of Training (T)	**1**	**875.01**	**37.93**	**<0.001***	**0.520**
Error (T x S)	35	23.07			
Total	71	1062.82			

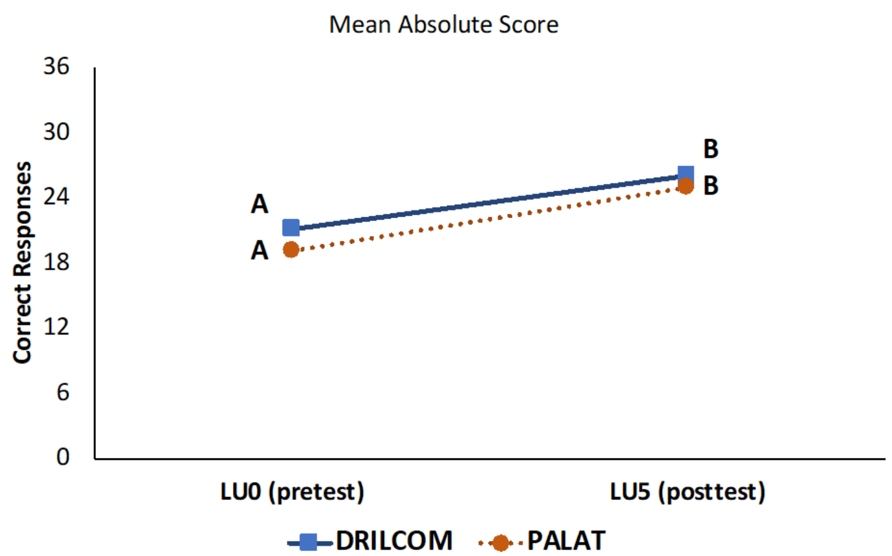


Figure 101. **Mean *absolute score* across all listening conditions at LU0 and LU5 by training system**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs.

Training effect by System under open ear condition on Absolute Score

A series of one-way repeated-measures ANOVAs were conducted to compare the *absolute scores* at LU0, pretest, and LU5, posttest, between the DRILCOM and PALAT systems under the open ear condition. First, an ANOVA was performed to determine if there was a difference in mean *absolute scores* at LU0 and LU5 while using the DRILCOM system and then the PALAT system. **A significant difference was found between the mean *absolute scores* at LU0 and LU5 while using the DRILCOM system under the open ear condition**, *F*(1,11)=5.80, *p*=0.035, η_p^2=0.35 (Table 47). **No significant difference was found between the mean *absolute scores* at LU0 and LU5 while using the PALAT system under the open ear condition**, *F*(1,11)=3.48, *p*=0.089, η_p^2=0.24 (Table 48). Given that the investigation was a full-factorial investigation, the investigator compared the *absolute scores* of the DRILCOM system at LU0 with the *absolute scores* of the PALAT system at LU5, and vice versa. **No significant difference was found between the mean *absolute scores* at LU0 on the DRILCOM system and LU5 on the PALAT system under the open ear condition**, *F*(1,11)=3.54, *p*=0.087, η_p^2=0.24 (Table 49). **A significant difference was found between the mean *absolute scores* at LU0 on the PALAT system and LU5 on the DRILCOM system under the open ear condition**, *F*(1,11)=5.65, *p*=0.037, η_p^2=0.34 (Table 50). Figure 102 displays the mean *absolute score* at LU0 (pretest) and LU5 (posttest) by training system with mean values given above the 95% confidence interval error bars about the mean. While the mean absolute score difference between the DRILCOM at LU0 and PALAT at LU5 was not statistically different, the graphs shows the same trend and increase from training effect under both systems. The mean score difference between the DRILCOM (*M*=30.3, *SD*=5.9) and PALAT (*M*=29.6, *SD*=6.8) system at

LU0 was a difference of 0.7 correct and the mean score difference between the DRILCOM (*M*=34.3, *SD*=1.5) and PALAT (*M*=33.4, *SD*=2.5) system at LU5 was a difference of 0.9 correct.

Table 47. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system under the open ear condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	20.95			
Within Subjects					
Stage of Training (T)	**1**	**92.04**	**5.803**	**0.035***	**0.345**
Error (T x S)	11	15.86			
Total	23	128.85			

Table 48. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system under the open ear condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	26.82			
Within Subjects					
Stage of Training (T)	1	88.17	3.48	0.089	0.240
Error (T x S)	11	25.35			
Total	23	140.34			

Table 49. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and the PALAT system at LU5 under the open ear condition** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	24.56			
Within Subjects					
Stage of Training (T)	1	57.04	3.54	0.087	0.243
Error (T x S)	11	16.13			
Total	23	97.73			

Table 50. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and the DRILCOM system at LU5 under the open ear condition** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	25.17			
Within Subjects					
Stage of Training (T)	**1**	**130.67**	**5.65**	**0.037***	**0.339**
Error (T x S)	11	23.12			
Total	23	178.96			

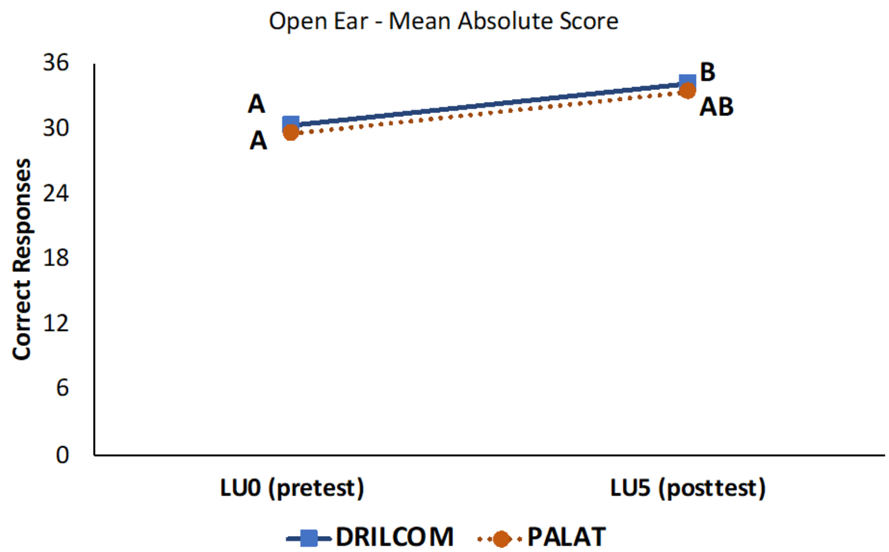


Figure 102. **Mean *absolute score* open ear condition at LU0 and LU5 by training system**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs.

Training effect by System under TEP-100 condition on Absolute Score

The same procedure presented above for open ear was performed for the TEP-100 listening condition. A series of one-way repeated-measures ANOVAs were conducted to compare the *absolute scores* at LU0, pretest, and LU5, posttest, between the DRILCOM and PALAT systems under the TEP-100 condition. First, an ANOVA was performed to determine if there was a difference in mean ***absolute scores*** at LU0 and LU5 while using the DRILCOM system and then the PALAT system. **A significant difference was found between the mean *absolute scores* at LU0 and LU5 while using the DRILCOM system under the TEP-100 condition**, $F(1,11)=25.71$, $p<0.001$, $\eta_p^2=0.70$ (Table 51). Likewise, a **significant difference was found between the mean *absolute scores* at LU0 and LU5 while using the PALAT system under the TEP-100 condition**, $F(1,11)=9.67$, $p=0.010$, $\eta_p^2=0.47$ (Table 52). Given that the investigation was a full-factorial investigation, the investigator compared the *absolute scores* of the DRILCOM system at LU0 with the *absolute scores* of the PALAT system at LU5, and vice versa. **A significant difference was found between the mean *absolute scores* at LU0 on the DRILCOM system and LU5 on the PALAT system under the TEP-100 condition**, $F(1,11)=5.742$, $p=0.035$, $\eta_p^2=0.34$ (Table 53). Also, **a significant difference was found between the mean *absolute scores* at LU0 on the PALAT system and LU5 on the DRILCOM system under the TEP-100 condition**, $F(1,11)=17.60$, $p=0.001$, $\eta_p^2=0.62$ (Table 54). Figure 103 displays the mean *absolute score* at LU0 (pretest) and LU5 (posttest) by training system with mean values given above the 95% confidence interval error bars about the mean.

Table 51. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system under the TEP-100 condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	96.05			
Within Subjects					
Stage of Training (T)	**1**	**308.17**	**25.71**	**<0.001***	**0.700**
Error (T x S)	11	11.99			
Total	23	416.21			

Table 52. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system under the TEP-100 condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	79.12			
Within Subjects					
Stage of Training (T)	**1**	**280.17**	**9.67**	**0.010***	**0.468**
Error (T x S)	11	28.99			
Total	23	388.28			

Table 53. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and the PALAT system at LU5 under the TEP-100 condition** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	94.41			
Within Subjects					
Stage of Training (T)	**1**	**126.04**	**5.74**	**0.035***	**0.343**
Error (T x S)	11	21.95			
Total	23	242.40			

Table 54. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and the DRILCOM system at LU5 under the TEP-100 condition** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	69.56			
Within Subjects					
Stage of Training (T)	**1**	**532.04**	**17.60**	**0.001***	**0.615**
Error (T x S)	11	30.22			
Total	23	631.82			

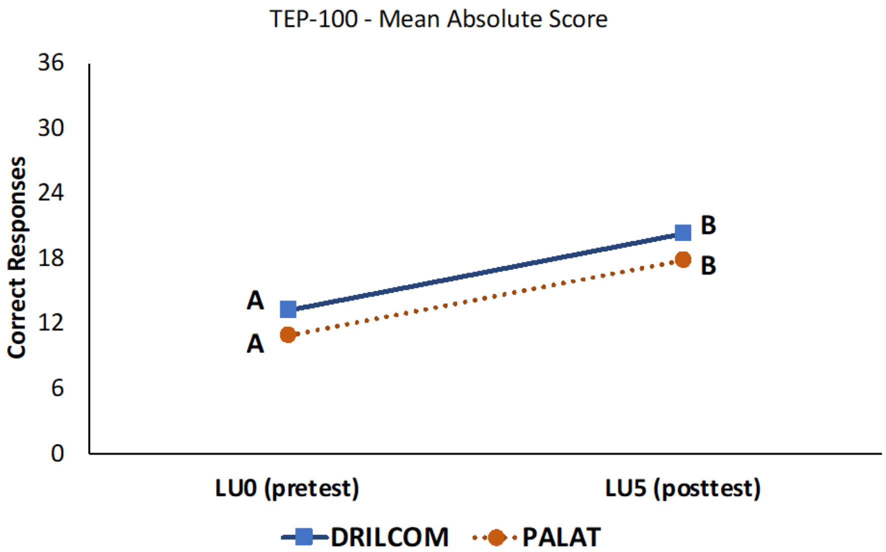


Figure 103. **Mean *absolute score* TEP-100 condition at LU0 and LU5 by training system**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs.

Training effect by System under ComTac™ III condition on Absolute Score

The same procedure presented above for open ear and TEP-100 was performed for the ComTac™ III listening condition. A series of one-way repeated-measures ANOVAs were conducted to compare the *absolute scores* at LU0, pretest, and LU5, posttest, between the DRILCOM and PALAT systems under the ComTac™ III condition. First, an ANOVA was performed to determine if there was a difference in mean *absolute scores* at LU0 and LU5 while using the DRILCOM system and then the PALAT system. **A significant difference was found between the mean *absolute scores* at LU0 and LU5 while using the DRILCOM system under the ComTac™ III condition**, $F(1,11)=10.05$, $p=0.009$, $\eta_p^2=0.48$ (Table 55). Likewise, a **significant difference was found between the mean *absolute scores* at LU0 and LU5 while using the PALAT system under the ComTac™ III condition**, $F(1,11)=18.04$, $p=0.001$, $\eta_p^2=0.62$ (Table 56). Given that the investigation was a full-factorial investigation, the investigator compared the *absolute scores* of the DRILCOM system at LU0 with the *absolute* *scores* of the PALAT system at LU5, and vice versa. **A significant difference was found between the mean *absolute scores* at LU0 on the DRILCOM system and LU5 on the PALAT system under the ComTac™ III condition**, $F(1,11)=7.41$, $p=0.020$, $\eta_p^2=0.40$ (Table 57). Also, **a significant difference was found between the mean *absolute scores* at LU0 on the PALAT system and LU5 on the DRILCOM system under the ComTac™ III condition**, $F(1,11)=20.17$, $p=0.001$, $\eta_p^2=0.65$ (Table 58). Figure 104 displays the mean *absolute score* at LU0 (pretest) and LU5 (posttest) by training system with mean values given above the 95% confidence interval error bars about the mean.

Table 55. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system under the ComTac™ III condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	111.68			
Within Subjects					
Stage of Training (T)	**1**	**100.04**	**10.05**	**0.009***	**0.478**
Error (T x S)	11	9.95			
Total	23	221.67			

Table 56. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system under the ComTac™ III condition at LU0 and LU5** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	97.35			
Within Subjects					
Stage of Training (T)	**1**	**280.17**	**18.04**	**0.001***	**0.621**
Error (T x S)	11	15.53			
Total	23	393.05			

Table 57. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the DRILCOM system at LU0 and the PALAT system at LU5 under the ComTac™ III condition** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	83.59			
Within Subjects					
Stage of Training (T)	**1**	**100.04**	**7.41**	**0.020***	**0.403**
Error (T x S)	11	13.50			
Total	23	197.13			

Table 58. One-way repeated-measures ANOVA table evaluating **differences in *absolute score* on the PALAT system at LU0 and the DRILCOM system at LU5 under the ComTac™ III condition** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	123.53			
Within Subjects					
Stage of Training (T)	**1**	**280.17**	**20.17**	**0.001***	**0.647**
Error (T x S)	11	13.89			
Total	23	417.59			

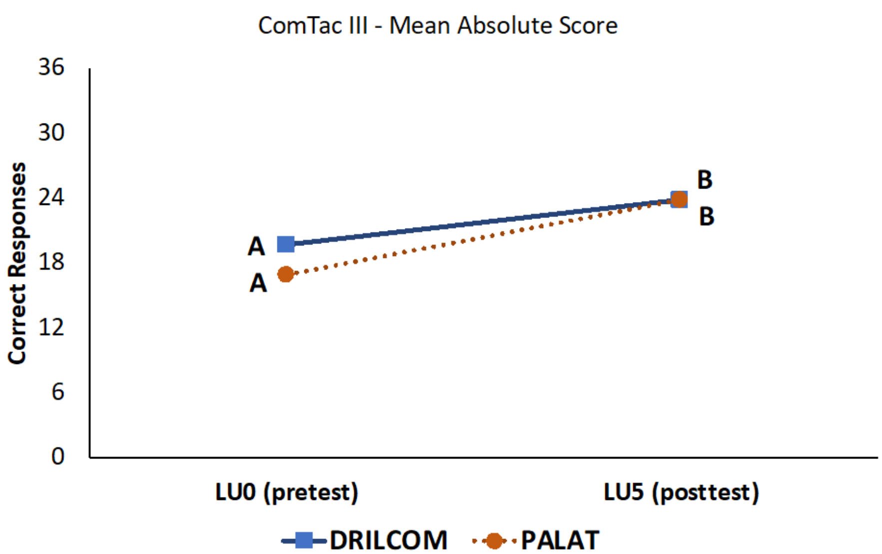


Figure 104. **Mean *absolute score* ComTac™ III condition at LU0 and LU5 by training system**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs.

Training Effect over Learning Units (LU0 through LU5) by Slope of Absolute Score

Figure 80 displayed similar trends in the mean training rates while using both the DRILCOM and PALAT systems for each listening condition. These training rates were depicted by similar regression line *slopes* of *absolute score* from LU0 through LU5 as well as the percent improvement from LU0 to LU5 shown on the right side of the graph. Pairwise comparisons using paired-samples *t*-test were performed to evaluate differences in mean training rates, measured by the *slope* of the absolute correct response scores from LU0 through LU5, between the DRILCOM system and PALAT system for each listening condition. First, a simple linear regression was performed to fit a regression line to each participant's absolute scores from LU0 to LU5 by training system and listening condition. (These linear regression equations appear later in Table 64, while the analyses on the equations ensue here.) The slopes of the regression lines for each participant were then evaluated to determine if a difference in *slope* means existed by using a paired-samples *t*-test comparing performance on the DRILCOM system with performance on the PALAT system for each listening condition. Paired-samples *t*-test were used since the study involved a completely within-subjects experiment and each participant completed the auditory localization test on both systems in a counterbalanced order (Scheaffer & McClave, 1990). Paired-samples *t*-tests were evaluated at $\alpha=0.05$ with a Bonferroni adjustment ($\alpha=0.05/3 = 0.0167$) to control the Type I error rate for multiple comparisons (Keppel & Wickens, 2004). The SPSS® Statistics software adjusted the *p*-values allowing comparison of tabled results at $\alpha=0.05$. **No significant differences were found on mean *slope* between the DRILCOM and PALAT systems for the open ear condition** ($t[11]=-0.66$, $p=0.521$), **TEP-100 condition** ($t[11]=-0.49$, $p=0.633$), **and ComTac™ III condition** ($t[11]=-0.90$, $p=0.390$). Descriptive statistics and paired-samples *t*-test results for *slope* of mean absolute scores from LU0 through LU5 are shown in Table 59.

Table 59. **Comparisons of *slope* for each listening condition between training systems (DRILCOM and PALAT) from LU0 (pretest) through LU5 (posttest)**.

Source		*n*	*M*	*SD*	*t*	*p*
Open ear – DRILCOM versus PALAT					-0.66	0.521
	DRILCOM	12	0.68	0.89		
	PALAT	12	1.02	1.56		
TEP-100 – DRILCOM versus PALAT					-0.49	0.633
	DRILCOM	12	1.29	0.78		
	PALAT	12	1.57	1.73		
ComTac™ III – DRILCOM versus PALAT					-0.90	0.390
	DRILCOM	12	0.92	1.00		
	PALAT	12	1.31	1.13		

*Training effect of Listening Condition on Slope from LU0 through LU5 using **DRILCOM***

Figure 80 displayed similar trends in mean training rates between listening conditions while using the DRILCOM system. To test this theory, a one-way repeated-measures ANOVA was performed to evaluate the effect of listening condition on regression line *slope* of absolute score from LU0 through LU5 while using the DRILCOM system. The Mauchly's Test of Sphericity resulted in no violations of homogeneity of variances of the differences for the listening condition variable (Table 60). **The ANOVA resulted in no significant differences between the mean slope for listening conditions**, $F(2,22)=1.68$, $p=0.209$, $\eta_p^2=0.13$ (Table 61).

Table 60. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the DRILCOM system on *slope* from LU0 through LU5.

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.99	0.05	2	0.98	0.99	1

Table 61. One-way repeated-measures ANOVA table evaluating **differences in _slope_ from LU0 through LU5 on the DRILCOM system by listening condition** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	1.04			
Within Subjects					
Listening Condition (C)	2	1.15	1.68	0.209	0.133
Error (C x S)	22	0.68			
Total	35	2.87			

Training effect of Listening Condition on Slope from LU0 through LU5 using ***PALAT***

The investigator hypothesized that the PALAT system would be sensitive to auditory localization performance differences among the different listening conditions. To test this theory, an identical one-way repeated-measures ANOVA was performed to evaluate the effect of listening condition on regression line *slope* of absolute score from LU0 through LU5 while using the PALAT system. The Mauchly's Test of Sphericity resulted in no violations of homogeneity of variances of the differences for the listening condition variable (Table 62). **The ANOVA resulted in no significant differences between the mean slope for listening conditions**, $F(2,22)=1.68$, $p=0.209$, $\eta_p^2=0.13$ (Table 63).

Table 62. Mauchly's test of sphericity for one-way ANOVA for the effect of listening condition while using the PALAT system on *slope* from LU0 through LU5.

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.93	0.78	2	0.68	0.93	1

Table 63. One-way repeated-measures ANOVA table evaluating **differences in *slope* from LU0 through LU5 on the PALAT system by listening condition** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	2.34			
Within Subjects					
Listening Condition (C)	2	0.92	0.42	0.662	0.037
Error (C x S)	22	2.18			
Total	35	5.45			

Figure 105 shows no significant difference in the mean regression line slopes of absolute score from LU0 through LU5 for listening condition or training system. The slopes were slightly higher on the PALAT system for all listening conditions.

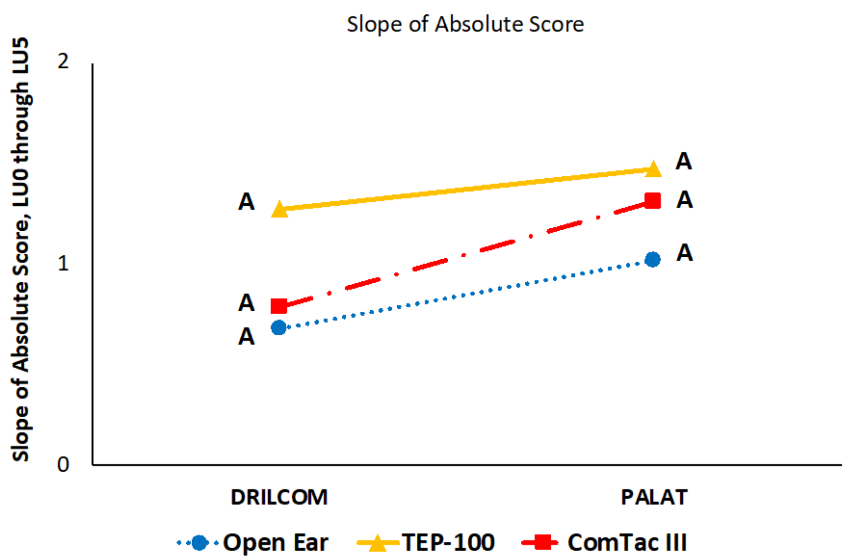


Figure 105. **Mean slope of absolute score from LU0 through LU5 by listening condition on the DRILCOM and PALAT systems**. Different letters indicate significant differences at $p<0.05$ per a one-way repeated-measures ANOVAs.

The PALAT system was designed to provide a capability of evaluating and training U.S. service members' auditory localization skills. In order to quickly assess the level of training needed, the results of Phase II were used to develop initial learning rates to predict potential localization performance. Thus, the simple linear regression lines that were calculated for both training systems at each listening condition, with the purpose of fitting a training rate model from LU0 (pretest) through LU5 (posttest) to be used to predict future training under similar conditions, are useful for predictive purposes. Thus, Table 64 provides the resulting mean *absolute score* linear regression equation by training system and listening condition along with the standard error and R-squared value.

Table 64. Linear regression equations for *absolute score* by training system and listening condition from LU0 (pretest) through LU5 (posttest).

System	Listening Condition	Linear Regression Equation	Std. Error	R-squared
	Open ear	absolute score = 31.37 + 0.679*LU	3.242	0.34
DRILCOM	TEP-100	absolute score = 15.60 + 1.271*LU	7.524	0.28
	ComTac™ III	absolute score = 20.48 + 0.786*LU	7.676	0.18
	Open ear	absolute score = 28.64 + 1.017*LU	4.765	0.45
PALAT	TEP-100	absolute score = 12.16 + 1.474*LU	6.927	0.35
	ComTac™ III	absolute score = 18.44 + 1.314*LU	7.445	0.29

Directional Accuracy by System, Listening Condition, and Stage of Training

The investigator evaluated localization accuracy by loudspeaker location in order to compare the PALAT system with the DRILCOM system by listening condition and stage of training. Statistical analyses described in the previous sections demonstrated similar ability between the PALAT and DRILCOM systems to impart localization skills and to differentiate localization performance between listening conditions. Given the within-subjects investigation, the investigator hypothesized that localization performance should also follow similar directional patterns when using both the PALAT and DRILCOM systems. Localization accuracy was measured by calculating the percent absolute correct for all 12 participants at each loudspeaker position for LU0 and LU5 by listening condition. Correlation analysis was performed to evaluate the relationship between mean percent accuracy while using the PALAT system and while using the DRILCOM system by loudspeaker location and under each listening condition. A

Spearman's rank order correlation was conducted because the percent accuracy data were not normally distributed when tested using a Shapiro-Wilk's test for normality, (PALAT percent accuracy data $W(72)=0.96$, $p=0.027$, DRILCOM percent accuracy data $W(72)=0.95$, $p=0.018$). There was a **strong, positive correlation between percent accuracy scores using the PALAT system and percent accuracy scores using the DRILCOM system by loudspeaker location for all listening conditions** ($r_s(70)=0.92$, $p<0.001$). This indicated that participants performed well or poorly in the same azimuthal locations on both the DRILCOM and PALAT systems. Table 65 shows the descriptive statistics for the overall correlation and correlation under each listening condition between the PALAT and DRILCOM system at loudspeaker location.

Table 65. Spearman correlation for percent accuracy while using the PALAT system and the DRILCOM system by listening condition.

Source	Spearman rho (r_s)	*M*	*SD*	*p*
All Listening Conditions	**0.92**			**<0.001***
PALAT percent accuracy		0.61	0.24	
DRILCOM percent accuracy		0.66	0.22	
Open ear	**0.54**			**0.007***
PALAT percent accuracy		0.87	0.10	
DRILCOM percent accuracy		0.89	0.09	
TEP-100	**0.90**			**<0.001***
PALAT percent accuracy		0.40	0.14	
DRILCOM percent accuracy		0.47	0.15	
ComTac™ III	**0.80**			**<0.001***
PALAT percent accuracy		0.57	0.18	
DRILCOM percent accuracy		0.61	0.16	

Radial plots were created and used to visually compare *percent accuracy* between the two training systems. Figure 106 displays localization *percent accuracy* between training system for each listening condition at LU0, pretest, and LU5, posttest. Figure 107 displays localization *percent accuracy* between LU0, pretest, and LU5, posttest, by training system for each listening condition. As hypothesized, participants' localization accuracy by loudspeaker location demonstrated similar patterns on the PALAT and DRILCOM systems. Percent accuracy under the open ear condition at LU0 showed a diamond shape radial plot where participants were most accurate at the 12, 3, 6, and 9 o'clock positions. The percent accuracy reduced in between each of the four cardinal direction positions. A similar pattern is seen in the percent accuracy plots under TEP-100 and ComTac™ III, with the reduced performance in between 12, 3, 6, and 9 o'clock positions being most prominent under the TEP-100 condition. The radial plots at LU5 also demonstrate similar patterns displayed by overlapping shapes representing increased percent accuracy at every loudspeaker position on both the PALAT and DRILCOM systems. As confirmed by the ANOVA analysis, participants were most accurate under the open ear condition and worst under the TEP-100 condition at LU0 and LU5. *Participants performed extremely poorly under the TEP-100 condition at all loudspeaker positions.* Participants scored less than 54% correct from every loudspeaker position with the exception of the 12 o'clock position on the pretest at LU0 with the TEP-100. However, a similar training effect is shown at LU5 by the increased percent accuracy at all loudspeaker locations on both the PALAT and DRILCOM systems. For the TEP-100, the LU5, posttest, radial plot is shaped much more like an oval than the LU0, pretest, indicating that performance is more consistent at all azimuthal locations with the most accurate performance directly in front or behind the participant (Figures 106 and 107). A similar pattern was observed under the ComTac™ III condition, with much more oval shape (i.e., more accuracy in 360-degrees). Participants initially had difficulty localizing the dissonant signal from the 4, 5, 7, and 8 o'clock positions located behind the participant with both TCAPS. Figure 107 displays the training effect under the ComTac™ III condition by training system represented by the increased localization performance behind the participant.

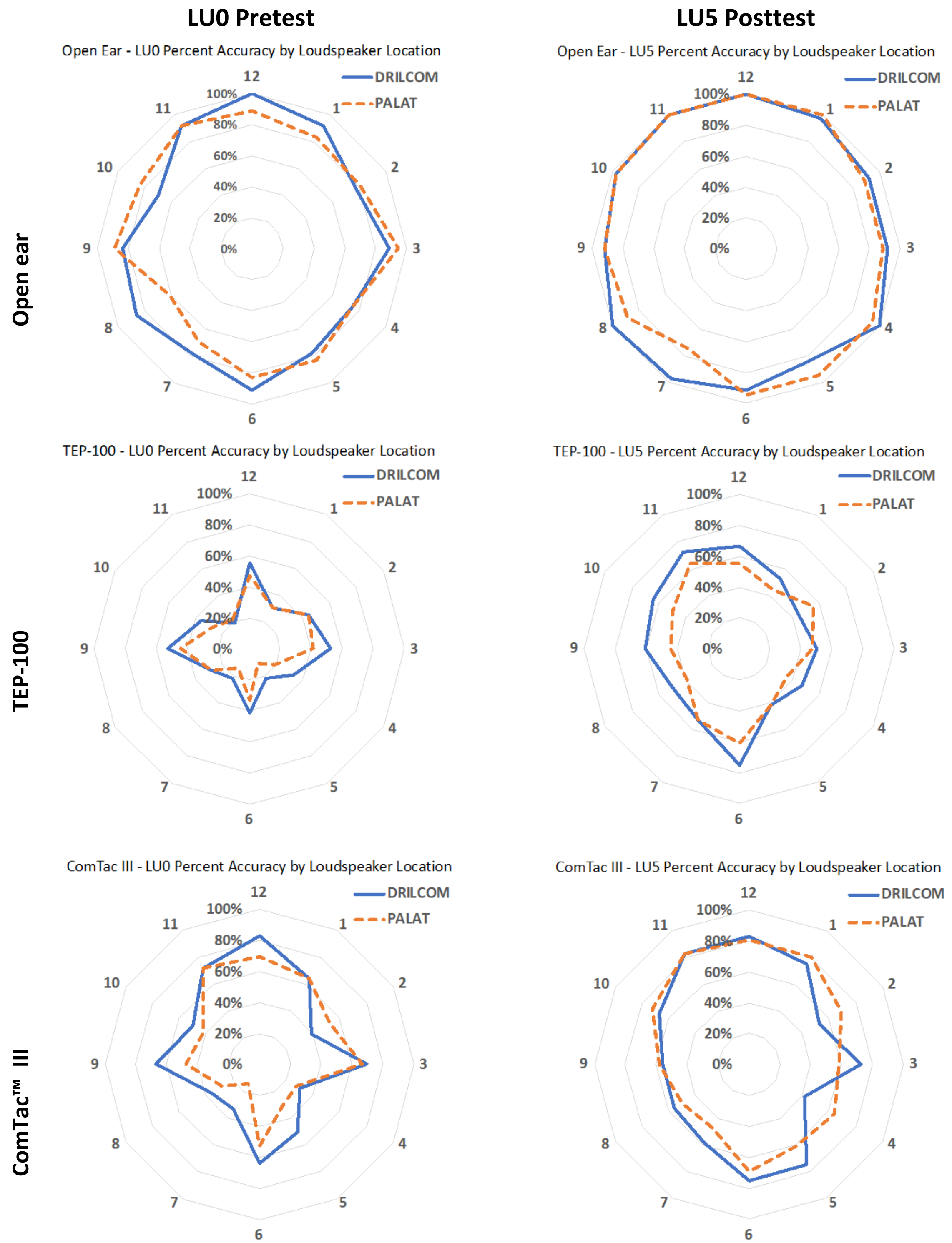


Figure 106. Radial plots of mean absolute correct accuracy percentage by loudspeaker source location for each listening condition comparing training system, DRILCOM (solid blue line) and PALAT (dashed orange line) by stage of training, LU0, pretest, (left column) and LU5, posttest, (right column). 12 represents the position directly in front of the participant, at 12 o’clock or 0 degrees azimuth.

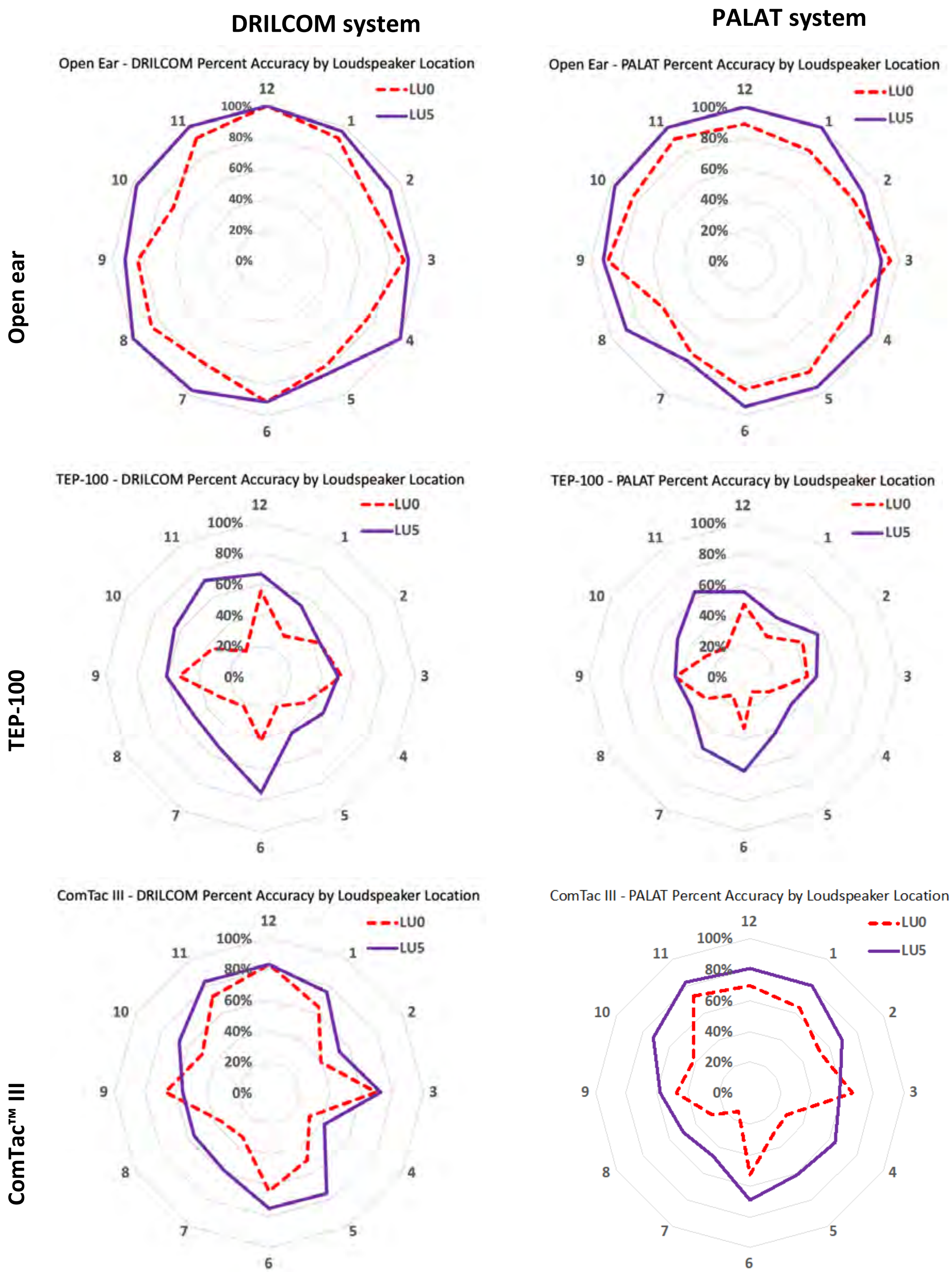


Figure 107. Radial plots of mean absolute correct accuracy percentage by loudspeaker source location for each listening condition comparing LU0, pretest, (dashed red line) and LU5, posttest, (solid purple line) by training system, DRILCOM (left column) and PALAT (right column). 12 represents the position directly in front of the participant, at 12 o’clock or 0 degrees azimuth.

Front-back Reversal Errors Analyses

<u>Effects of System, Listening Condition, and Stage of Training on Front-back Errors</u>

In addition to absolute correct responses, analyses were performed on the dependent measure of the number of *<u>front-back reversal errors</u>* out of 36 signal presentations during each localization test at the end of each learning unit. A full factorial repeated-measures ANOVA was performed to analyze the training effect of training system (DRILCOM and PALAT), listening condition (open ear, TEP-100, and ComTac™ III), and stage of training (LU0, pretest, and LU5, posttest) on the measure of *<u>front-back reversal errors</u>*. As noted previously, a front-back reversal error occurred if a signal was presented from a loudspeaker in positions in front of the participant between 10 o'clock and 2 o'clock (clockwise) and a participant responded with a loudspeaker position behind them between 4 o'clock and 8 o'clock (clockwise) or vice versa. Results were considered significant at α=0.05. Mauchly's Test of Sphericity was performed for all independent variables with more than two levels. The interaction between listening condition and stage of training resulted in a violation of the assumption of sphericity ($\chi^2(2)=8.69, p=0.013$). As a result of the violation, the Greenhouse-Geisser estimator was used to evaluate significance in the difference in means of the listening condition and stage of training interaction. No other independent variables violated the assumption of homogeneity of variances (Table 66). The **main effect for training system was not significant**, $F(1,11)=0.63$, $p=0.45$, $\eta_p^2=0.05$ (Table 67). **No significant differences were found in the interactions between training system and listening condition**, $F(2,22)=1.05$, $p=0.37$, $\eta_p^2=0.09$, **between training system and stage of training**, $F(1,11)=4.57$, $p=0.06$, $\eta_p^2=0.29$, **between listening conditions and stage of training using Greenhouse-Geisser adjusted degrees of freedom**, $F(1.27,13.92)=1.67$, $p=0.22$,

η_p^2=0.13, or **between the three way interaction between training system, listening condition, and stage of training**, $F(2,22)=0.31$, $p=0.74$, η_p^2=0.03. **Significant differences between the means of *front-back reversal errors* were found for the main effect of listening condition**, $F(2,22)=35.29$, $p<0.001$, η_p^2=0.76 and for the **main effect of stage of training**, $F(1,11)=32.58$, $p<0.001$, η_p^2=0.75.

Table 66. Mauchly's test of sphericity for full factorial repeated-measures ANOVA for the effect of training system, listening condition, and stage of training on *front-back reversal errors* (*bold text indicates significant values at the α=0.05 significance level).

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.88	1.21	2	0.55	0.90	1
Training System x Listening Condition	0.80	2.22	2	0.33	0.83	0.96
Listening Condition x Stage of training	**0.42**	**8.69**	**2**	**0.01***	**0.63**	**0.68**
Training System x Listening Condition x Stage of training	0.64	4.39	2	0.11	0.74	0.82

Table 67. Full-factorial ANOVA table evaluating **differences in *front-back reversal errors* according to training system, listening condition, and stage of training (LU0 and LU5)**. Greenhouse-Geisser corrected *p*-value evaluated for significance at α=0.05 (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	23.72			
Within Subjects					
Training System (A)	1	5.06	0.63	0.446	0.054
Error (A x S)	11	8.09			
Listening Condition (C)	**2**	**404.11**	**35.29**	**<0.001***	**0.762**
Error (C x S)	22	11.45			
Stage of Training (T)	**1**	**76.56**	**32.58**	**<0.001***	**0.748**
Error (T x S)	11	2.35			
A x C	2	4.00	1.05	0.369	0.087
Error(A x C x S)	22	3.83			
A x T	1	7.56	4.57	0.056	0.294
Error(A x T x S)	11	1.65			
C x T	1.27	18.97	1.67	0.221	0.132
Error(C x T x S)	13.91	11.34			
A x C x T	2	1.00	0.31	0.735	0.028
Error (A x C x T x S)	22	3.21			
Total	134.18	496.34			

Listening Condition Main Effect: Post hoc test for training effect on Front-back Reversal Errors

Paired-samples *t*-test pairwise comparisons were conducted using a Bonferroni correction (α = 0.05/3) for each listening condition (within the main effect of listening condition) on the dependent measure of *front-back reversal errors*. Post-hoc comparisons showed a **significant difference between the open ear condition** (*M*=0.27, *SD*=0.76) a**nd both TCAPS listening conditions, TEP-100 condition** (*M*=5.94, *SD*=3.56) **and ComTac™ III condition** (*M*=4.19, *SD*=3.12). **No significant difference in mean *front-back reversal errors* existed between the TEP-100 condition and ComTac™ III condition** (Table 68). Figure 108 displays the mean *front-back reversal errors* for each listening condition and 95% confidence intervals about the means. For consistency of scale with the absolute score measure, the mean front-back reversal errors were plotted against a possible 36 responses.

Table 68. Paired-samples *t*-test pairwise **comparisons between listening conditions on *front-back reversal errors*** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**-5.67**	**0.79**	**<0.001***
Open ear	**ComTac™ III**	**-3.92**	**0.58**	**<0.001***
TEP-100	ComTac™ III	1.75	0.69	0.082

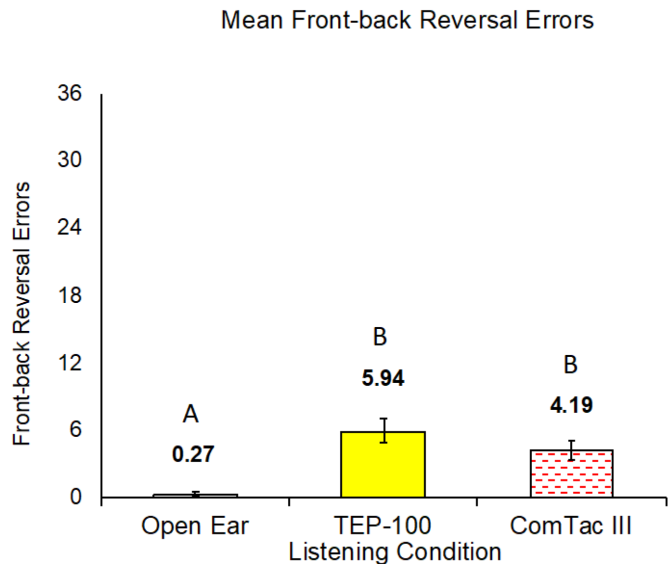


Figure 108. **Mean *front-back reversal errors* for each listening condition across both training systems at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test.

No significant differences were found for the interaction between training system and listening condition on *front-back reversal errors* at LU0 (pretest) and LU5 (posttest). For comparison purposes, Figure 109 displays the mean *front-back reversal errors* for each listening condition by training system with mean values given above the 95% confidence interval error bars about the mean.

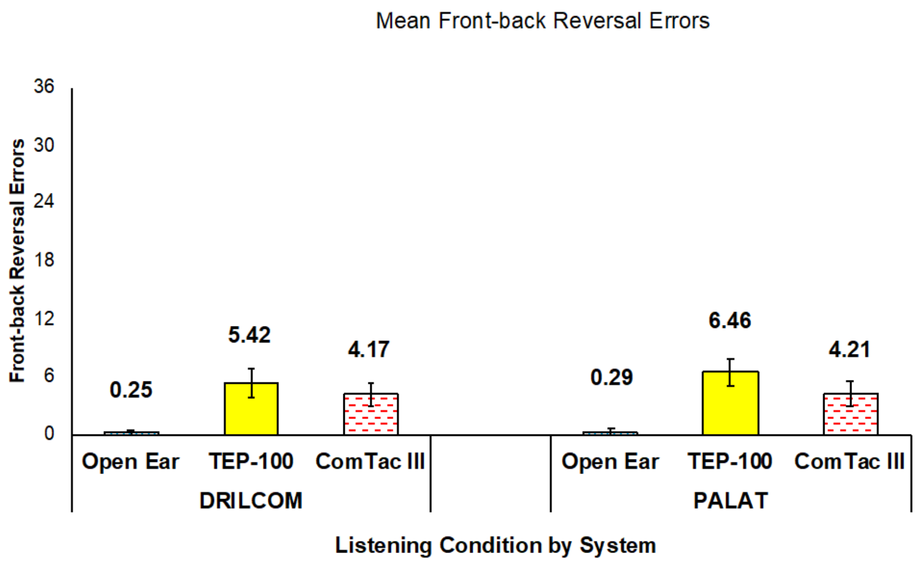


Figure 109. **Mean *<u>front-back reversal errors</u>* by training systems and listening condition at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means.

<u>Stage of Training Main Effect: Descriptive statistics on Front-back Reversal Errors</u>

No post hoc pairwise comparison test was conducted for stage of training since there were only two levels, LU0 and LU5, and the factorial ANOVA showed a significant difference. Comparing the mean values showed that participants recorded significantly fewer front-back reversal errors during LU5 (posttest) (M=2.74, SD=3.31) than prior to training at LU0 (pretest) (M=4.19, SD=3.82). Figure 110 displays the mean *<u>front-back reversal errors</u>* at LU0 (pretest) and LU5 (posttest) with 95% confidence intervals about the means.

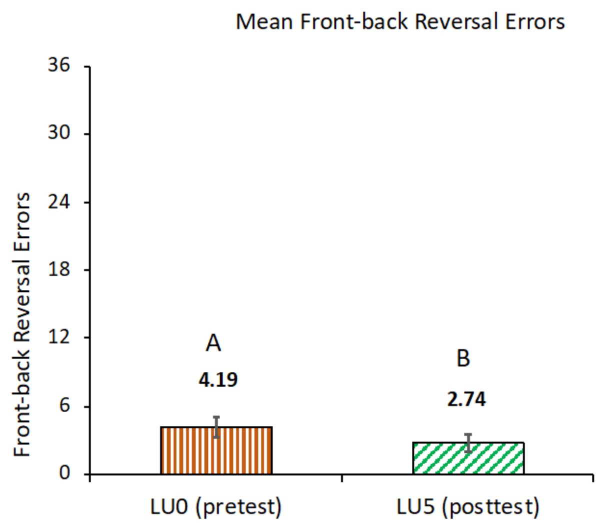


Figure 110. **Mean *front-back reversal errors* at LU0 (pretest) and LU5 (posttest) across all training systems and listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a repeated-measures ANOVA.

No significant differences were found for the interaction between training system and stage of training on *front-back reversal errors*. However, to continue to evaluate the DRILCOM and PALAT systems' abilities to instill the same training value, Figure 111 displays the mean *front-back reversal errors* at LU0 (pretest) and LU5 (posttest) by training system with mean values given above the 95% confidence interval error bars about the mean. Participants recorded a higher number of front-back reversal errors during the pretest using both the DRILCOM and PALAT systems than after training on the posttest. In addition, the rate of change or difference between front-back reversal errors from pretest to posttest dropped at a similar rate on both training systems.

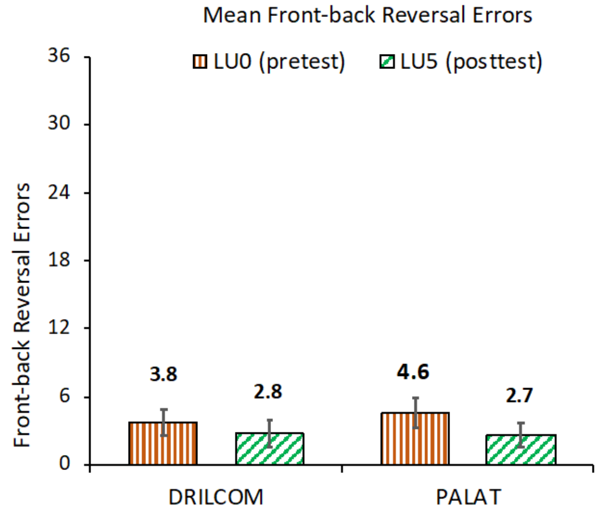


Figure 111. **Mean *front-back reversal errors* by training systems at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Numbers above the error bars are means.

Response Time

Training Effect by System, Listening Condition, and Stage of Training on Response Time

In addition to analyses on absolute score and front-back reversal errors, the dependent measure of *response time* was analyzed against all independent measures to determine if the training systems differed. A full factorial repeated-measures ANOVA was conducted to analyze the training effect of training system (DRILCOM and PALAT), listening condition (open ear, TEP-100, and ComTac™ III), and stage of training (LU0, pretest, to LU5, posttest) on the measure of *response time* to determine if a significant difference in mean response times existed. ANOVA results were considered significant at α=0.05. No violations of homogeneity of variances were found in the Mauchly's Test of Sphericity for all within-subjects independent variables (Table 69). **A statistically significant difference in means on *response time* existed for the main effect for training system**, $F(1,11)=24.87$, $p<0.001$, $\eta_p^2=0.69$ (Table 70). **Significant differences were also found for the main effect of listening condition on *response***

time, $F(2,22)$=12.62, p<0.001, η_p^2=0.53. **No significant differences were found for the main effect of stage of training**, $F(1,11)$=0.81, p=0.39, η_p^2=0.07, **or in the interactions between training system and listening condition,** $F(2,22)$=0.55, p=0.59, η_p^2=0.05, **between training system and stage of training**, $F(1,11)$=0.43, p=0.52, η_p^2=0.04, **between listening conditions and stage of training**, $F(2,22)$=1.17, p=0.33, η_p^2=0.10, and **between the three way interaction between training system, listening condition, and stage of training**, $F(2,22)$=2.44, p=0.11, η_p^2=0.18.

Table 69. Mauchly's test of sphericity for full factorial repeated-measures ANOVA for the effect of training system, listening condition, and stage of training on *response time*.

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	p	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.82	1.98	2	0.37	0.85	0.99
Training System x Listening Condition	0.75	2.83	2	0.24	0.80	0.92
Listening Condition x Stage of training	0.69	3.70	2	0.16	0.76	0.86
Training System x Listening Condition x Stage of training	0.65	4.34	2	0.11	0.74	0.83

Table 70. Full-factorial ANOVA table evaluating **differences in *response time* according to training system, listening condition, and stage of training (LU0 and LU5)** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	1.42			
Within Subjects					
Training System (A)	**1**	**4.16**	**24.87**	**<0.001***	**0.693**
Error (A x S)	11	0.17			
Listening Condition (C)	**2**	**5.96**	**12.62**	**<0.001***	**0.534**
Error (C x S)	22	0.47			
Stage of Training (T)	1	0.16	0.81	0.387	0.069
Error (T x S)	11	0.19			
A x C	2	0.08	0.55	0.586	0.047
Error(A x C x S)	22	0.15			
A x T	1	0.05	0.43	0.524	0.038
Error(A x T x S)	11	0.11			
C x T	2	0.12	1.17	0.329	0.096
Error(C x T x S)	22	0.09			
A x C x T	2	0.41	2.44	0.111	0.181
Error (A x C x T x S)	22	0.17			
Total	143	13.71			

Training System Main Effect: Descriptive statistics for Response Time

No post hoc pairwise comparison was conducted on the two level independent variable of training system. The significant difference detected in the ANOVA table and comparison of means showed that participants recorded **significantly higher response times while using the DRILCOM system** (*M*=2.46, *SD*=0.71) **compared to the PALAT system** (*M*=2.12, *SD*=0.50) across all listening conditions and stages of training. Figure 112 displays the mean response time for the DRILCOM system and PALAT system and 95% confidence intervals about the means.

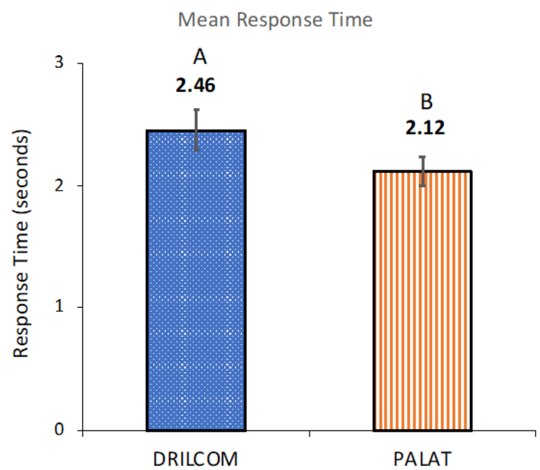


Figure 112. **Mean *response time* for the DRILCOM and PALAT systems across all listening conditions at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a repeated-measures ANOVA.

Listening Condition Main Effect: Post hoc test for training effect on Response Time

Pairwise comparisons were conducted using a paired-samples *t*-test with Bonferroni correction (α = 0.05/3) for each listening condition (within the main effect of listening condition) using the measure of *response time*. Post-hoc comparisons showed a **significant difference between the open ear condition** (*M*=1.91, *SD*=0.44) **and both TCAPS listening conditions, TEP-100 condition** (*M*=2.60, *SD*=0.60) **and ComTac™ III condition** (*M*=2.36, *SD*=0.65). **No significant difference in *response time* existed between the TEP-100 condition and ComTac™ III condition** (Table 71). Figure 113 displays the mean response time for each listening condition and 95% confidence intervals about the means.

Table 71. Paired-samples *t*-test pairwise **comparisons between listening conditions on *response time* across both training systems at LU0 and LU5** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**-0.69**	**0.14**	**0.002***
Open ear	**ComTac™ III**	**-0.45**	**0.11**	**0.005***
TEP-100	ComTac™ III	0.24	0.16	0.497

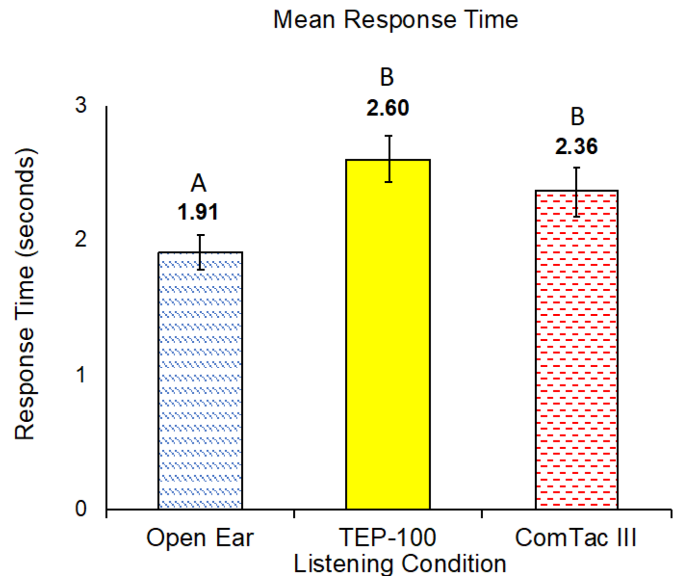


Figure 113. **Mean *response time* for each listening condition across both training systems at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples t-test.

As previously discussed, the PALAT system's user interface was specifically designed to be more efficient for the user to operate than the DRILCOM system. The PALAT system's tablet touch screen interface offered a significantly quicker input mechanism compared to the computer mouse used in the DRILCOM system. As a result, the statistically-significant differences between the training systems on response time are more likely due to the differences in user interface than a quicker participant response time, or reaction to the auditory signal. Due to the differences in user interface between the DRILCOM system and PALAT system, the investigator decided a priori to the study that response time could not be compared directly but the response time dependent measure was amenable to analyses between listening conditions and training effects on each training system independently.

DRILCOM system Response Time by Listening Condition and Stage of Training

A two-way repeated-measures ANOVA was conducted to analyze the effect of listening condition (open ear, TEP-100, and ComTac™ III) and stage of training (LU0, pretest, and LU5, posttest) on the measure of *response time* for tests conducted on the DRILCOM system. Results were considered significant at α=0.05. The Mauchly's Test of Sphericity performed for all independent variables resulted in no violations of homogeneity of variances (Table 72). **A statistically significant difference in means of _response time_ existed for the main effect of listening condition**, $F(2,22)=3.59$, $p=0.002$, $\eta_p^2=0.42$ (Table 73). **No significant differences were found for the main effect of stage of training**, $F(1,11)=0.08$, $p=0.79$, $\eta_p^2=0.01$, **or in the interactions between listening condition and stage of training**, $F(2,22)=2.24$, $p=0.13$, $\eta_p^2=0.17$.

Table 72. Mauchly's test of sphericity for DRILCOM system two-way repeated-measures ANOVA for the effect of listening condition and stage of training on *response time*.

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.84	1.733	2	0.42	0.86	1.00
Listening Condition x Stage of Training	0.64	4.393	2	0.11	0.74	0.82

Table 73. Two-way repeated-measures ANOVA table evaluating **differences in _response time_ for the DRILCOM system according to listening condition and stage of training** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	1.03			
Within Subjects					
Listening Condition (C)	**2**	**3.59**	**8.02**	**0.002***	**0.421**
Error (C x S)	22	0.45			
Stage of Training (T)	1	0.02	0.08	0.788	0.007
Error (T x S)	11	0.20			
C x T	2	0.46	2.24	0.130	0.169
Error (C x T x S)	22	0.20			
Total	71	5.95			

DRILCOM system Listening Condition Main Effect: Post hoc test on Response Time

Pairwise comparisons were conducted using a paired-samples *t*-test with Bonferroni correction ($\alpha = 0.05/3$) for each listening condition (within the main effect of listening condition) using the measure of ***response time*** on the DRILCOM system. Post-hoc comparisons showed a **significant difference between the open ear condition** (*M*=2.03, *SD*=0.49) **and both TCAPS listening conditions, TEP-100 condition** (*M*=2.78, *SD*=0.64) **and ComTac™ III condition** (*M*=2.56, *SD*=0.78). **No significant difference in *response time* existed between the TEP-100 condition and ComTac™ III condition** (Table 74). Figure 114 displays the mean response time for each listening condition and 95% confidence intervals about the means.

Table 74. Paired-samples *t*-test pairwise **comparisons between listening conditions on *response time* on the DRILCOM system** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**-0.75**	**0.17**	**0.003***
Open ear	**ComTac™ III**	**-0.53**	**0.18**	**0.035***
TEP-100	ComTac™ III	0.22	0.23	1.000

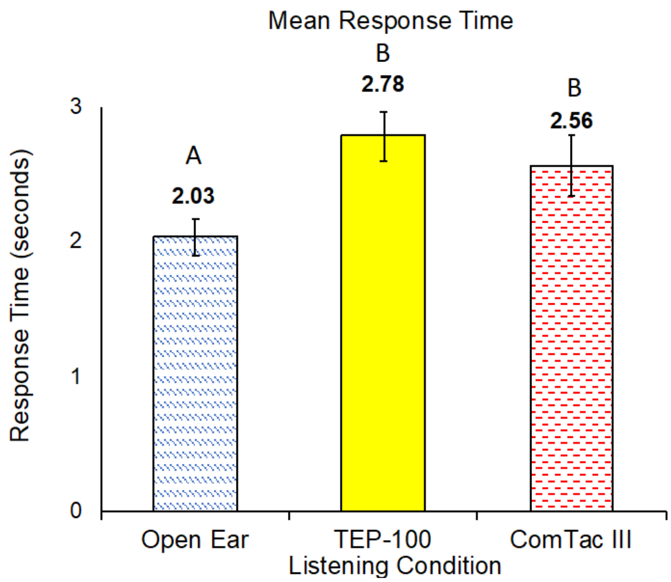


Figure 114. **Mean *response time* for each listening condition on the DRILCOM system at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples *t*-test.

PALAT system Response Time by Listening Condition and Stage of Training

An identical two-way repeated-measures ANOVA was conducted to analyze the effect of listening condition (open ear, TEP-100, and ComTac™ III) and stage of training (LU0, pretest, and LU5, posttest) on the measure of ***response time*** for tests conducted on the PALAT system. Results were considered significant at α=0.05. The Mauchly's Test of Sphericity resulted in a violation of the sphericity assumption for the independent variable listening condition ($\chi^2(2)=6.78, p=0.034$). As a result of the violation, the Greenhouse-Geisser estimator was used to evaluate significance in the difference in means for the main effect of listening condition. No other independent variables violated the assumption of homogeneity of variances (Table 75). **A statistically significant difference in means of *response time* existed for the main effect of listening condition using the Greenhouse-Geisser adjustment for degrees of freedom**, $F(1.34,14.74)=14.45$, $p=0.001$, $\eta_p^2=0.57$ (Table 76). **No significant differences were found for the main effect of stage of training**, $F(1,11)=1.90$, $p=0.20$, $\eta_p^2=0.15$, **or in the interactions between listening condition and stage of training**, $F(2,22)=1.06$, $p=0.36$, $\eta_p^2=0.09$.

Table 75. Mauchly's test of sphericity for PALAT system two-way repeated-measures ANOVA for the effect of listening condition and stage of training on ***response time***.

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	**0.51**	**6.781**	**2**	**0.03***	**0.67**	**0.73**
Listening Condition x Stage of Training	0.70	3.55	2	0.17	0.77	0.87

Table 76. Two-way repeated-measures ANOVA table evaluating **differences in *response time* for the PALAT system according to listening condition and stage of training** (*bold text indicates significant values at the α=0.05 significance level).

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Subjects (S)	11	0.55			
Within Subjects					
Listening Condition (C)	**1.34**	**3.66**	**14.45**	**0.001***	**0.568**
Error (C x S)	14.74	0.25			
Stage of Training (T)	1	0.18	1.90	0.196	0.147
Error (T x S)	11	0.10			
C x T	2	0.07	1.06	0.362	0.088
Error (C x T x S)	22	0.06			
Total	63.08	4.87			

PALAT system Listening Condition Main Effect: Post hoc test on Response Time

Post hoc testing consisted of pairwise comparisons using a paired-samples *t*-test with Bonferroni correction (α = 0.05/3) for each listening condition (within the main effect of listening condition) using the measure of *response time* on the PALAT system. Post-hoc comparisons showed **a significant difference in mean *response times* between the open ear condition** (*M*=1.78, *SD*=0.35) **and both TCAPS listening conditions, TEP-100 condition** (*M*=2.42, *SD*=0.51) **and ComTac™ III condition** (*M*=2.16, *SD*=0.40). **No significant difference in *response time* existed between the TEP-100 condition and ComTac™ III condition** (Table 77). Figure 115 displays the mean response time for each listening condition and 95% confidence intervals about the means.

Table 77. Paired-samples *t*-test pairwise **comparisons between listening conditions on *response time* on the PALAT system** with a Bonferroni adjustment (*bold text indicates significant values at the α=0.05 significance level).

Listening Condition		*Mean Difference*	*Standard Error*	*p*
Open ear	**TEP-100**	**-0.64**	**0.14**	**0.003***
Open ear	**ComTac™ III**	**-0.38**	**0.07**	**<0.001***
TEP-100	ComTac™ III	0.26	0.14	0.238

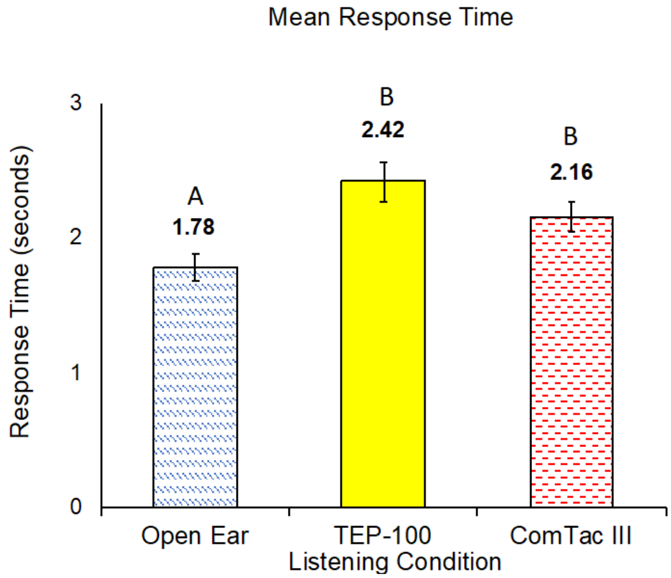


Figure 115. **Mean *response time* for each listening condition on the PALAT system at LU0 and LU5** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a paired-samples t-test.

Directional Response Time by Training System, Listening Condition, and Stage of Training

Radial plots were created and used to visually compare *response times* between the two training systems. Figure 116 displays localization *response times* between training system for each listening condition at LU0, pretest, and LU5, posttest. Figure 117 displays localization *response times* between LU0, pretest, and LU5, posttest, by training system for each listening condition. As hypothesized, participants' response times by loudspeaker location were significantly quicker while using the PALAT system compared to when using the DRILCOM system (Table 70). The quicker mean response times on the PALAT system corroborate the improvements to portable systems' user interface. In addition, mean response times were significantly quicker under the open ear condition than both TCAPS conditions. Response times under all three conditions at LU0 showed a generally square shape radial plot where participants' response times were quickest at the 12, 3, 6, and 9 o'clock positions. The mean response times increased in between each of the four cardinal direction positions. The radial plots at LU5 demonstrate a slightly rounder pattern indicating that mean response times are more similar at every loudspeaker signal location. As confirmed by the ANOVA analysis, participants demonstrated the quickest response times under the open ear condition and worst under the TEP-100 condition at LU0 and LU5 on both the PALAT and DRILCOM systems (Tables 74 and 77). While response time performance was best while using the PALAT system, a similar training effect is shown at LU5 by the slightly quicker response times at most loudspeaker locations on both the PALAT and DRILCOM systems with the exception of the TEP-100 on the DRILCOM system. Response times for the TEP-100 on the DRILCOM system at LU5, posttest, slightly increased possibly indicating that participants were being deliberate when deciding where the signal originated (Figures 116 and 117).

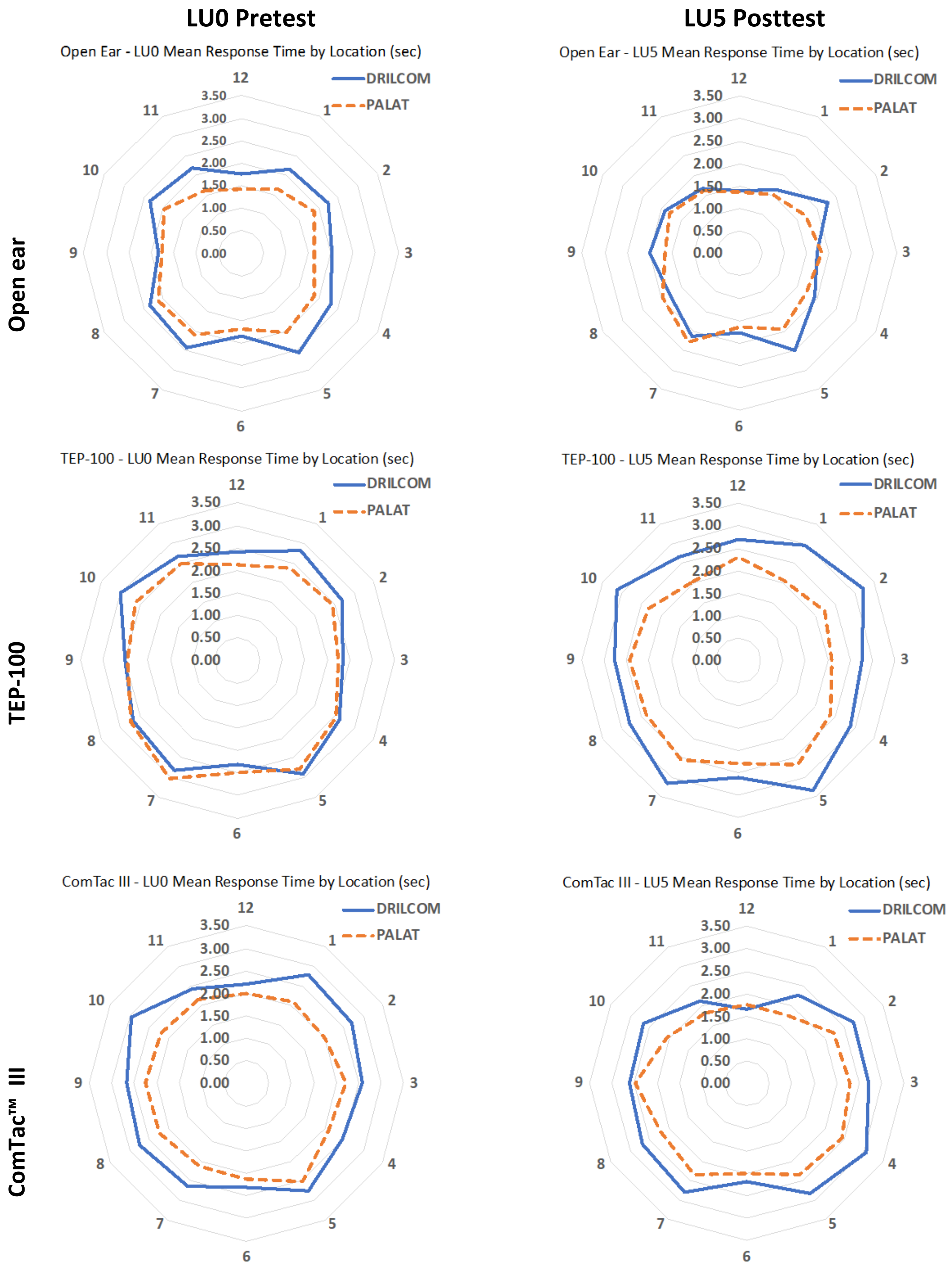


Figure 116. Radial plots of mean response time by loudspeaker source location for each listening condition comparing training system, DRILCOM (solid blue line) and PALAT (dashed orange line) by stage of training, LU0, pretest, (left column) and LU5, posttest, (right column). 12 represents the position directly in front of the participant, at 12 o'clock or 0 degrees azimuth.

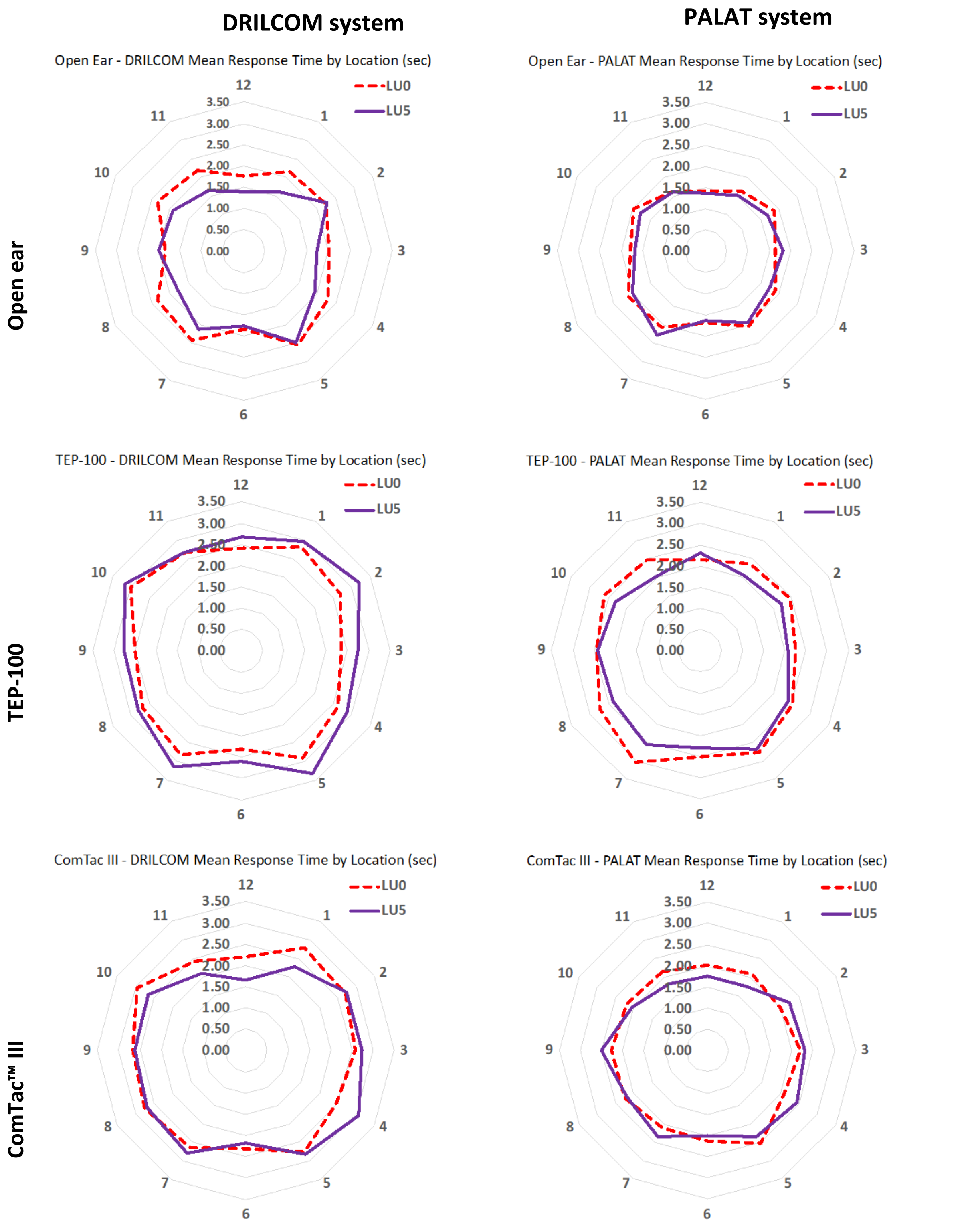


Figure 117. Radial plots of mean response time by loudspeaker source location for each listening condition comparing LU0, pretest, (dashed red line) and LU5, posttest, (solid purple line) by training system, DRILCOM (left column) and PALAT (right column). 12 represents the position directly in front of the participant, at 12 o’clock or 0 degrees azimuth.

3.6.4 Phase II Subjective Measures Statistical Analyses

Following the completion of each LU5 posttest, each participant was asked to evaluate the DRILCOM system or PALAT system under the tested listening condition. A paper-based questionnaire was administered following each of the six training sessions to measure perceived performance and capture participants' subjective ratings in order to compare the training systems. In order to reduce bias from order effects, the within-subjects investigation was completely counterbalanced using two identical 3x6 Latin squares resulting in two sets of every combination of localization training system and listening condition order. All questionnaires shared 10 common questions focused on identifying how the following items impacted training or the ability to localize auditory signals: 1) perceived confidence, 2) loudspeaker proximity, 3) system ease of use, 4) room environment impact on localization, 5) training impacts, 6) difficulty of judging sound location, 7) response time, 8) user interface, 9) room environment on response time, and 10) loudspeaker visibility impacts on localization. All questions used a bipolar seven-point Likert scale from -3 to 3 with 0 as the middle ranking (see Appendix F for the complete set of Likert scales). Questions 11 through 14 were included following the completion of each listening condition (2nd, 4th, and 6th sessions) and asked the participant to compare the PALAT and DRILCOM system under the recently completed listening condition. Lastly, questions 15 through 20 asked participants to directly choose their preference between the DRILCOM system (-3) and PALAT system (3) with an indifferent rating of 0 in the middle. Questions 11 through 20 were very similar to one of the 10 common questions and were used as a questionnaire validity measurement to confirm the statistical findings of a similar Question 1-10. The results of Questions 11 through 20 are reported with descriptive statistics after the statistical results of the similar question they helped to confirm.

The Phase II investigation had three main objectives: 1) Evaluate and validate the effectiveness of the PALAT system compared to the DRILCOM system, 2) Investigate the auditory localization skills acquisition under three listening conditions, and 3) Determine the TCAPS effects on localization accuracy and response time. The first 10 questions, common to all questionnaires, were designed to allow participants to evaluate the three main objectives. The responses to all 10 common questions were evaluated using non-parametric statistical analyses.

Subjective Evaluation Overview of the PALAT System Effectiveness

In order to answer the primary objective of the study, the investigator first used Wilcoxon signed-rank tests to compare differences in participant ratings between the DRILCOM and PALAT systems across all listening conditions. The Wilcoxon signed-rank test allowed for comparison of ordinal scores from two-related populations or repeated-measures on a single sample (Scheaffer & McClave, 1990). Under the Wilcoxon test, participant ratings are ranked and mean rankings are compared to identify if the differences between means follow a symmetric distribution (Scheaffer & McClave, 1990). A significant difference represents an unequal number of positive and negative ranks, or non-symmetric distribution around the mean rankings (Scheaffer & McClave, 1990). The Wilcoxon signed-rank test statistic was evaluated using a significance level of $\alpha=0.05$. A significant finding of the Wilcoxon test demonstrates that participants perceived a difference between the two systems across all listening conditions. Results for each question are reported below.

Subjective Evaluation Overview of Auditory Localization Skills Acquisition

To address the second objective of comparing training effects under different listening conditions, the investigator compared participant' ratings for the DRILCOM system versus the PALAT system under each listening condition in order to detect any perceived differences between the two training systems under the same listening conditions. Three separate Wilcoxon signed-rank test pairwise comparisons were performed for each question in order to compare the DRILCOM and PALAT systems under the open ear condition, TEP-100 condition, and ComTac™ III condition. Each Wilcoxon signed-rank test statistic was evaluated using a significance level of $\alpha=0.05$.

Subjective Evaluation Overview of TCAPS Effects on Auditory Localization Performance

To investigate the third objective of TCAPS effects on localization accuracy and response time, a Friedman test was applied to compare participant responses between listening conditions across both systems. As observed in objective absolute score performance, listening condition played a significant role in the ability to localize the dissonant signal. The non-parametric Friedman test allowed for comparisons between the ordinal rankings of the within-subjects repeated measurements from the three listening conditions across both training systems. The Friedman test assigned ranks for each participant rating across the listening conditions and then ranked the ratings within each listening condition (Portney & Watkins, 2009). The null hypothesis was that there were no significant differences between the open ear, TEP-100, and ComTac™ III ratings for each question. A significant finding indicated a difference was detected among one of the listening conditions. Significant results of the Friedman's test were followed by three Wilcoxon signed-rank test pairwise comparisons between each listening condition. The Friedman's test statistic was evaluated using a significance level of $\alpha=0.05$. Post hoc pairwise comparisons used a Bonferroni correction of $\alpha=0.05/3 = 0.017$ to control for the increase risk of Type I errors due to multiple comparisons.

Lastly, the investigator evaluated whether there were any differences between participant ratings for each listening condition by training system, DRILCOM and PALAT. The testing followed the same procedure for comparing listening conditions across both systems. A Friedman test was applied to compare participant responses between listening conditions on each training system for a total of three comparisons on each system. Significant results of the Friedman's test were followed by a Wilcoxon signed-rank test for pairwise comparisons between each listening condition. The Friedman's test statistic was evaluated using a significance level of α=0.05. Post hoc pairwise comparisons used a Bonferroni correction of α=0.05/3 = 0.017 to control for the increase risk of Type I errors due to multiple comparisons.

The following sections report the results of participant subjective ratings for each question. Data tables are provided for all statistical analyses. Graphs are provided when significant differences were found. However, non-significant findings are not presented in graphs.

Question 1. Training impact on perceived confidence in ability to localize

Participants were asked to respond to the following: "Rate how **training** using the DRILCOM (or PALAT) system **impacted your confidence** in your ability to localize sounds, from **before to after** all the training you received using this system," from -3 (extremely less confident) to 3 (extremely more confident). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 1 Wilcoxon results (Table 78) showed **no significant difference in mean rankings of perceived confidence between the DRILCOM and PALAT system** (Z=-0.54, p=0.590). Figure 118 displays the mean *subjective ratings* for perceived confidence on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 78. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 1, perceived confidence (bolded text in the table indicates a significant test result at $p<0.05$.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	-0.54	0.590	1.58	1.05	1.50	1.08

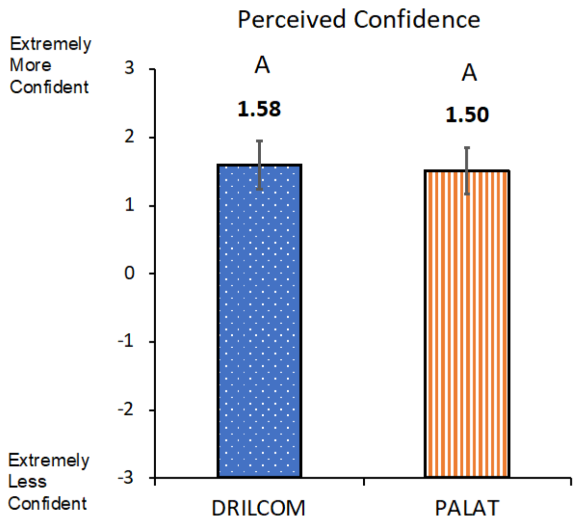


Figure 118. **Mean *subjective ratings* for Question 1, perceived confidence on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 1 pairwise comparison Wilcoxon test results (Table 79) showed **no significant difference in mean rankings of perceived confidence between the DRILCOM and PALAT system for any of the listening conditions**, open ear (Z=-0.25, p=0.803), TEP-100 (Z=-1.19, p=0.234), and ComTac™ III (Z=-0.71, p=0.480). Figure 119 displays the mean *subjective ratings* for perceived confidence on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 79. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 1, perceived confidence (bolded text in the table indicates a significant test result at $p<0.05$.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-0.25	0.803	1.83	0.72	1.75	0.87
TEP-100	-1.19	0.234	1.42	1.38	1.08	1.51
ComTac™ III	-0.71	0.480	1.50	1.00	1.67	0.65

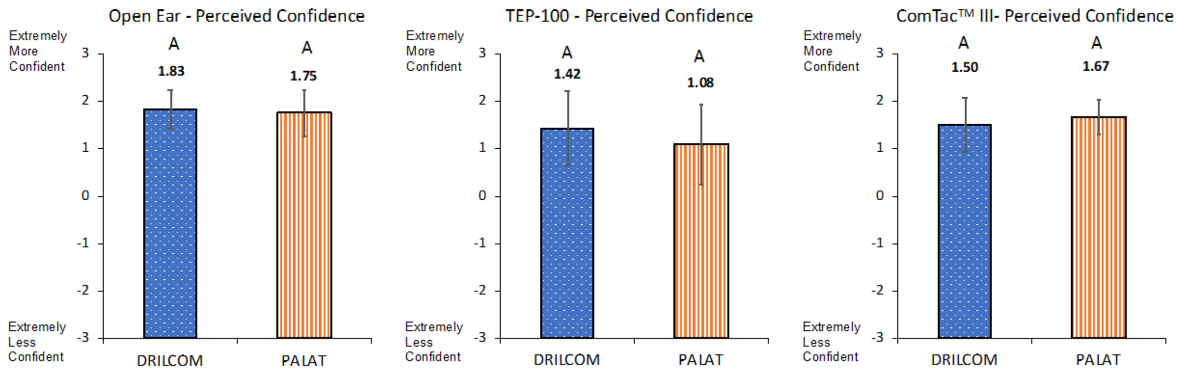


Figure 119. **Mean *subjective ratings* for Question 1, perceived confidence on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both training systems. The Friedman's test for perceived confidence resulted in **no significant differences between mean rankings for listening conditions**, open ear (*M*=1.79, *SD*=0.78), TEP-100 (*M*=1.25, *SD*=1.42), and ComTac™ III (*M*=1.58, *SD*=0.83) across both systems ($\chi^2[2]=0.351$, $p=0.839$).

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each localization training systems, DRILCOM and PALAT. The Friedman's test for perceived confidence while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=1.75, *SD*=0.87), TEP-100 (*M*=1.08, *SD*=1.51), and ComTac™ III (*M*=1.67, *SD*=0.65), (χ^2[2]=0.65, *p*=0.723). Likewise, the Friedman's test for perceived confidence while using the **DRILCOM system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=1.83, SD=0.72), TEP-100 (M=1.42, SD=1.38), and ComTac™ III (M=1.50, SD=1.00), (χ^2[2]=0.05, *p*=0.973).

Participant ratings for how training impacted their confidence in ability to localize sounds showed no significant differences between the DRILCOM and PALAT systems. In addition, there were no significant differences in the mean rankings of perceived confidence between listening conditions. These findings were supported by the mean participant responses to Question 11 and Question 15. Question 11 asked, "Compared to the previously used system, rate how confident you are in your ability to localize sounds using the most recently used system," from -3 (extremely less confident) to 3 (extremely more confident). The investigator anticipated that confidence would be slightly higher on the second or last training system since the participant would have previously trained on the other system under the same listening condition. Indeed, this was borne out in the results - participants who used the DRILCOM system last slightly preferred the DRILCOM system (*M*=0.61, *SD*=1.19), while participants who used the PALAT system last slightly preferred the PALAT system (*M*=0.33, *SD*=1.32). Question 15 asked participants to rate their system preference for "confidence in accurately localizing sounds" between the DRILCOM (-3) or PALAT (3) system. The mean ratings showed no clear preference (*M*=-0.08, *SD*=1.48).

Question 2. Impact of the proximity of the loudspeakers on ability to train to localize sound

Participants were asked to respond to the following: "Rate the **impact** you felt the **proximity (distance) of the loudspeakers** of the DRILCOM (or PALAT) system contributed to your **ability to train to localize** sounds," from -3 (extremely negative impact) to 3 (extremely positive impact). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 2 Wilcoxon results (Table 80) showed **no significant difference in mean rankings of loudspeaker proximity impacts between the DRILCOM and PALAT system** (Z=-1.25, p=0.212). Figure 120 displays the mean *subjective ratings* for loudspeaker proximity impacts using the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 80. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 2, loudspeaker proximity impacts (bolded text in the table indicates a significant test result at p<0.05.)

Source			DRILCOM (n=12)		PALAT (n=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	-1.25	0.212	1.11	1.06	0.75	1.11

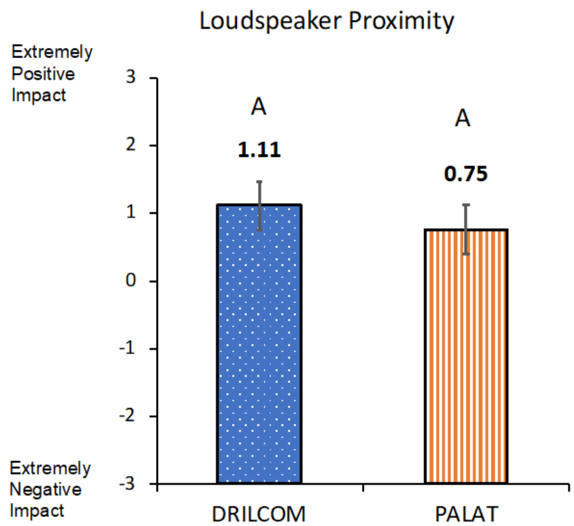


Figure 120. **Mean *subjective ratings* for Question 2, loudspeaker proximity impacts using the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 2 pairwise comparison Wilcoxon test results (Table 81) showed **no significant difference in mean rankings of loudspeaker proximity impacts between the DRILCOM and PALAT system for any of the listening conditions**, open ear (Z=-0.79, p=0.429), TEP-100 (Z=-0.86, p=0.388), and ComTac™ III (Z=-0.33, p=0.739). Figure 121 displays the mean *subjective ratings* for loudspeaker proximity impacts on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 81. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 2, loudspeaker proximity impacts (bolded text in the table indicates a significant test result at $p<0.05$.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-0.79	0.429	1.33	1.23	1.00	0.95
TEP-100	-0.86	0.388	1.08	0.90	0.67	1.30
ComTac™ III	-0.33	0.739	0.92	1.08	0.58	1.08

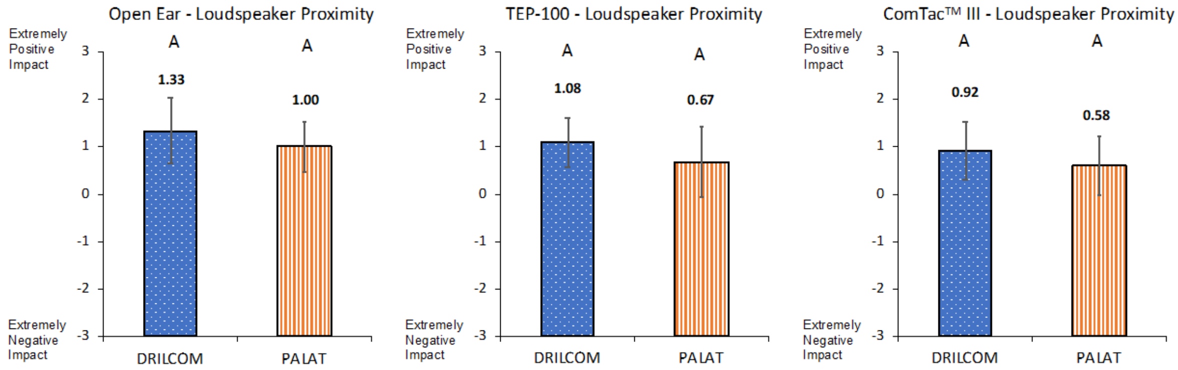


Figure 121. **Mean *subjective ratings* for Question 2, loudspeaker proximity impacts on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at p<0.05 within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both training systems. The Friedman's test for loudspeaker proximity impacts resulted in **no significant differences between mean rankings for listening conditions**, open ear (M=1.17, SD=1.09), TEP-100 (M=0.88, SD=1.16), and ComTac™ III (M=0.75, SD=1.07) across both systems (χ^2[2]=1.55, p=0.461).

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each training system, DRILCOM and PALAT. The Friedman's test for loudspeaker proximity impacts while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (M=1.00, SD=0.95), TEP-100 (M=0.67, SD=1.30), and ComTac™ III (M=0.58, SD=1.08), (χ^2[2]=1.27, p=0.531). Likewise, the Friedman's test for loudspeaker proximity impacts while using the **DRILCOM system resulted in no significant**

differences between mean rankings for listening conditions, open ear (*M*=1.33, SD=1.23), TEP-100 (M=1.08, SD=0.90), and ComTac™ III (M=0.92, SD=1.08), (χ^2[2]=0.44, *p*=0.804).

Participant ratings for how the proximity of the loudspeaker impacted their ability to localize sounds showed no signs of significant differences between the DRILCOM and PALAT systems. In addition, there were no significant differences in the mean rankings of loudspeaker proximity impacts between listening conditions. These findings were supported by the mean participant responses to Question 18. Question 18 asked participants to rate their system preference as to "the loudspeaker configuration and proximity" between the DRILCOM (-3) or PALAT (3) system. The mean ratings showed no clear preference (*M*=0.03, *SD*=1.23).

Question 3. Ease of use to operate the system

Participants were asked to respond to the following: "Rate how **easy it was to operate** the DRILCOM (or PALAT) system hardware and software during your localization training," from -3 (extremely difficult) to 3 (extremely easy). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 3 Wilcoxon test (Table 82) **resulted in a significant difference in mean rankings on ease of use (usability) of the system between the DRILCOM and PALAT system** (*Z*=-3.864, $p<0.001$). Participants rated the PALAT system as being significantly easier to operate than the DRILCOM system. Figure 122 displays the mean *subjective ratings* for ease of use on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 82. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 3, ease of use (bolded text in the table indicates a significant test result at *p*<0.05.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	**-3.86**	**<0.001***	**1.39**	**1.60**	**2.67**	**0.59**

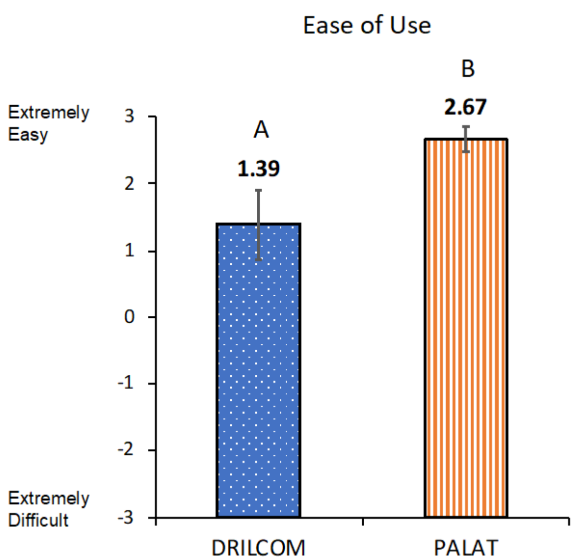


Figure 122. **Mean *subjective ratings* for Question 3, ease of use on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 3 pairwise comparison Wilcoxon test results (Table 83) **resulted in significant differences in mean rankings of ease of use between the DRILCOM and PALAT system under all listening conditions**, open ear condition (*Z*=-2.23, *p*=0.026), TEP-100 (*Z*=-2.23, *p*=0.026), and ComTac™ III (*Z*=-2.39, *p*=0.026). For all listening conditions, participants rated the PALAT system as being easier to operate than the DRILCOM system. Figure 123 displays the mean *subjective ratings* for ease of use on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 83. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 3, ease of use (bolded text in the table indicates a significant test result at *p*<0.05.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	**-2.23**	**0.026***	**1.50**	**1.57**	**2.67**	**0.65**
TEP-100	**-2.23**	**0.026***	**1.25**	**1.82**	**2.67**	**0.49**
ComTac™ III	**-2.39**	**0.017***	**1.42**	**1.51**	**2.67**	**0.65**

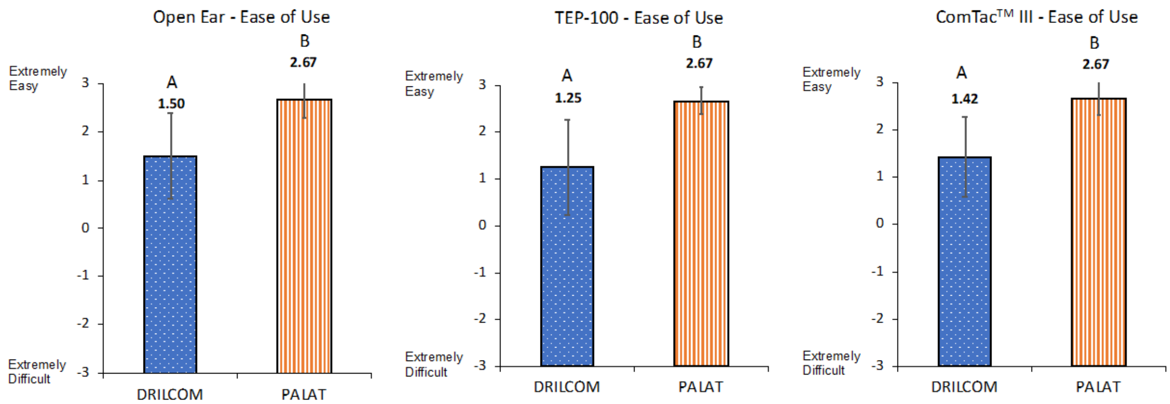


Figure 123. **Mean *subjective ratings* for Question 3, ease of use on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both training systems for ease of use. The Friedman's test for ease of use resulted in **no significant differences between mean rankings for listening conditions**, open ear (*M*=2.08, *SD*=1.32), TEP-100 (*M*=1.96, *SD*=1.49), and ComTac™ III (*M*=2.04, *SD*=1.30) across both systems (χ^2[2]=1.31, *p*=0.519).

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each training system, DRILCOM and PALAT. The Friedman's test for ease of use while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=2.67, *SD*=0.65), TEP-100 (*M*=2.67, *SD*=0.49), and ComTac™ III (*M*=2.67, *SD*=0.65), (χ^2[2]=0.00, *p*=1.00). Likewise, the Friedman's test for ease of use while using the **DRILCOM system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=1.50, SD=1.57), TEP-100 (M=1.25, SD=1.82), and ComTac™ III (M=1.42, SD=1.51), (χ^2[2]=1.83, *p*=0.401).

Participant ratings for how easy it was to operate one of the training systems resulted in significant differences between the DRILCOM and PALAT systems. Participants reported that it was significantly easier to use the PALAT system compared to DRILCOM system under all three listening conditions. There were no significant differences in the mean rankings of ease of use between listening conditions.

Question 4. Impact of the room environment on ability to train to localize sound

Participants were asked to respond to the following: "Rate the **impact** you felt the **room environment** of the DRILCOM (or PALAT) system contributed to your **ability to train to localize** sounds," from -3 (extremely negative impact) to 3 (extremely positive impact).

A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 4 Wilcoxon (Table 84) **resulted in a significant difference in mean rankings on the impact of room environment between the DRILCOM and PALAT**

system (*Z*=-2.059, *p*=0.039). The DRILCOM system room environment was rated as having a higher positive impact on contributing to the ability to localize sounds than the PALAT system. This was expected due to the absorptive hemi-anechoic DRILCOM room, which provided a more acoustically-directional environment than the office environment of the PALAT system which had hard, reflective wall surfaces. Figure 124 displays the mean ***subjective ratings*** for room effects on localization on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 84. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 4, room effects on localization (bolded text in the table indicates a significant test result at $p<0.05$.)

Source			DRILCOM (n=12)		PALAT (n=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	**-2.06**	**0.039***	**1.11**	**1.01**	**0.75**	**0.97**

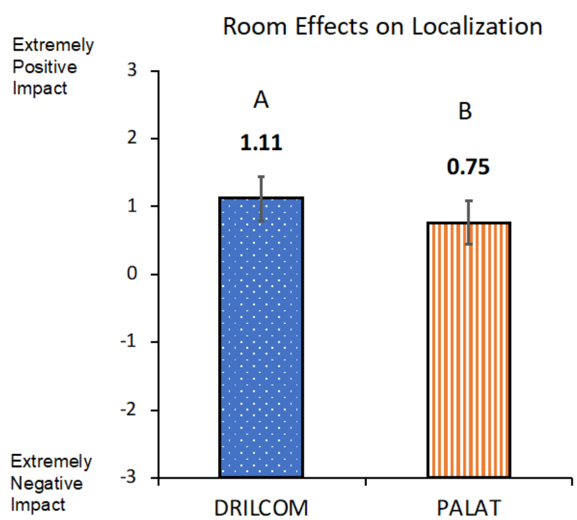


Figure 124. **Mean *subjective ratings* for Question 4, room effects on localization on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 4 pairwise comparison

Wilcoxon test results (Table 85) showed **no significant difference in mean rankings of room effects on localization between the DRILCOM and PALAT system for any of the listening conditions**, open ear (Z=-1.27, p=0.206), TEP-100 (Z=-0.37, p=0.713), and ComTac™ III (Z=-1.84, p=0.066). Figure 125 displays the mean *subjective ratings* for room effects on localization on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 85. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 1, room effects on localization (bolded text in the table indicates a significant test result at p<0.05.)

Source			DRILCOM (n=12)		PALAT (n=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-1.27	0.203	1.25	0.97	0.92	1.00
TEP-100	-0.36	0.713	1.08	1.08	1.00	1.04
ComTac™ III	-1.84	0.066	1.00	1.04	0.33	0.78

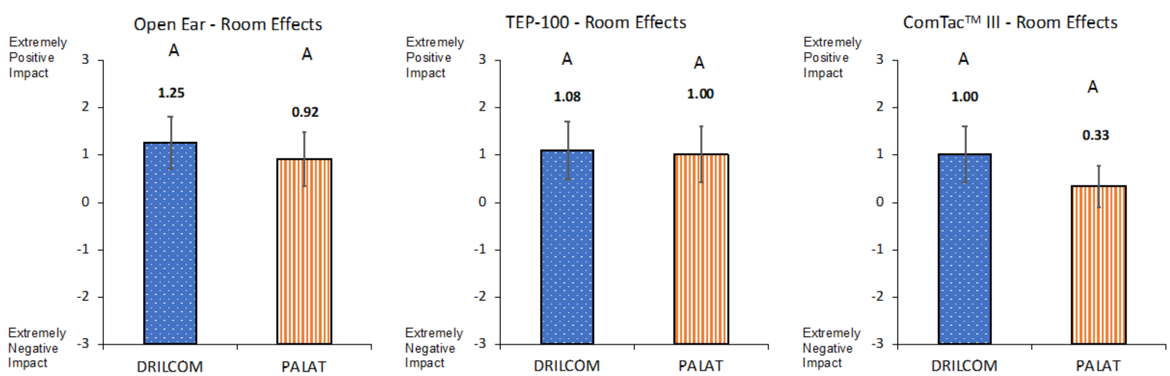


Figure 125. **Mean *subjective ratings* for Question 4, room effects on localization on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at p<0.05 within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both training systems. The Friedman's test for room effects on localization resulted in **no significant**

differences between mean rankings for listening conditions, open ear (*M*=1.08, *SD*=0.97), TEP-100 (*M*=1.04, *SD*=1.04), and ComTac™ III (*M*=0.67, *SD*=0.96) across both systems (χ^2[2]=2.46, *p*=0.292).

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each training system, DRILCOM and PALAT. The Friedman's test for room effects on localization while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions,** open ear (*M*=0.92, *SD*=1.00), TEP-100 (*M*=1.00, *SD*=1.04), and ComTac™ III (*M*=0.33, *SD*=0.78), (χ^2[2]=2.24, *p*=0.326). Likewise, the Friedman's test for room effects on localization while using the **DRILCOM system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=1.25, SD=0.97), TEP-100 (M=0.50, SD=1.17), and ComTac™ III (M=1.00, SD=1.04), (χ^2[2]=4.75, *p*=0.093).

Participant ratings for how the room environment impacted their ability to localize sounds showed a significant difference between the DRILCOM and PALAT system. However, no significant differences were found in the mean rankings of room effects on localization between training systems for any of the listening conditions. Questions 14 and 17 seemed to support participants being indifferent to the room impacts on localization. Question 14 asked, "Compared to the previous used system, please rate how much the room environment of the most recently used system impacted your ability to localize sounds," from -3 (extremely negative impact) to 3 (extremely positive impact). Ratings were identical for both systems depending on which training system was most recently used. Participants who used the DRILCOM system last slightly preferred the DRILCOM system (*M*=0.22, *SD*=0.43), while participants who used the PALAT system last slightly preferred the PALAT system (*M*=0.22, *SD*=0.65). Question 17 asked participants to rate their system preference for "room environment for training for sound localization" between the DRILCOM (-3) or PALAT (3) system. The mean ratings showed no clear preference (*M*=-0.22, *SD*=0.80).

Question 5. Training impacts on ability to localize sound

Participants were asked to respond to the following: "Rate how much you feel your **ability** to determine sound location improved as a **result of training** with this system," from -3 (extremely less capable) to 3 (extremely more capable). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 5 Wilcoxon results (Table 86) showed **no significant difference in mean rankings on localization ability improvement from training between the DRILCOM and PALAT system** (Z=-1.11, p=0.268). Figure 126 displays the mean *subjective ratings* for training impacts on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 86. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 5, training improvements (bolded text in the table indicates a significant test result at p<0.05.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	-1.11	0.268	1.50	0.91	1.64	0.76

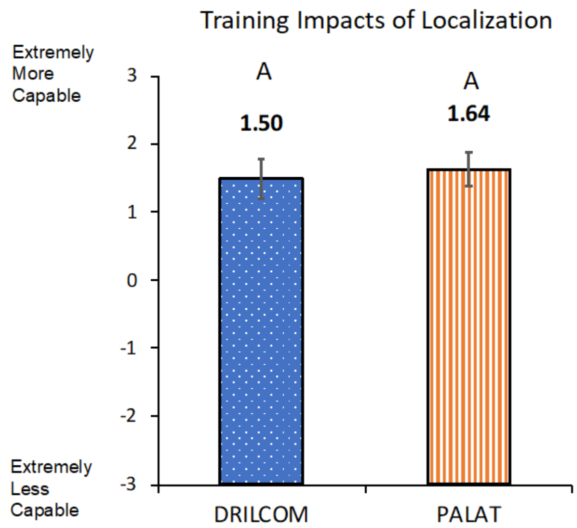


Figure 126. **Mean *subjective ratings* for Question 5, training impacts on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 5 pairwise comparison Wilcoxon test results (Table 87) showed **no significant difference in mean rankings of training impacts between the DRILCOM and PALAT system for any of the listening conditions**, open ear ($Z=-1.13$, $p=0.257$), TEP-100 ($Z=-0.71$, $p=0.480$), and ComTac™ III ($Z=-0.00$, $p=1.00$). Figure 127 displays the mean *subjective ratings* for training impacts on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 87. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 5, training impacts (bolded text in the table indicates a significant test result at $p<0.05$.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-1.13	0.257	1.58	0.80	1.83	0.84
TEP-100	-0.71	0.480	1.25	1.22	1.42	0.79
ComTac™ III	0.00	1.000	1.67	0.65	1.67	0.65

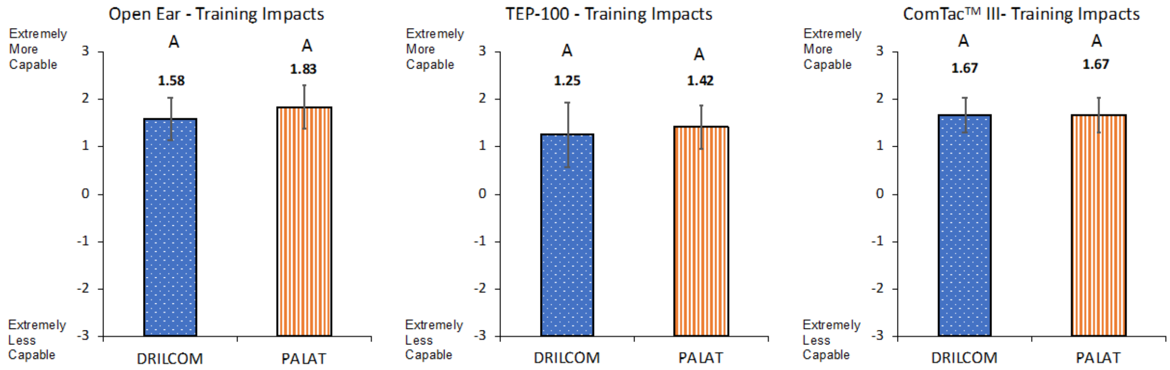


Figure 127. **Mean _subjective ratings_ for Question 5, training impacts on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both training systems. The Friedman's test for training impacts resulted in **no significant differences between mean rankings for listening conditions**, open ear (M=1.71, SD=0.81), TEP-100 (M=1.33, SD=1.01), and ComTac™ III (M=1.67, SD=0.64) across both systems ($\chi^2[2]=0.58$, p=0.748).

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each training system, DRILCOM and PALAT. The Friedman's test for training impacts while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (M=1.83, SD=0.84), TEP-100 (M=1.42, SD=0.79), and ComTac™ III (M=1.67, SD=0.65), ($\chi^2[2]=1.00$, p=0.607). Likewise, the Friedman's test for training impacts while using the **DRILCOM system resulted in no significant differences**

between mean rankings for listening conditions, open ear (*M*=1.58, SD=0.79), TEP-100 (M=1.25, SD=1.22), and ComTac™ III (M=1.67, SD=0.65), (χ^2[2]=0.26, *p*=0.879).

Participant ratings for how training improved their ability to localize sounds showed no significant differences between the DRILCOM and PALAT systems. In addition, there were no significant differences in the mean rankings of training impacts between listening conditions. These findings were supported by the mean participant responses to Question 13 and Question 20. Question 13 asked, “Compared to the previously used system, rate how much of an impact training with the most recently used system had on your ability to localize sounds,” from -3 (extremely negative impact) to 3 (extremely positive impact). Participants who used the DRILCOM system last slightly preferred the DRILCOM system (*M*=0.50, *SD*=0.79), while participants who used the PALAT system last slightly preferred the PALAT system (*M*=0.56, *SD*=1.34). Question 20 asked participants to rate their system preference for “confidence in the benefits achieved with the training for sound localization,” between the DRILCOM (-3) or PALAT (3) system. The mean ratings showed no clear preference (*M*=0.06, *SD*=1.22).

Question 6. Difficulty in judging the signal location

Participants were asked to respond to the following: “Rate how **difficult** it was to judge the **location** of the sounds **using this system**,” from -3 (extremely difficult) to 3 (extremely easy). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 6 Wilcoxon results (Table 88) showed **no significant difference in mean rankings on difficulty judging signal location between the DRILCOM and PALAT system** (*Z*=-1.03, *p*=0.301). Figure 128 displays the mean *subjective ratings* for difficulty judging signal location on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 88. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 6, difficulty judging signal location (bolded text in the table indicates a significant test result at *p*<0.05.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	-1.03	0.301	0.28	1.67	0.03	1.65

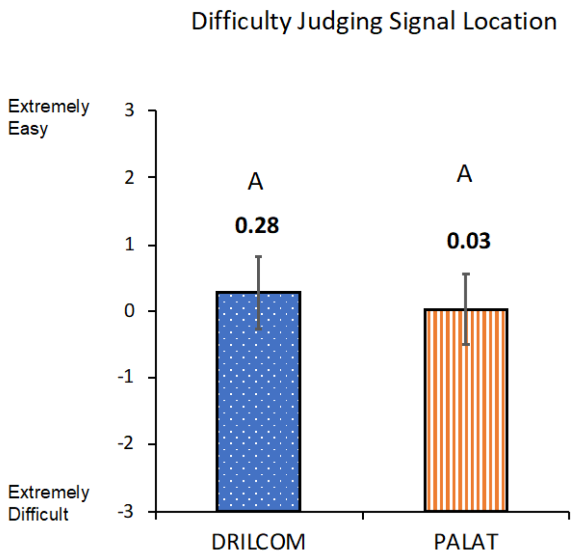


Figure 128. **Mean *subjective ratings* for Question 6, difficulty judging signal location on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 6 pairwise comparison Wilcoxon test results (Table 89) showed **no significant difference in mean rankings of difficulty judging signal location between the DRILCOM and PALAT system for any of the listening conditions**, open ear (*Z*=-1.42, *p*=0.155), TEP-100 (*Z*=-0.19, *p*=0.852), and ComTac™ III (*Z*=-0.32, *p*=0.748). Figure 129 displays the mean *subjective ratings* for difficulty judging signal location on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 89. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 6, difficulty judging signal location (bolded text in the table indicates a significant test result at *p*<0.05.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-1.42	0.155	1.58	1.24	0.92	1.38
TEP-100	-0.19	0.852	-0.75	1.49	-0.75	1.77
ComTac™ III	-0.32	0.748	0.00	1.41	-0.08	1.44

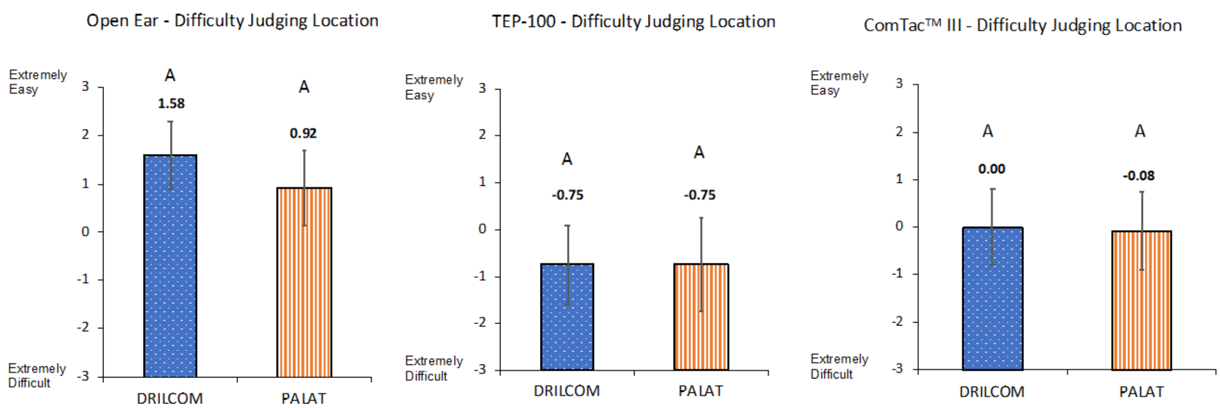


Figure 129. **Mean _subjective ratings_ for Question 6, difficulty judging signal location on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both localization training systems. The Friedman's test for difficulty judging signal location resulted in **a significant difference between mean rankings for listening conditions**, open ear (*M*=1.25, *SD*=1.33), TEP-100 (*M*=-0.75, *SD*=1.60), and ComTac™ III (*M*=-0.04, *SD*=1.40) across both systems (χ^2[2]=13.93, *p*=0.001). Post hoc pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare each listening condition. Pairwise comparisons used a Bonferroni adjusted α level of 0.017. The Wilcoxon test results (Table 90) showed **a significant difference in mean rankings of difficulty judging signal location between open ear versus TEP-100** (Z=-3.45, p=0.001)**, and the open ear versus ComTac™ III** (Z=-3.37, p=0.001). **No significant difference was found between TEP-100 and ComTac™ III** (Z=-1.90, p=0.057). Figure 130 displays the mean ***subjective ratings*** for difficulty judging signal location between each listening condition with 95% confidence interval error bars.

Table 90. Wilcoxon results pairwise comparisons between each listening condition for Question 6, difficulty judging signal location (bolded text in the table indicates a significant test result at p<0.017.)

Listening Condition	Z	*p*
Open - TEP 100	**-3.45**	**0.001***
Open - ComTac™ III	**-3.37**	**0.001***
TEP-100 – ComTac™ III	-1.90	0.057

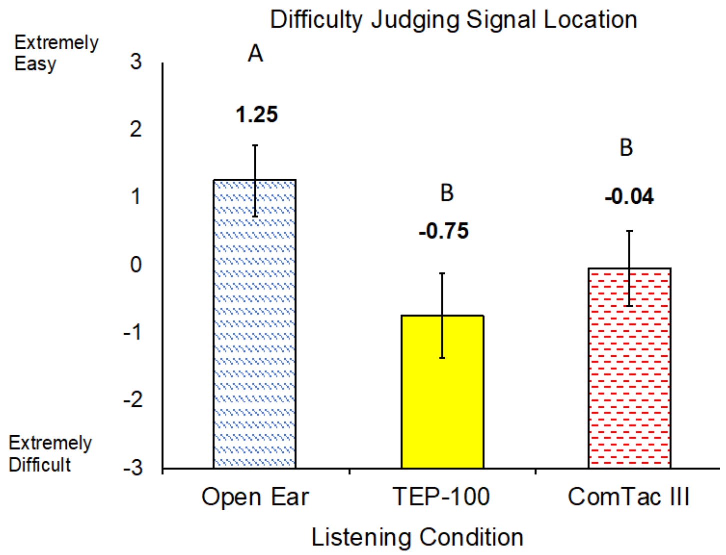


Figure 130. **Mean *subjective ratings* for Question 6, difficulty judging signal location between each listening condition**, with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at p<0.05 within each listening condition per a Wilcoxon signed-rank test.

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each training system, DRILCOM and PALAT. The Friedman's test for difficulty judging signal location while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=0.92, *SD*=1.38), TEP-100 (*M*=-0.75, *SD*=1.77), and ComTac™ III (*M*=-0.08, *SD*=1.44), (χ^2[2]=5.286, *p*=0.071).

The Friedman's test for difficulty judging signal location while using the **DRILCOM system resulted in a significant difference between mean rankings for listening conditions**, open ear (*M*=1.58, SD=1.24), TEP-100 (M=-0.75, SD=1.49), and ComTac™ III (M=0.00, SD=1.41), (χ^2[2]=9.05, *p*=0.011). Post hoc Wilcoxon signed-rank test pairwise comparisons with Bonferroni correction (α= 0.017) results (Table 91) showed **a significant difference in mean rankings of difficulty judging signal location on the DRILCOM system between open ear versus TEP-100** (*Z*=-2.50, *p*=0.012)**, and the open ear versus ComTac™ III** (*Z*=-2.69, *p*=0.001). No significant difference was found between TEP-100 and ComTac™ III(*Z*=-1.24, *p*=0.216). Figure 131 displays the mean *subjective ratings* for difficulty judging signal location by training system between each listening condition with 95% confidence interval error bars.

Table 91. Wilcoxon pairwise comparisons between each listening condition on DRILCOM system for Question 6, difficulty judging signal location (bolded text in the table indicates a significant test result at *p*<0.05.)

System	Listening Condition	Z	*p*
DRILCOM			
	Open - TEP 100	**-2.50**	**0.012***
	Open - ComTac™ III	**-2.69**	**0.007***
	TEP-100 – ComTac™ III	-1.24	0.216

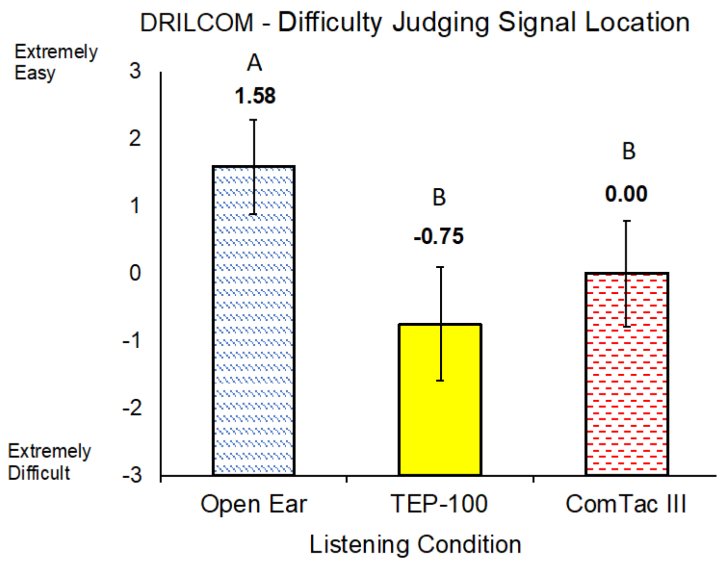


Figure 131. **Mean *<u>subjective ratings</u>* for Question 6, difficulty judging signal location between each listening condition**, with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ within each listening condition per a Wilcoxon signed-rank test.

<u>Question 7. Impact on reaction time before to after training</u>

Participants were asked to respond to the following: "Rate how **training** using the DRILCOM (or PALAT) system **impacted your reaction time** in determining sound location, from **before to after** all the training you received using this system," from -3 (extremely slower reaction time) to 3 (extremely faster reaction time). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 7 Wilcoxon results (Table 92) showed **no significant difference in mean rankings on reaction time between the DRILCOM and PALAT system** ($Z=-1.58$, $p=0.114$). Figure 132 displays the mean *<u>subjective ratings</u>* for training impacts on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 92. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 7, reaction time (bolded text in the table indicates a significant test result at *p*<0.05.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	-1.58	0.114	0.92	1.03	1.22	0.80

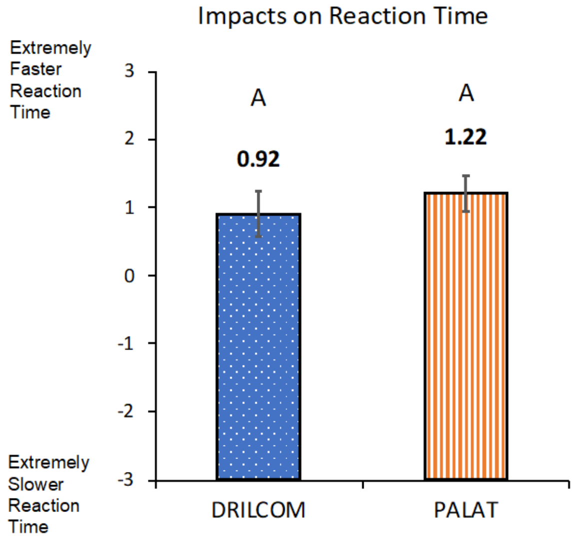


Figure 132. **Mean *subjective ratings* for Question 7, impact on reaction time on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 7 pairwise comparison Wilcoxon test results (Table 93) showed **no significant difference in mean rankings of impact on reaction time between the DRILCOM and PALAT system for the open ear** (*Z*=-0.52, *p*=0.603) and **ComTac™ III** (*Z*=-0.00, *p*=1.00). There was a **significant difference in mean rankings of impact on reaction time between the DRILCOM and PALAT system for the TEP-100** (*Z*=-1.98, *p*=0.047). Participant ratings indicated that there was a faster perceived reaction time while using the PALAT system. Figure 133 displays the mean *subjective ratings* for impact on reaction time on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 93. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 7, impact on reaction time (bolded text in the table indicates a significant test result at *p*<0.05.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-0.52	0.603	1.33	0.99	1.50	0.80
TEP-100	**-1.98**	**0.047***	**0.42**	**1.08**	**1.17**	**0.94**
ComTac™ III	0.00	1.000	1.00	0.85	1.00	0.60

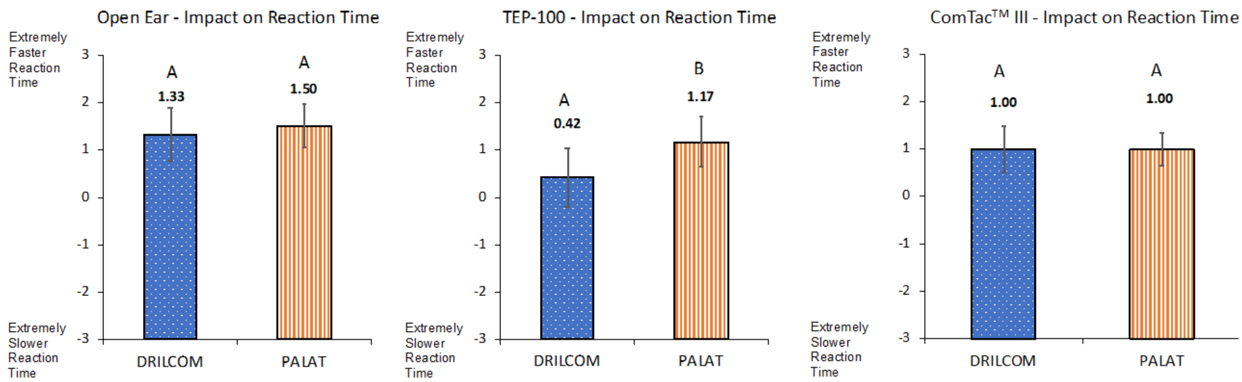


Figure 133. **Mean *subjective ratings* for Question 7, impact on reaction time on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both training systems. The Friedman's test for impacts on reaction time resulted in **a significant difference between mean rankings for listening conditions**, open ear (*M*=1.42, *SD*=0.88), TEP-100 (*M*=0.79, *SD*=1.06), and ComTac™ III (*M*=1.00, *SD*=0.72), across both systems (χ^2[2]=8.25, *p*=0.016). Post hoc pairwise comparisons using Wilcoxon signed-rank tests (Table 94) showed that participants perceived a faster reaction time under open ear than with either of the TCAPS devices. Using a Bonferroni adjustment with α= 0.017 resulted in none of the pairwise comparisons meeting the threshold of significance. However, the difference in mean rankings of open ear versus TEP-100 was extremely close to significant (*Z*=-2.38, *p*=0.018) and open ear versus ComTac™ III was close to significant (*Z*=-2.14, *p*=0.032). There was no difference in mean rankings between the TEP-100 and ComTac™ III (*Z*=-0.85, *p*=0.398). Figure 134 displays the mean *subjective ratings* for impact on reaction time between each listening condition with 95% confidence interval error bars.

Table 94. Wilcoxon results pairwise comparisons between each listening condition for Question 7, impact on reaction time (bolded text in the table indicates a significant test result at $p<0.05$ with Bonferroni correction α= 0.017.)

Listening Condition	Z	*p*
Open - TEP 100	-2.37	0.018
Open - ComTac™ III	-2.14	0.032
TEP-100 – ComTac™ III	-0.85	0.398

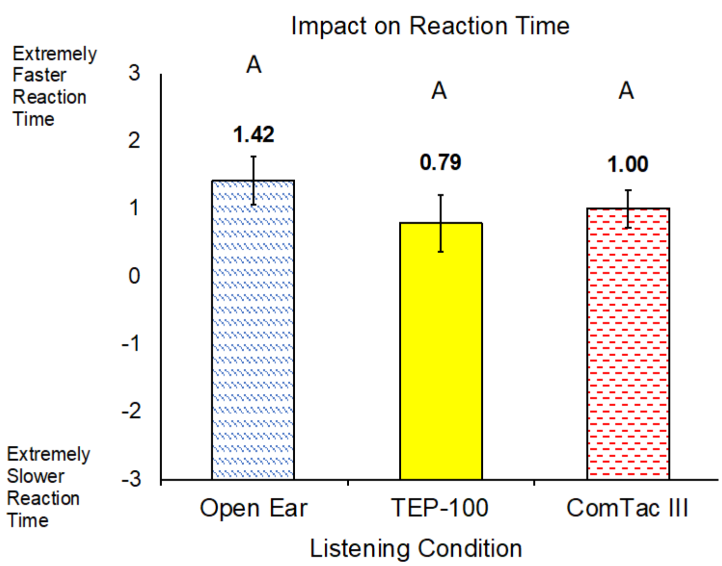


Figure 134. **Mean *subjective ratings* for Question 7, impact on reaction time between each listening condition**, with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ within each listening condition per a Wilcoxon signed-rank test.

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each training system, DRILCOM and PALAT. The Friedman's test for impact on reaction time while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=1.50, *SD*=0.798), TEP-100 (*M*=1.17, *SD*=0.937), and ComTac™ III (*M*=1.00, *SD*=0.603), (χ^2[2]=4.35, *p*=0.114). Likewise, the Friedman's test for impact on reaction time while using the **DRILCOM system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=1.33, SD=0.985), TEP-100 (M=0.42, SD=1.08), and ComTac™ III (M=1.00, SD=0.853), (χ^2[2]=5.72, *p*=0.057).

Participant ratings for the impact of training with a particular system on reaction time showed no significant differences between the DRILCOM and PALAT systems except for under the TEP-100 condition where participants perceived reaction time as faster using the PALAT system. These findings were supported by the mean participant responses to Question 16. Question 16 asked participants to rate their system preference for "confidence in making quick decisions (reaction time) about the location of the sounds," between the DRILCOM (-3) or PALAT (3) system. The mean ratings showed a slight preference toward the PALAT system but a standard deviation that showed mixed preference from participants (*M*=0.56, *SD*=1.34).

Question 8. Impact of the user interface on ability to train to localize sound

Participants were asked to respond to the following: "Rate how much of an **impact** the DRILCOM (PALAT) system **user interface** (monitor, software, loudspeakers, wires, etc.) had on your **ability to train your sound localization skills**," from -3 (extremely negative impact) to 3 (extremely positive impact). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 8 Wilcoxon test (Table 95) resulted in a **significant difference in mean rankings on the impact of the user interface on ability to train localization between the DRILCOM and PALAT systems** (*Z*=-3.29, *p*=0.001). Participants rated the PALAT system user interface has having a more positive impact on the ability to localize sounds than the DRILCOM system. Figure 135 displays the mean ***subjective ratings*** for impact of user interface on localization on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 95. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 8, User interface (bolded text in the table indicates a significant test result at $p<0.05$.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	**-3.29**	**0.001***	**0.69**	**1.22**	**1.50**	**0.85**

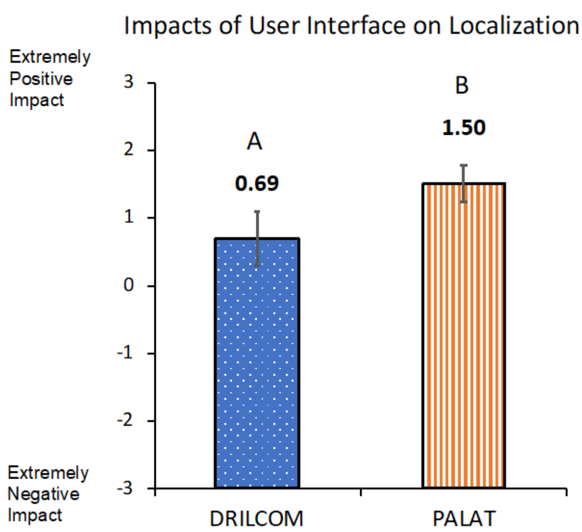


Figure 135. **Mean *subjective ratings* for Question 8, impact of user interface on localization on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 8 pairwise comparison Wilcoxon test results (Table 96) **resulted in significant differences in mean rankings of impact of user interface on localization between the DRILCOM and PALAT system under both TCAPS devices**, TEP-100 (Z=-2.21, p=0.027) and ComTac™ III (Z=2.11, p=0.035). Under both TCAPS listening conditions, participants rated the PALAT system user interface as having a more positive impact on localization than the DRILCOM system user interface. There was no significant difference in mean rankings of impact of user interface on localization between the DRILCOM and PALAT system under the open ear condition (Z=-1.57, p=0.117). Figure 136 displays the mean *subjective ratings* for impact of user interface on localization on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 96. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 8, impact of user interface on localization (bolded text in the table indicates a significant test result at p<0.05.)

Source			DRILCOM (n=12)		PALAT (n=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-1.57	0.117	0.75	1.29	1.50	0.91
TEP-100	**-2.21**	**0.027***	**0.67**	**1.16**	**1.67**	**1.07**
ComTac™ III	**-2.11**	**0.035***	**0.67**	**1.30**	**1.33**	**0.49**

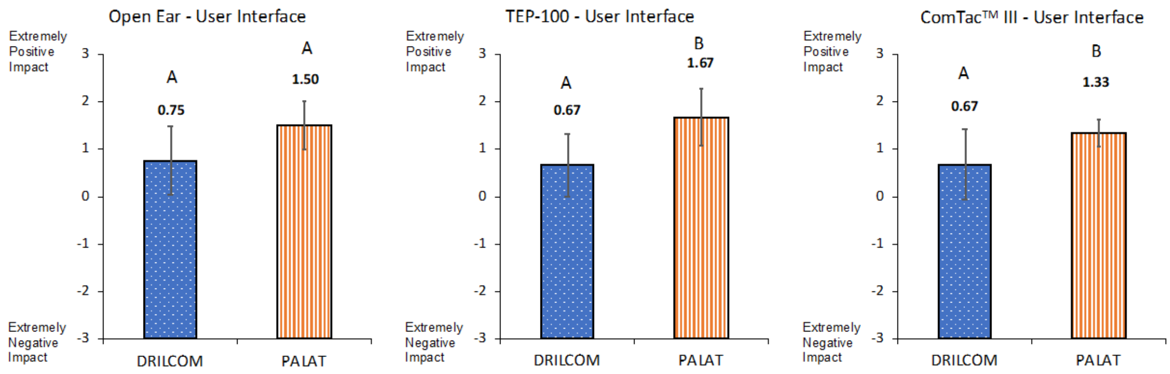


Figure 136. **Mean *subjective ratings* for Question 8, impact of user interface on localization on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both training systems for impact of user interface on localization. The Friedman's test for impact of user interface on localization resulted in **no significant differences between mean rankings for listening conditions**, open ear (*M*=1.13, *SD*=1.15), TEP-100 (*M*=1.17, *SD*=1.20), and ComTac™ III (*M*=1.00, *SD*=1.02) across both systems ($\chi^2[2]=0.419$, $p=0.811$).

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each training system, DRILCOM and PALAT. The Friedman's test for impact of user interface on localization while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=1.50, *SD*=0.91), TEP-100 (*M*=1.67, *SD*=1.07), and ComTac™ III (*M*=1.33, *SD*=0.49), ($\chi^2[2]=1.52$, $p=0.469$). Likewise, the Friedman's test for impact of user interface on localization while using the **DRILCOM system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=0.75, SD=1.29), TEP-100 (M=0.67, SD=1.16), and ComTac™ III (M=0.67, SD=1.30), (χ^2[2]=0.21, *p*=0.902).

Participant ratings for the impact of user interface on ability to localize sounds resulted in significant differences between the DRILCOM and PALAT systems. Participants reported that the PALAT system user interface had a significantly more positive impact on the ability to localize sounds compared to DRILCOM system user interface. These findings were supported by the mean participant responses to Question 12 and Question 19. Question 12 asked, "Compared to the previously used system, rate how much the user interface (computer, software, loudspeakers, etc.) of the most recently used system impacted your ability to train to localize sounds," from -3 (extremely negative impact) to 3 (extremely positive impact). Participants who used the PALAT system last preferred the PALAT system (*M*=1.17, *SD*=1.15), while participants who used the DRILCOM system last *still* slightly preferred the PALAT system (*M*=-0.61, *SD*=1.09). Question 19 asked participants to rate their system preference as to the "user interface for responding to the location of the sound" between the DRILCOM (-3) or PALAT (3) system. The mean ratings showed participants preferred the PALAT system user interface (*M*=1.83, *SD*=1.11).

Question 9. Impact of room environment on reaction time

Participants were asked to respond to the following: "Rate how training in the **room environment** of the DRILCOM (or PALAT) system **impacted your reaction time** in determining sound location," from -3 (extremely slower reaction time) to 3 (extremely faster reaction time). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 9 Wilcoxon results (Table 97) showed **no significant difference in mean rankings on the impact of room environment on reaction time between the DRILCOM and PALAT system** (Z=-0.69, p=0.491). Figure 137 displays the mean ***subjective ratings*** for the impact of room environment on reaction time on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 97. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 9, Room environment impact on reaction time (bolded text in the table indicates a significant test result at p<0.05.)

Source			DRILCOM (n=12)		PALAT (n=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	-0.69	0.491	0.58	1.00	0.53	0.65

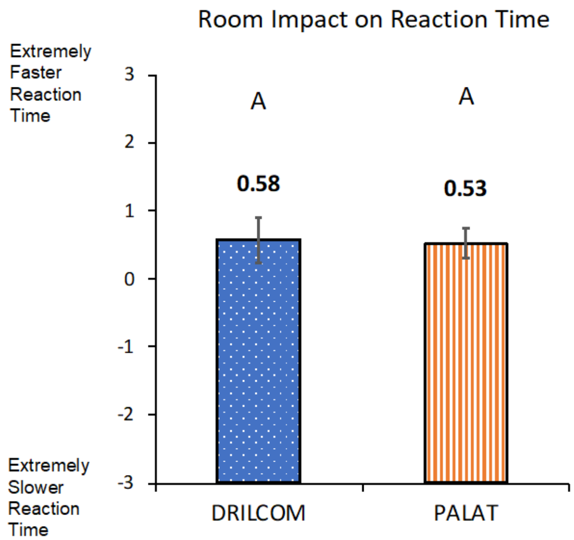


Figure 137. **Mean *subjective ratings* for Question 9, room environment impacts on reaction time on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at p<0.05 per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 9 pairwise comparison Wilcoxon test results (Table 98) showed **no significant difference in mean rankings of room**

environment impact on reaction time between the DRILCOM and PALAT system for all listening conditions, open ear (*Z*=-0.82, *p*=0.414), TEP-100 (*Z*=0.00, *p*=1.000), and ComTac™ III (*Z*=-1.89, *p*=0.059). Figure 138 displays the mean *subjective ratings* for room environment impact on reaction time on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 98. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 9, room environment impact on reaction time (bolded text in the table indicates a significant test result at *p*<0.05.)

Source			DRILCOM (*n*=12)		PALAT (*n*=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-0.82	0.414	0.50	1.00	0.67	0.65
TEP-100	0.00	1.000	0.50	1.17	0.58	0.79
ComTac™ III	-1.89	0.059	0.75	0.87	0.33	0.49

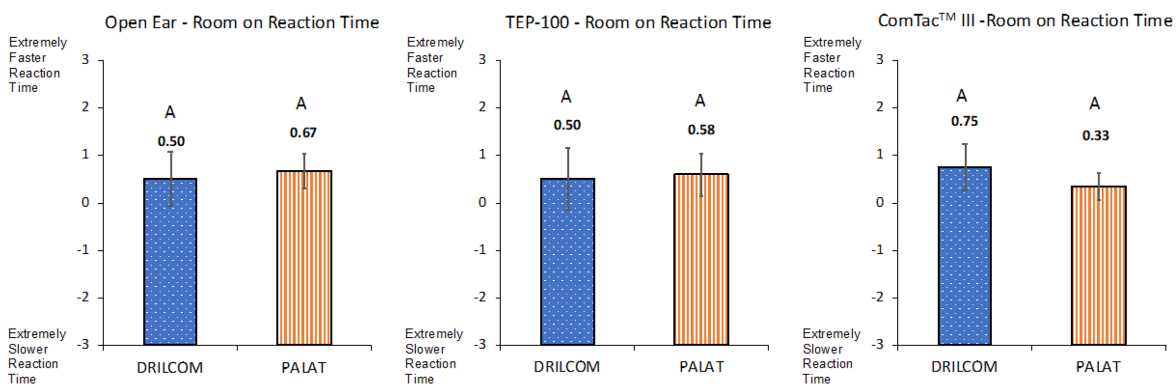


Figure 138. **Mean *subjective ratings* for Question 9, room environment impact on reaction time on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at *p*<0.05 within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both localization training systems. The Friedman's test for room environment impact on reaction time resulted in **no significant differences between mean rankings for listening conditions**, open ear (*M*=0.58, *SD*=0.83), TEP-100 (*M*=0.54, *SD*=0.98), and ComTac™ III (*M*=0.54, *SD*=0.72) across both systems (χ^2[2]=0.95, *p*=0.622).

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each localization training systems, DRILCOM and PALAT. The Friedman's test for room environment impact on reaction time while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=0.67, *SD*=0.65), TEP-100 (*M*=0.58, *SD*=0.79), and ComTac™ III (*M*=0.33, *SD*=0.49), (χ^2[2]=4.33, *p*=0.115). Likewise, the Friedman's test for room environment impact on reaction time while using the **DRILCOM system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=0.50, SD=0.1.00), TEP-100 (M=0.50, SD=1.17), and ComTac™ III (M=0.75, SD=0.87), (χ^2[2]=0.64, *p*=0.727).

Question 10. Impact of loudspeaker visibility on ability to train to localize

Participants were asked to respond to the following: "Rate the **impact** you felt the **hidden loudspeakers** of the DRILCOM system (visible loudspeakers of the PALAT system) contributed to your **ability to train to localize** sounds," from -3 (extremely negative impact) to 3 (extremely positive impact). A Wilcoxon signed-rank test was used to compare participant ratings of using the DRILCOM system with the ratings of using the PALAT system during training and testing under all three listening conditions. Question 10 Wilcoxon test (Table 99) resulted in **a significant difference in mean rankings on the impact of loudspeaker visibility on ability to localize sounds between the DRILCOM and PALAT system** (*Z*=-2.14, *p*=0.032). Participants rated the PALAT system loudspeaker visibility has having a more positive impact on the ability to localize sounds than the DRILCOM system. Figure 139 displays the mean *subjective ratings* for impact of loudspeaker visibility on localization on the DRILCOM and PALAT systems across all listening conditions with 95% confidence interval error bars.

Table 99. Wilcoxon results comparing ratings for DRILCOM versus PALAT across all listening conditions for Question 10, Loudspeaker visibility impact of localization (bolded text in the table indicates a significant test result at $p<0.05$.)

Source			DRILCOM (n=12)		PALAT (n=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
DRILCOM vs. PALAT	**-2.14**	**0.032**	**0.28**	**0.91**	**0.78**	**0.90**

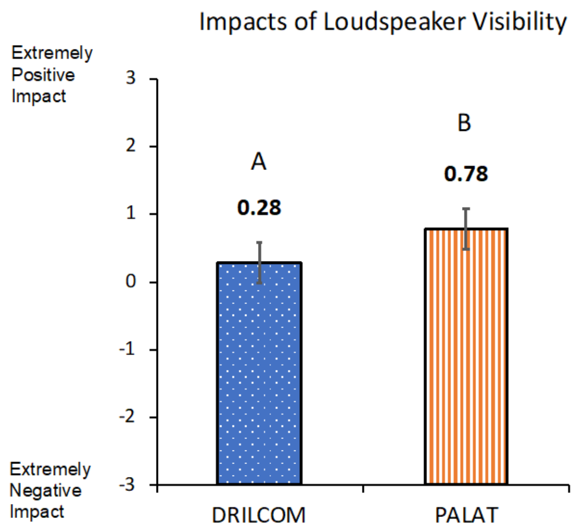


Figure 139. **Mean *subjective ratings* for Question 10, impact of loudspeaker visibility on localization on the DRILCOM and PALAT systems across all listening conditions** with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at $p<0.05$ per a Wilcoxon signed-rank test.

Pairwise comparisons were conducted using Wilcoxon signed-rank tests to compare DRILCOM versus PALAT under each listening condition. Question 10 pairwise comparison Wilcoxon test results (Table 100) **resulted in no significant differences in mean rankings of impact of loudspeaker visibility on localization between the DRILCOM and PALAT systems for all listening conditions**, open ear (Z=-1.51, p=0.132). TEP-100 (Z=-1.27, p=0.204) and ComTac™ III (Z=-0.91, p=0.366). Figure 140 displays the mean *subjective ratings* for impact of loudspeaker visibility on localization on the DRILCOM and PALAT systems for each listening condition with 95% confidence interval error bars.

Table 100. Wilcoxon results comparing ratings for DRILCOM versus PALAT for each listening condition for Question 10, impact of loudspeaker visibility on localization (bolded text in the table indicates a significant test result at p<0.05.)

Source			DRILCOM (n=12)		PALAT (n=12)	
	Z	*p*	*M*	*SD*	*M*	*SD*
Open ear	-1.51	0.132	0.42	0.90	1.00	0.85
TEP-100	-1.27	0.204	0.17	1.03	0.83	1.19
ComTac™ III	-0.91	0.366	0.25	0.87	0.50	0.52

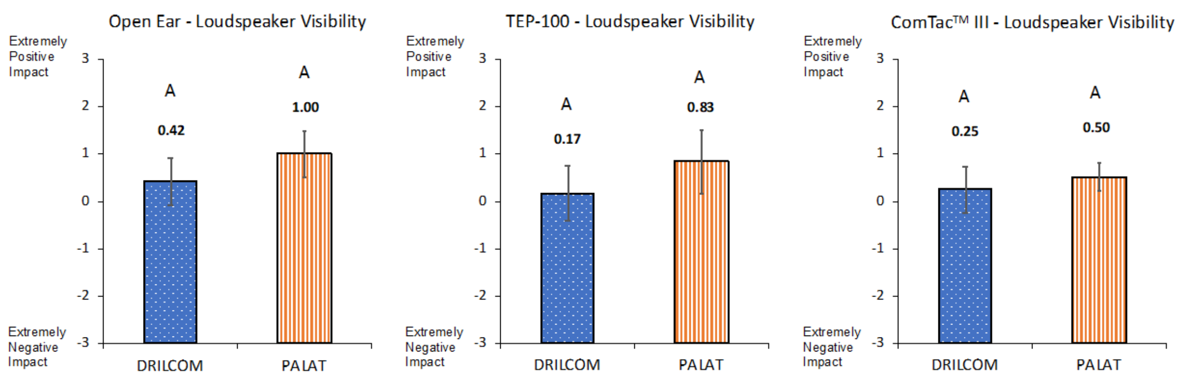


Figure 140. **Mean *subjective ratings* for Question 10, impact of loudspeaker visibility on localization on the DRILCOM and PALAT systems for each listening condition**, open ear (left), TEP-100 (middle), and ComTac™ III (right), with 95% confidence intervals plotted around the mean values on each bar. Different letters indicate significant differences at p<0.05 within each listening condition per a Wilcoxon signed-rank test.

A Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, across both training systems for impact of loudspeaker visibility on localization. The Friedman's test for impact of loudspeaker visibility on localization resulted in **no significant differences between mean rankings for listening conditions**, open ear (M=0.71, SD=0.91), TEP-100 (M=0.50, SD=1.14), and ComTac™ III (M=0.38, SD=0.71) across both systems (χ^2[2]=1.57, p=0.457).

Lastly, a Friedman non-parametric test was performed to detect differences between mean subjective ratings by listening condition, open ear, TEP-100, and ComTac™ III, for each training system, DRILCOM and PALAT. The Friedman's test for impact of loudspeaker visibility on localization while using the **PALAT system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=1.00, *SD*=0.85), TEP-100 (*M*=0.83, *SD*=1.19), and ComTac™ III (*M*=0.50, *SD*=0.52), (χ^2[2]=2.48, *p*=0.289). Likewise, the Friedman's test for impact of loudspeaker visibility on localization while using the **DRILCOM system resulted in no significant differences between mean rankings for listening conditions**, open ear (*M*=0.42, SD=0.90), TEP-100 (M=0.17, SD=1.03), and ComTac™ III (M=0.25, SD=0.87), (χ^2[2]=0.00, *p*=1.000).

3.7 Phase II Discussion and Conclusions

Conclusions Overview

The results of the Phase II investigation supported the hypotheses that a Portable Auditory Localization Acclimation Training (PALAT) system could be developed using primarily off-the-shelf components, operated by a trainee, and provide acoustically-accurate localization cues that imparted training benefits in a non-laboratory indoor environment that demonstrated similar training benefit to that imparted by the full-scale DRILCOM-based system in the same number of learning unit (LU) sessions (that is, a total of 5 LUs). Multiple analysis of variance tests showed no significant difference between the dependent measures of absolute correct response scores and front-back reversal errors while using the DRILCOM and PALAT systems. While from a statistical inference standpoint the null hypothesis cannot be proven true, meaning it cannot be proven that the DRILCOM and PALAT systems are equal, multiple statistical tests failed to show differences in participants' localization performance while using the DRILCOM and PALAT systems, so from a training effectiveness standpoint, there is no demonstrated difference between the systems. Again, the objective of the PALAT system was to provide similar auditory localization training benefits as the DRILCOM system. As a result, the investigator concluded that the PALAT system is capable of providing the necessary localization training benefits in a non-laboratory (i.e., office or barracks) environment to increase military service members' situation awareness. This was further evidenced in the ensuing Phase III research, which entailed a field (outdoor) validation of the PALAT training against gunshot localization in the field.

In addition to providing similar training benefit, results supported that training on the PALAT system was sensitive to auditory localization performance differences with the open ear and with an in-the-ear and over-the-ear TCAPS. Multiple ANOVA tests found significant differences between the absolute score and front-back reversal errors between listening conditions. Analyses showed no significant interactions between localization training system and listening condition indicating that the PALAT system was able to detect localization performance differences between listening conditions similar to the DRILCOM system. One-way repeated-measures ANOVAs confirmed participants performed best on the localization accuracy measures under the open ear condition and worst under the TEP-100 condition.

While the PALAT system was designed to impart similar auditory localization training effects, improvements were made to the user interface and software to allow trainees to operate the system without an experimenter. As a result of the improved system, the investigator expected to observe differences in response times and participant system preference between the DRILCOM and PALAT systems. Results showed significantly faster response times while using the PALAT system compared to the DRILCOM system. In addition, participants indicated a significant preference toward the PALAT system user interface and rated the PALAT system easier to operate than the DRILCOM system.

The following sections discuss the findings from the Phase II experiment and the impacts of the PALAT system on auditory localization training. Data discussed under each section cover findings from all dependent measures under each subheading.

Similar Training Benefit between PALAT and DRILCOM: Objective and Subjective Measures

As noted above, while it is not possible to prove the null hypothesis, the study found no significant differences between the PALAT and DRILCOM systems at every stage of the investigation on absolute score or front-back reversal errors. There were multiple design differences between the DRILCOM and PALAT systems that could have affected auditory localization performance. The DRILCOM system consisted of 12 azimuthal loudspeakers that were hidden under acoustically transparent cloth and only presented 12 loudspeaker response options on the participant training screen. In contrast, the PALAT system consisted of 24 azimuthal loudspeakers that were clearly visible along with 24 loudspeaker response options on the participant training screen. The larger diameter of the DRILCOM system equated to a 0.78 meter (2.6 foot) separation between loudspeaker placement from the participant’s point of view, whereas, the PALAT system loudspeaker were separated by 0.52 meters (1.7 feet) with a loudspeaker located directly in between each of the 12 loudspeakers used during the study (Figure 53). While the azimuthal angular separation was consistent between the two systems at 30-degrees between actual signal-producing loudspeakers in the experiment, the smaller visual separation of the PALAT loudspeakers could make it more challenging for the participant to accurately determine the signal location.

The room environment was also very different between the two systems. The DRILCOM system was located in a hemi-anechoic room with no windows and the participant faced toward a large black acoustically transparent cloth providing very little distraction during training and testing. The PALAT system was located in an office room with large windows covered by metal blinds, a white board directly in front of the participant, and several large pieces of metal and wood office furniture located around the room (Figure 71); this was intended to represent a typical field testing environment. Acoustic testing of the reverberation time in the PALAT and DRILCOM rooms showed a higher RT60 time in the PALAT office room at the 500 Hz to 8000 Hz 1/3 octave-band frequencies as well as a higher noise floor in the 250 Hz and 500 Hz 1/3 octave-band frequencies (Table 16). In addition, the smaller PALAT loudspeakers were not able to produce the same sound pressure level of the dissonant tonal signal at the lowest 104 Hz frequency (Figure 60).

Despite substantial differences in design features and room environment appearance and acoustics, as well as small differences in test signal spectra at the lowest dissonant signal frequency, mean absolute correct responses for each listening condition at LU0 pretest and LU5 posttest resulted in no significant differences between the two localization training systems. Participants, on average, did perform slightly better during the pretest on the DRILCOM system under the open ear condition (2.2% better), TEP-100 condition (6.3% better) and ComTac™ III condition (7.6% better). However, the absolute score deltas either remained consistent or shrank by the posttest, open ear condition (2.3% better), TEP-100 condition (7.1% better) and ComTac™ III condition (0.0% better). The consistent performance was supported by the non-significant findings for interaction effects of training systems and stage of training (LU0 and LU5). Similar performance was also shown with no significant difference in the training effect as measured by regression line slopes of absolute score from LU0 through LU5 between the DRILCOM and PALAT systems collapsed across all listening conditions and under each listening condition. Auditory localization performance improved at a similar rate from LU0 to LU5 on the DRILCOM and PALAT systems for each listening condition; open ear mean absolute score improved by 11% on both systems, TEP-100 mean absolute score improved by 20% on DRILCOM and 19% on PALAT, and ComTac^{TM} III mean absolute score improved by 11% on DRILCOM and 19% on PALAT (Figure 80). In effect, these data show that the PALAT system provided more training benefit for the over-the-ear TCAPS than the DRILCOM systems, and equivalent benefit for the in-the-ear TCAPS and open ear.

In addition to absolute correct response score, participant front-back reversal errors and azimuthal accuracy was consistent between the DRILCOM and PALAT systems. No significant differences were found on mean front-back reversal errors at LU0 or LU5 between DRILCOM and PALAT collapsed across all listening conditions and within each listening condition. In addition, there was a strong correlation between the DRILCOM and PALAT systems on percent accuracy by loudspeaker location indicating that participants performed just as well at certain target locations and had difficulty judging sound direction for other target locations at a similar rate at each loudspeaker location. As depicted in Figure 106, the overlapping lines on the radial plots for each listening condition show that participants were consistently accurate by loudspeaker location before and after training, LU0 and LU5, between the two training systems.

Participant responses to the questionnaires showed that the DRILCOM and PALAT systems were also perceived to provide similar auditory localization training benefit. No significant difference was found for perceived confidence in the *ability to localize the sound* between the two systems for each listening condition. When asked to directly compare the two systems on *perceived confidence,* the mean participant rating indicated no preference for either system. Questionnaire results also showed no significant difference in *training impacts by system* or *impacts of proximity of loudspeakers effect on ability to localize* between the DRILCOM and PALAT systems. Response scores did indicate a significant difference in *perceived room environment impacts on ability to localize* the sound with a higher positive impact score for the DRILCOM room (Question 4). However, two follow-up questions (Questions 14 and 17) seemed to show there was not a strong difference between the two room environments when it comes to localizing sounds. Question 17 asked participants to rate their system preference for *"room environment for training for sound localization"* between the DRILCOM (-3) or PALAT (3) system and found only a slight preference for the DRILCOM system (*M*=-0.22, *SD*=0.80). Lastly, participants were asked to rate *how difficult it was to judge the location of the sounds* using each system from Extremely Difficult (-3) to Extremely Easy (3). No significant differences in difficulty ratings were found between the two systems, DRILCOM (*M*=0.28, *SD*=1.67) and PALAT (*M*=0.03, *SD*=1.65).

Similar Training Benefits Across Training Sessions

To review, the full-factorial repeated-measures experiment was completely counterbalanced on training system and listening condition in order to reduce order effects. As a result, half of the participants trained first on the DRILCOM system under one listening condition completing the full complement of learning units before switching to the PALAT system while the other half started on PALAT and switched to DRILCOM. Once complete with both systems under the first assigned listening condition, each participant returned to the first system and began training and testing under the second listening condition. On average, training sessions were separated by one day (*M*=23.3 hours) with a minimum of a three hour break and a maximum of 96 hours between sessions. Previously discussed statistical analyses evaluated the dependent measures for all 12 participants after completion of the entire study. However, the investigator wanted to see if there were noticeable differences between absolute score if the participant was assigned to the DRILCOM or PALAT system first, on all three listening conditions.

The investigator evaluated the training benefit imparted by each localization training system over the course of the two training sessions by calculating the mean absolute score at each learning unit on the first session and second session by training system. Figure 141 displays the open ear mean ***absolute score*** by training session in the assigned order. Each line represents half of the participants' mean ***absolute scores*** on the respective system. The participants switched systems after the first session, which is depicted on the graph by the dotted line between LU5 of the first system and LU0 of the second system. The mean absolute score line is coded to show participant performance on the DRILCOM system (blue solid line) and PALAT system (orange dashed line) with the participants switching between systems halfway through the graph. Thus, the DRILCOM mean absolute score blue solid line turns into the PALAT mean absolute score orange dashed line in between the first and second session.

Figure 141 for the open ear clearly supports that the PALAT system is able to impart a similar training benefit as the DRILCOM system during each training session. Participants assigned to both systems started the first pretest with mean absolute scores within 1% of each other with the mean on the PALAT system slightly lower. Participants assigned to the PALAT system demonstrated a slower training rate in LU1 and LU2 but were achieving similar scores as participants assigned to the DRILCOM system by LU3 through LU5. Switching systems, including the one-day break, had very little impact on localization performance on either system under the open ear condition. Participants assigned to both systems in the second training session achieved similar training rates and finished with the virtually the same mean absolute score, a mean of 34.5 correct for participants who started on the DRILCOM system for the first session (*SD*=2.7) and a mean of 34.8 correct for participants who started on the PALAT system for the first session (*SD*=0.8). Of note, there was an increase of 4.6% from LU5 posttest on the first session to LU5 posttest on the second session indicating there was still room for auditory localization improvement given an additional 1.5 hour training session.

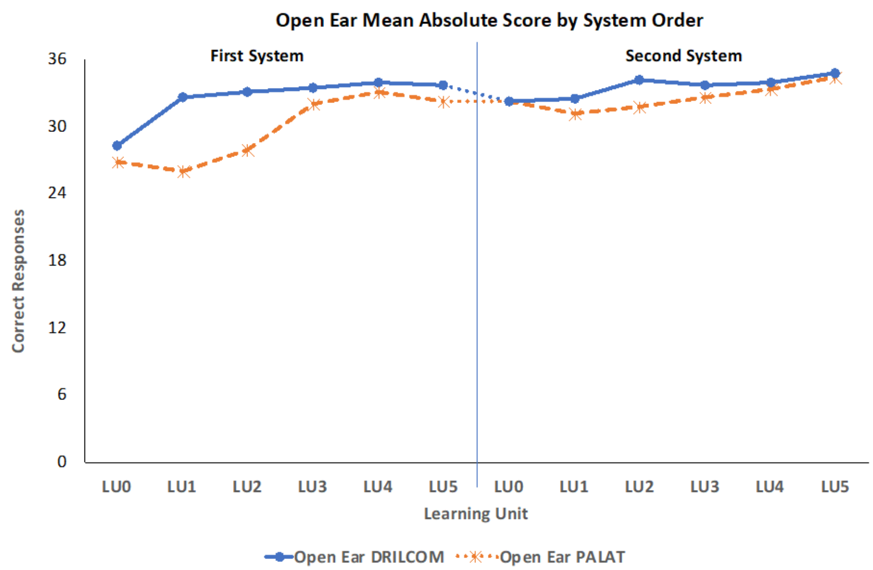


Figure 141. Mean *absolute score* under open ear condition by training system by order of assignment, DRILCOM system (blue solid line) and PALAT system (orange dashed line).

A similar trend was observed under the TEP-100 condition. Figure 142 indicates that participants using the PALAT system were able to achieve similar training benefits as imparted by the DRILCOM system. However, mean absolute score for participants assigned first to the PALAT system scored 9.3% lower during the pretest than those assigned first to the DRILCOM system. The PALAT-first participants achieved a higher training rate through LU4 reducing the difference to 4.6%. Switching systems resulted in a decrease in localization performance on the pretest under the new training system. However, mean absolute score on both PALAT and

DRILCOM shot back up to the highest scores after the first learning unit and continue to increase until LU4. Mean participant scores on both systems declined on the last posttest and ended within 3%, or 1 correct response, a mean of 19.5 correct for participants who started on the DRILCOM system for the first session (*SD*=9.2) and a mean of 20.7 correct for participants who started on the PALAT system for the first session (*SD*=9.3).

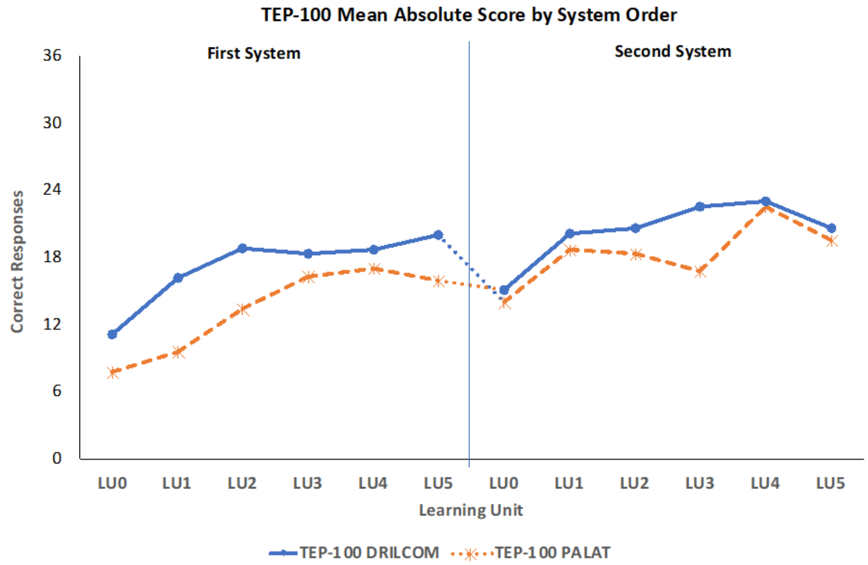


Figure 142. Mean *absolute score* under TEP-100 condition by training system by order of assignment, DRILCOM system (blue solid line) and PALAT system (orange dashed line).

The mean absolute scores under the TEP-100 condition were, by far, the lowest of the three conditions and after a total of 10 learning units, the mean absolute score across both systems was only 56% correct. While this was an improvement from the 9% mean absolute correct at the first pretest, the low accuracy coupled with the highest number of front-back reversal errors is alarming for a U.S. military-fielded TCAPS on which warfighters are actively relying. The poor localization performance under the TEP-100 condition is cause for concern for U.S. military service members as the TEP-100 was selected by the U.S. Army as the program of record for hearing protection devices in 2017, and is already fielded to military units performing combat operations (DePass, 2017). The true implications of reduced auditory situation awareness for military service members is hard to define and even more challenging to simulate in military operational settings. As a result, there is a dearth of prior published research to enable an assessment of the military operational performance impacts of 56% absolute correct localization or even worse, 9% absolute correct localization. To say the least, both these pre-training and after-training levels of performance are low, and especially so when compared to those of the open ear at the same states (Figure 80). One relevant study is that of Brungart and Sheffield (2016) who studied hearing loss and poor localization in simulated combat drills. The results showed that service members with simulated hearing loss were able to detect a degradation in hearing and tended to not take risks. However, service members with simulated poor localization were not able to detect the degradation and made risky decisions that reduced their survivability (Brungart & Sheffield, 2016). To better understand why the TEP-100 resulted in poor localization performance, the investigator conducted ad hoc testing of the TEP-100 device (discussed later herein).

Participants who began training on the DRILCOM system under the ComTac™ III condition recorded a mean absolute score 14% higher on the pretest than participants who began training on the PALAT system. However, participants who started on the PALAT system achieved higher training benefits and ended the first session with a mean absolute score within 0.9% of the DRILCOM first participant mean absolute score (Figure 143). Switching systems resulted in an 11% decline in mean absolute score on both training systems. Participants assigned to both the PALAT and DRILCOM systems for the second session were able to make up for the initial decline and achieve a 5% increase in mean absolute score from LU5 of the first session to LU5 of the second session.

Participants assigned to both systems in the second session achieved similar training rates and finished with the similar mean absolute scores, a mean of 24.8 correct for participants who started on the DRILCOM system for the first session (*SD*=8.1) and a mean of 24.5 correct for participants who started on the PALAT system for the first session (*SD*=8.3). As with the open ear condition, there was a mean absolute score increase of 4.6% from LU5 posttest on the first session to LU5 posttest on the second session indicating there was still room for auditory localization improvement given an additional 1.5-hour training session.

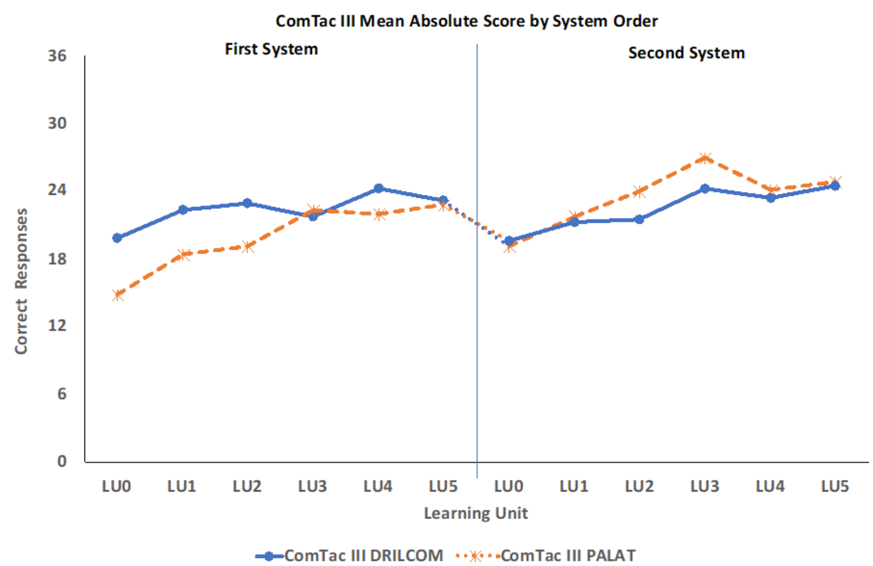


Figure 143. Mean <u>*absolute score*</u> under ComTac^{TM} III condition by training system by order of assignment, DRILCOM system (blue solid line) and PALAT system (orange dashed line).

Participant localization accuracy performance under all three listening conditions supported the hypothesis that the PALAT system imparts similar training benefits as the proven, full-scale laboratory grade DRILCOM system. First session pretest mean absolute scores for all three listening conditions were lower on the PALAT system, possibly indicating that localization using the PALAT system or in the office room environment is a slightly harder task. This may also be the reason for lower mean absolute correct scores in LU1 and LU2 on the PALAT system. If the PALAT system is indeed more difficult than the DRILCOM at the pre-training stage, this could be an advantage for the PALAT in a training sense, because it may be a closer representation of localization tasks in real-world scenarios as demonstrated in Phase III (discussed in the following chapter). A lower pre-training baseline accuracy score also reduces the possibility of a ceiling effect that occurs as a result of limited room for improvement due to a near perfect baseline score; the open ear condition is the most likely listening condition to be impacted by ceiling effects. However, after three learning units, mean absolute scores on the PALAT system were in line with mean absolute scores on the DRILCOM system. Analyzing the mean absolute scores by training session order also confirmed the ability of the PALAT system to support the secondary hypothesis of being able to detect differences in localization performance between TCAPS devices, which is important not only for ascertaining the training burden imposed by a product, but also whether that product is contraindicated for fielding.

Sensitivity to Device Differences in Auditory Localization Training

A secondary objective of the PALAT system was to provide the capability of detecting device differences in auditory localization training. A series of auditory localization testing conducted at VT-ASL demonstrated performance differences between the open ear and HPDs (Casali & Robinette, 2014; Casali & Lee, 2016a; Casali & Lee, 2016b). Abel (2008) and Bevis et al. (2014) highlighted the necessity of being able to measure performance differences between hearing protection devices by reporting warfighters' reported impacts of reduced situation awareness while using military HPDs. Both studies detailed focus group findings from military service members who described making conscious decisions to forgo hearing protection due to the deleterious effects on localization (Abel, 2008; Bevis et al., 2014). The studies conducted at VT-ASL demonstrated, via objective measurements, the adverse effects TCAPS caused on localization performance, finding significant differences in accuracy, front-back reversal errors, and response times between the open ear and TCAPS conditions (Casali & Robinette, 2014; Casali & Lee, 2016a; Casali & Lee, 2016b).

The Phase II study contained multiple unique factors that made it difficult to compare directly with previous studies. This was the first study to use the PALAT system or to test participant localization on the DRILCOM system in conjunction with another training system. The training protocol used was developed in Phase I of the overarching study by K. Cave and implemented for the first time during Phase II (Cave et al., 2019). This was also the first known study to test the TEP-100 on localization performance. Lastly, the full-factorial repeated-measures experiment was completely counterbalanced meaning two-thirds of the results on any listening condition were from participants who had already completed two training sessions under a different listening condition (a session on DRILCOM and PALAT) and one-third of the results on any listening condition were from participants had completed four training sessions. In addition, participants were trained on both the PALAT and DRILCOM system under each listening condition prior to switching to a new listening condition. As a result, half of the results for any listening condition were from participants completing their second session of training under the same listening condition. Figures 141, 142, and 143 demonstrated that mean absolute scores continued to improve during the second session causing the mean absolute score for each system to be slightly higher than other studies after five equivalent training sessions. However, results can still be compared to previous studies to confirm the ability of the PALAT system to effectively detect differences between listening conditions.

Consistent with previous localization studies, the PALAT system clearly demonstrated the ability to distinguish localization performance between listening conditions in the same manner as the DRILCOM system if not better. On both the DRILCOM and PALAT systems, participants under open ear significantly outperformed both TCAPS conditions on absolute score and front-back reversal errors at LU0 and LU5. The PALAT system also detected a significant difference in mean absolute score between the ComTac™ III condition and TEP-100 condition at LU0, though the DRILCOM system did not exhibit similar sensitivity. As with previous localization studies, the open ear condition resulted in the highest localization performance across all stages of training on both PALAT and DRILCOM. Surprisingly, participants under the in-the-ear TCAPS device, TEP-100, performed worse than while using the over-the-ear TCAPS device, ComTac™ III on both PALAT and DRILCOM. In addition, both the DRILCOM and PALAT systems detected significant differences in front-back reversal errors between the open ear condition and both TCAPS devices. Participants recorded the highest number of front-back reversal errors under the TEP-100 condition followed by the ComTac™ III and the least amount, by a significant difference, under the open ear condition. Again, the poor performance under the TEP-100 condition was counter to previous studies with in-the-ear TCAPS or HPDs compared to over-the-ear TCAPS or HPDs, thus it is believed that this device exhibited unique dynamic characteristics which led to its relatively poor localization accuracy, regardless of the training system used. Thus, the implications of this discovery were further explored in the next section.

Previous localization studies found that various in-the-ear TCAPS retained localization cues and as a result typically outperformed over-the-ear TCAPS devices. Abel et al. (2007) found that earmuff style HPDs were particularly disruptive to sound localization because they prohibited pinna effects. Talcott et al. (2012) found the PELTOR™ ComTac™ II earmuff ranked lowest in localization behind an in-the-ear HPD and that the open ear outperformed all HPDs in both absolute and ballpark localization accuracy. Casali and Lee (2016a) used the DRILCOM system with the original training protocol to test localization accuracy of the

ComTac™ III against open ear and two in-the-ear TCAPS devices on absolute correct score. The open ear significantly outperformed all TCAPS and the in-the-ear TCAPS devices performed higher than the ComTac™ III. While the Phase II study supports the hypothesis that the PALAT system provides the capability to detect differences between listening conditions, the poor performance of the in-the-ear TEP-100 TCAPS warranted further investigation.

Investigation of Localization Performance under TEP-100

As alluded to above, the poor localization performance, absolute score, front-back reversal errors, and response times, under TEP-100 condition in the Phase II study was concerning and inconsistent with previous studies using in-the-ear HPDs or TCAPS. While not directly comparable, Figure 144 displays the results from a previous DRILCOM study along with the TEP-100 mean absolute scores from the Phase II study. Casali and Lee (2020) tested localization performance of the Etymotic ER125-GSE in-the-ear HPD using the DRILCOM system with the original training protocol consisting of five replications of sequential and five random presentations per learning unit prior to each test. The Etymotic ER125-GSE, a device that has similar form to the TEP-100 but very different dynamic signal pass-through and gain compression characteristics, achieved considerably higher localization absolute scores and improved by 29% from LU0 to LU5 (Casali & Lee, 2020). To achieve a closer comparison, the Phase II mean absolute score shown for DRILCOM and PALAT were calculated for half the study population who were assigned to the DRILCOM system for the first training session (blue solid line) and PALAT system for the first training session (orange dotted line). Figure 142 shows how mean absolute scores during the second training session continued to improve causing the total mean absolute score by system to increase slightly. Figure 144 shows the mean absolute scores for participants assigned first to either DRILCOM or PALAT systems below the

Etymotic ER125-GSE mean absolute score at all learning units. In addition, the TEP-100 resulted in a lower percentage of training improvement from LU0 to LU5, DRILCOM improved by 24% and PALAT improved by 23%, compared to a 29% improvement under the Etymotic ER125-GSE.

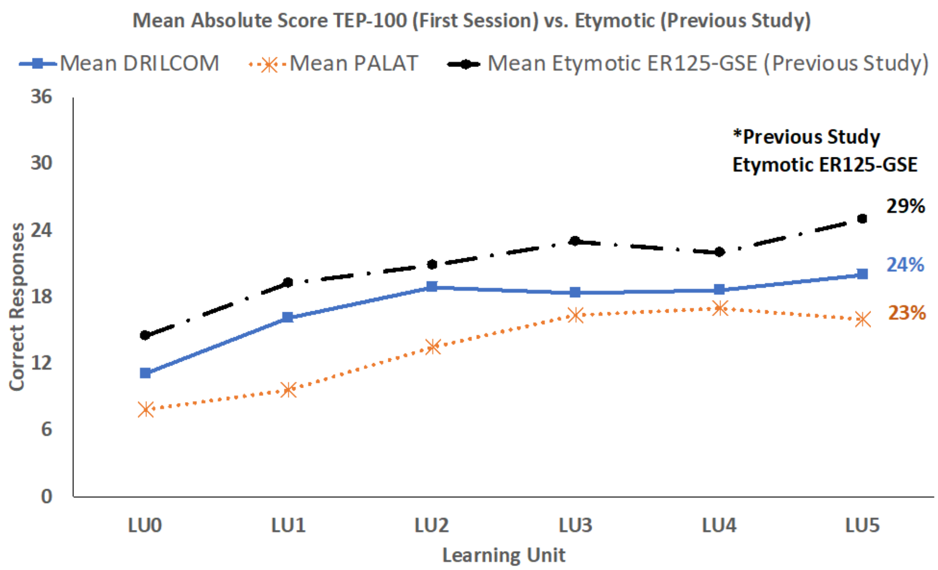


Figure 144. Phase II mean *absolute score* for six participants assigned to DRILCOM (blue solid line) or PALAT (orange dotted line) for the first training session versus mean absolute score from 11 participants tested using Etymotic ER125-GSE (black dash-dotted line) on DRILCOM in a previous study (Casali & Lee, 2020). Percent improvement from LU0 to LU5 of correct responses displayed on right side of chart.

In order to better understand the poor localization performance while using the TEP-100, the investigator tested the TEP-100 to measure the output sound pressure level and spectral content when presented with the dissonant tonal signal at two sound pressure levels, 55 dBA and 80 dBA. The investigator specifically wanted to test the effects of the compression technology employed by 3M™ in the TEP-100 software on output sound pressure level and the spectral content of the auditory signal. The TEP-100 uses compression technology to prevent hazardous steady state noise and impulse noise from damaging the ear (3M, 2016b). The compression technology is intended to reduce high sound pressure level noise to a safe decibel level while still allowing the auditory signal to be heard. The investigator believed that the compression software algorithm, the onset decibel level of the compression curve, the shape of the compression curve, and/or the fidelity of processing and passing-through localization cues were potential sources for the lower localization performance under the TEP-100 listening condition.

The investigator measured the resulting output sound pressure level of the TEP-100 using a KEMAR manikin when presented with the dissonant tonal signal at both 55 dBA and 80 dBA from the 12 loudspeakers in the DRILCOM system hemi-anechoic room. Figure 145 displays the sound pressure levels of the dissonant signal at both 55 dBA (left column graphs) and 80 dBA (right column graphs) at 800 Hz, 5000 Hz, and 8000 Hz 1/3 octave-band frequencies as measured using a Larson Davis® measurement microphone located inside the right ear canal of the KEMAR manikin under the open ear (blue dotted line) and TEP-100 (orange solid line) listening conditions. The dissonant signal was sounded and recorded three times from each loudspeaker location. The radial plots display the mean sound pressure level recorded at a specified frequency from each loudspeaker location. Based on the proximity to the loudspeaker and head shadow effect, the radial plot should display slightly higher sound pressure levels when the signal originates from a loudspeaker closer to the ear being measured, e.g. the right ear plot should display a higher sound pressure level from the 3 o’clock loudspeaker position than the 9 o’ clock loudspeaker position.

The resulting impacts of the compression technology are illustrated by comparing the radial plots at each 1/3 octave-band frequency between the 55 dBA (left column) and 80 dBA (right column) signals at each loudspeaker location. At 800, 5000 and 8000 Hz, the *difference* between the sound pressure level recorded with the open ear and the TEP-100 is higher from loudspeakers on the right side, or near ear, when presented with the 80 dBA signal than with the

55 dBA dissonant signal. The radial lines for open ear and TEP-100 are overlapping or much closer when presented with a 55 dBA signal demonstrating the TEP-100 was not compressing the signal. At 80 dBA, the convergence of the two sound pressure level radial lines displays the effect of the TEP-100 compression software. This attenuation of the signal in the near ear under the TEP-100 indicates the dissonant tonal signal set at 70 dBA during the Phase II study was impacted (i.e., reduced in level) by the TEP-100 compression circuit during the localization training and testing.

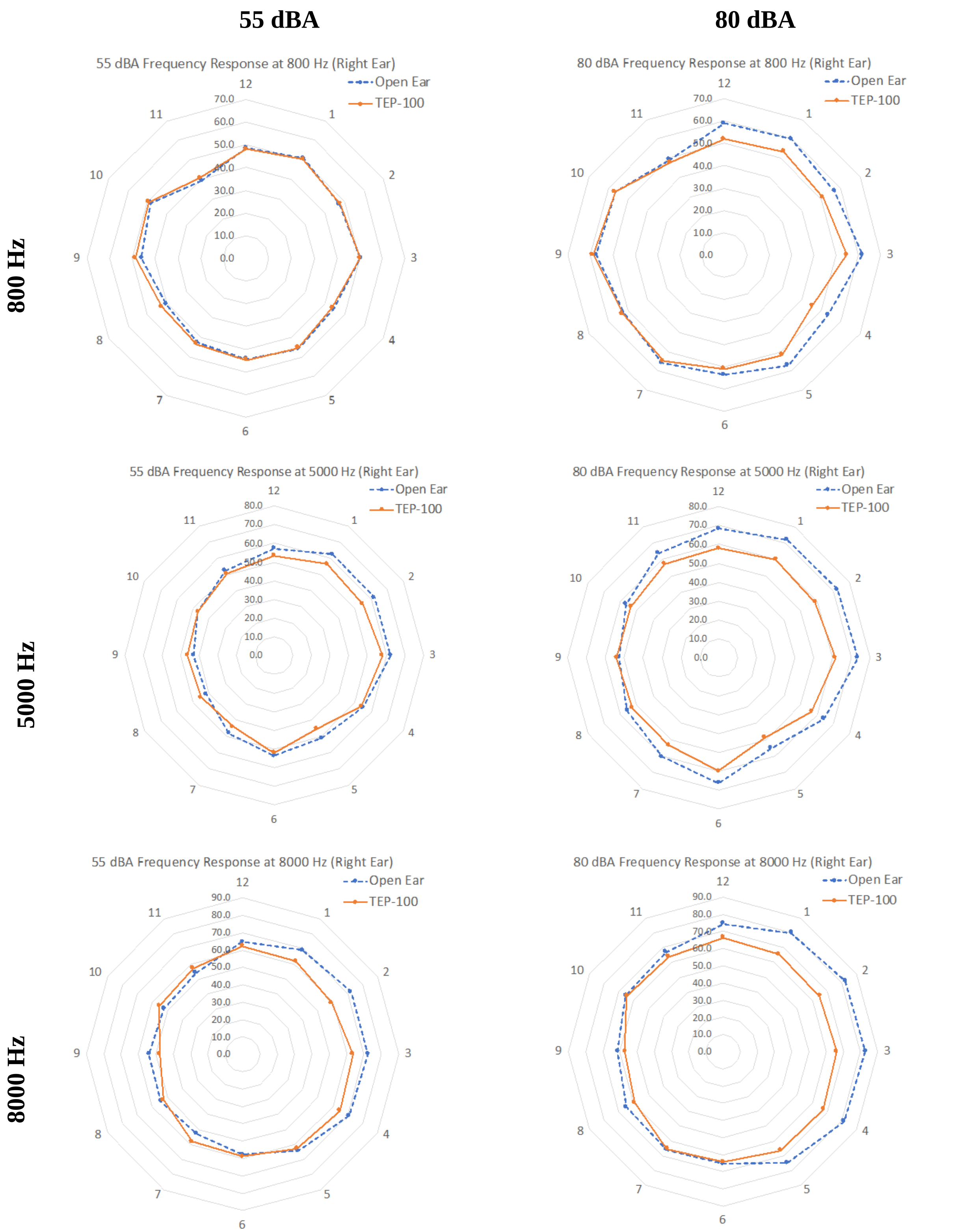


Figure 145. Frequency response radial plots of mean sound pressure level at a single 1/3 octave-band frequency. Measurements recorded using KEMAR manikin with open ear (blue dotted lines) and TEP-100 (orange solid lines) in right ear. Dissonant tone set at 55 dBA (left column) and 80 dBA (right column).

The investigator shared the unexpectedly poor localization results and attenuation measurements of the TEP-100 with representatives from 3M™ during the National Hearing Conservation Association Annual Conference 2019. The investigator acquired a frequency response curve chart depicting a representative TEP-100 compression curve from 3M™ in order to analyze the effects of compression onset and ramp-up on localization performance. Figure 146 confirms the TEP-100 begins compressing the auditory signal at 60 dBA as indicated by the bend or knee-point in the minimum volume line (lower red line curve) using a broadband pink noise signal (Stergar, Fackler, & Hamer, 2019). At 70 dBA, the sound pressure level used during the Phase II experiment, the gain circuit using the minimum volume or unity gain setting on the TEP-100 compresses the signal to an output level below the input level. The data shown in Figure 146 are representative results, but not considered exact for all samples, per manufacturer.

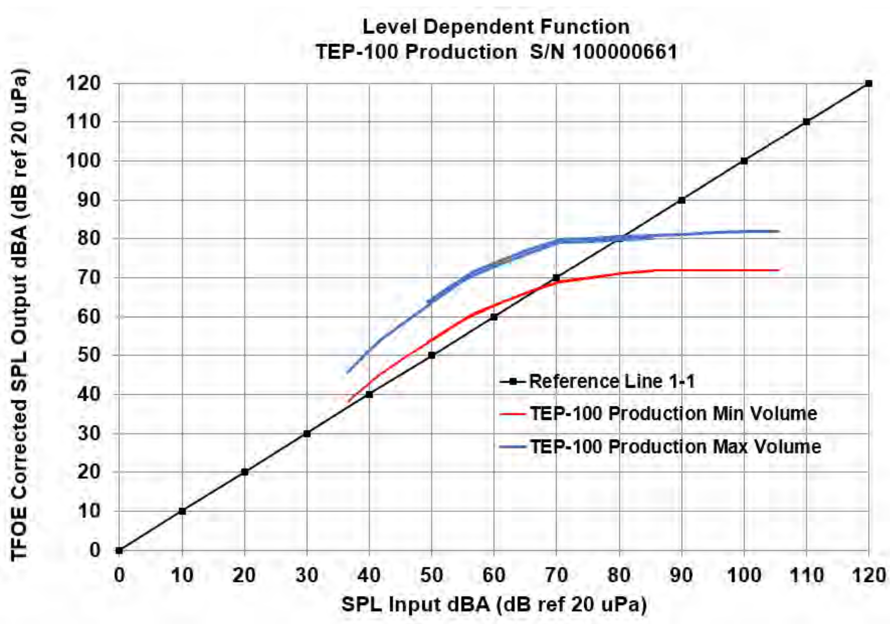


Figure 146. Level dependent function of the TEP-100 measured using a broadband pink noise signal; data are representative, but not considered exact for all samples, per manufacturer (Stergar et al., 2019).

The sound level reduction in a single ear may have contributed to the lower localization performance. In addition to the signal compression, discussions with the 3M^{TM} representatives also revealed that each TEP-100 earpiece acts independently with no communication with the earpiece in the opposite ear; this results in independent degrees of compression between the two earpieces for the arriving signal. As a result of the independent or unsynchronized compression functions, the head shadow effect which create the interaural level difference localization cues could result in the earpiece on one side of the participant's head (the near-signal side) to compress the signal while the opposite earpiece (the far side) fails to reach the compression threshold. This scenario was highly probable for the Phase II study given that the 70 dBA dissonant tonal signal aligns directly with the inflection point of the compression curve for the TEP-100 processing algorithm (Figure 146). Altering or eliminating interaural level differences that are normally relied upon by the open ear would significantly degrade localization performance with such an in-the-ear TCAPS. The results of the independent compression functions are evident in the radial plots (Figure 145). When the 80 dBA signal originated from the same side as the device (near ear), the compression circuit within the TEP-100 earpiece reduced the signal at all 1/3 octave-band frequencies, indicated by the gap between the open ear SPL (blue dotted line) and TEP-100 SPL (orange solid line) from loudspeakers located at 12 o'clock to 5 o'clock. Whereas, when the 80 dBA signal originated from the opposite side of the device (far ear), the compression circuit within the TEP-100 earpiece maintained unity gain amplification (and no compression) presenting a similar SPL as measured with the open ear listening condition, indicated by the overlapping lines of the open ear SPL, (blue dotted line) and TEP-100 SPL (orange solid line) from loudspeakers located at 6 o'clock to 11 o'clock. More testing is needed to confirm the impacts of the independent compression circuitry with each

TEP-100 earpiece (of a pair) on localization performance. However, preliminary results indicate the poor localization performance under the TEP-100 listening condition may in part be attributable to the independent or unsynchronized compression technology software algorithm, the onset and ramp-up of the compression in the pass-through gain circuit, and/or the fidelity of processing and passing-through localization cues.

Training System Impacts on Response Time

The PALAT system was designed based on lessons learned from previous DRILCOM studies to improve upon the user interface and allow for individual users to operate the system without an experimenter present. Specifically, a tablet computer with stylus pen was chosen as the user interface hardware in order to reduce the DRILCOM equipment consisting of a desktop computer, two monitors, keyboard, experimenter mouse, and user operated mouse. Questionnaire results confirmed the PALAT system user interface was rated as being easier to operate and perceived to provide a more positive impact on the ability to train localization. Participant ratings for ease of use to operate the system were significantly higher for the PALAT system under all three listening conditions. The PALAT system was also rated significantly higher to provide a more positive impact on the ability to train localization under both TCAPS conditions. One predicted impact of the user interface improvements was the lower response time scores when using the PALAT system.

Mean response times on the PALAT system were significantly faster than the times on the DRILCOM system for all listening conditions. The quicker response times are attributed more to the design principles of response selection, specifically location compatibility, since the direct position control of the PALAT system allows the user to physically touch the response location button with their finger or stylus as opposed to using the indirect position control computer mouse to provide feedback (Wickens, Lee, Liu, & Becker, 2004). However, when analyzing mean response times on the PALAT and DRILCOM separately, similar trends were observed between the two systems. Response times under the open ear condition were significantly lower than response times for both TCAPS conditions on both the PALAT and DRILCOM systems. There was no significant difference detected on mean response times between the TCAPS devices but mean response times were highest under the TEP-100 condition on both DRILCOM and PALAT system. Overall, the PALAT system was able to measure similar trends as the DRILCOM system in regards to response time and was sensitive to response time differences between listening conditions in the same manner as the DRILCOM system.

Phase II Overall Conclusions

The PALAT system demonstrated the ability to improve auditory localization skills over a regimented series of learning units in a similar manner as the full-scale DRILCOM system. No significant differences were found between the DRILCOM or PALAT system on absolute score or front-back reversal errors at all stages of training or training effects measured by absolute score regression line slopes through the entire complement of learning units. Results also supported the PALAT system ability to detect auditory localization performance differences between the open ear and an in-the-ear and an over-the-ear TCAPS. The PALAT system was able measure poor localization performance under a currently fielded TCAPS device potentially caused by design issues. Improvements made to the PALAT system user interface were perceived by participants to have a positive impact on the ability to train localization and were rated as making the PALAT system easier to operate than the DRILCOM system.

There are currently no studies that quantify a minimum or desired localization performance standard needed to successfully complete military operations. As such, it is difficult to determine if the training protocol developed and tested in Phase I combined with the PALAT system developed and tested in Phase II are capable of providing the requisite localization training needed in the U.S. military. A first step toward defining a localization standard is to test the external validity of the PALAT system and training protocol. Phase III of the overarching study addressed the external validity by training participants under the same conditions employed during Phase II followed by testing in a field environment using real gunshot stimuli signals.

3.8 Implications for the Phase III Experiment

The positive findings in Phase II warranted the necessity to conduct the Phase III experiment to evaluate the transfer of training of the PALAT system training effects to a real world, outdoor test environment using a military relevant gunshot signal. The Phase I main experiment found no significant training effect after the fourth learning unit (Cave et al., 2019). The recommendation was made to reduce the training protocol to a pretest, LU0, and four learning units to reduce the training time with the expectation that future training availability of U.S. military service members is limited. However, participants in Phase I were only tested under the open ear condition. The investigator chose to use five learning units during Phase II to provide a balance of an additional training session but limit training time to an hour and a half for a complete training session, LU0 through LU5. Analysis of mean absolute score performance over the course of two training sessions showed there is still room improvements in localization accuracy after LU5. However, the investigator recommended continuing to use only five learning units under each listening condition since each participant would complete training under all three listening conditions.

The Phase II investigation identified potential design issues within the TEP-100 device that resulted in lower localization performance than the other two listening conditions. However, participants still demonstrated the ability to improve localization performance over the course of the training protocol. As a result, the investigator recommended training and testing the same three listening conditions during the Phase III experiment to test the external validity of the training effects under each listening condition.

Finally, Phase II experimental instructions clearly specified that the DRILCOM and PALAT systems would only use 12 loudspeakers due to major design differences in the quantity and visibility of the loudspeakers. The specific instructions possibly resulted in the majority of participants only responding with locations associated with the 12 active loudspeakers eliminating the ability or need to analyze ballpark accuracy, or within ±15° of signal location. For Phase III, the investigator stressed that participants were able to respond using all 24 loudspeaker locations or 24 gunshot locations to facilitate the evaluation of ballpark localization accuracy.

CHAPTER 4. Phase III In-Field Investigation of Transfer-of Training

4.1 Phase III Objectives

The primary objective of the Phase III in-field experiment was to evaluate the transfer-of-training effects of conducting azimuthal localization training in-lab, using the PALAT system, on in-field localization performance. The experiment also investigated the sensitivity of the in-lab training and in-field testing to differences among listening conditions. As such, the evaluation incorporated three listening conditions: open ear (unoccluded), in-the-ear TCAPS device (TEP-100), and over-the-ear TCAPS device (ComTac™ III). A secondary objective was to evaluate the validity of using the in-lab PALAT system results to predict localization performance under all three listening conditions in a military operational setting. To meet these objectives, localization performance was compared between trained and untrained participants using an in-lab pretest and an in-field posttest using live (blank) gunshots.

4.2 Phase III Methodology

This investigation aimed to measure the transfer-of-training effect instilled by the PALAT system using the localization training protocol developed in Phase II and Phase I, respectively. A series of auditory localization studies conducted at the Virginia Tech Auditory Systems Laboratory previously measured localization ability in terms of accuracy, response time, and subjective rankings of participant-perceived localization ability in a single environment, either the laboratory (Casali & Robinette, 2014; Casali & Lee, 2016a; Casali & Lee, 2016b) or a field environment (Talcott et al., 2012). Unique to the Phase III experiment, participants trained and tested on a localization task in an office setting using a dissonant tonal signal, but tested in a field using a military-relevant signal. Previous VT-ASL experiments did not incorporate lab and field and environments within the same study nor did participants train on signals different from the test stimuli.

The participant first signed a consent form and then was audiometrically and demographically screened. Next, the participant underwent pretesting in the PALAT system under all three listening conditions (i.e., open ear, TEP-100, and ComTac™ III) using the dissonant training signal. Figure 147 displays a participant's progression through the experiment. The PALAT system was located in an academic building office space at Virginia Tech. The pretest order was counterbalanced and the participant's order was maintained throughout the stages of the study. After the pretest, half of the participants were randomly selected to undergo localization training on the PALAT system. In the trained group, or experimental group, participants underwent the localization training protocol developed in Phase I of the overarching research effort (Cave et al., 2019; Cave, 2019). The training sessions consisted of three, one-hour sessions of localization training using the PALAT system. Training occurred under one listening condition during each one-hour session. Each session included all five learning units of localization training and testing. The three training sessions were completed within three days of the pretest date with a maximum of two training sessions per day and at least two hours separation between any training session. Within one to three days after training completion under all listening conditions, the trained and untrained groups underwent field testing. The field site for the posttest experiment was designed to simulate a scenario where a U.S. Military service member listens for enemy threats in a lightly-wooded field. The same field site previously used in Casali et al. (2012) was used in this experiment. Results from the Subject Matter Expert survey conducted by Cave et al. (2019) showed the most prominent enemy threat facing U.S.

Armed Forces ground combat personnel was gunshots (Cave, 2019). As a result, gunshots from Fiocchi .22 caliber long rifle blanks were used as the posttest stimuli.

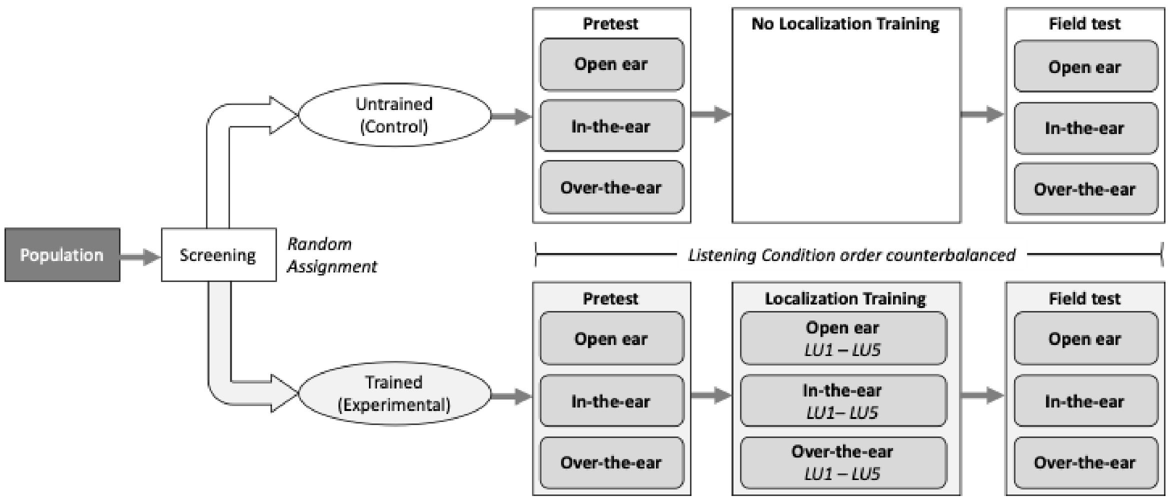
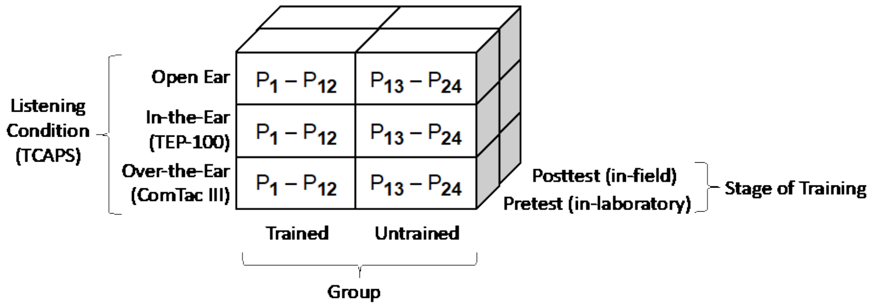


Figure 147. Phase III experimental design order.

4.3 Phase III Experimental Design

Phase III consisted of a pretest-posttest, control group design experiment involving 24 normal-hearing participants (Figure 148) who had neither prior experience in localization testing nor TCAPS use. The 2 x 3 x 2 mixed factor design involved three independent variables: a between-subjects factor of group with two levels (trained and untrained) and two within-subjects factors: listening condition with three levels (open ear, TEP-100, and ComTac™ III) and stage of training with two levels (pretest and posttest). Results were measured using three groups of dependent measures: localization accuracy, response time, and participant subjective responses.



<u>Experimental Design</u>

1. Mixed-factor, pretest posttest control group design
   Between-subjects
   - 2x Group: Trained (Experimental) and Untrained (Control)
   Within-subjects
   - 3x Listening conditions: Open ear, In-the-ear, Over-the-ear
   - 2x Stage of training: Pretest (In-laboratory) and Posttest (In-field)
2. Order of:
   - Group assignment – Random
   - Listening condition – Latin square, counterbalanced order
   - Stimulus presentation azimuth – Random during testing
3. Participants
   - 18-45 years of age
   - 75-85% male to generalize military population
   - Normal Hearing
   - No experience with PALAT system, localization training, or TCAPS

<u>Dependent Variables</u>

1. Localization accuracy
   - Percent absolute accurate (response matches signal azimuth)
   - Percent ballpark accurate (response ±15° signal azimuth)
   - Front-back confusion percent
2. Response time (seconds, with msec precision)
3. Perceived localization performance rating
   - Confidence in accuracy for each listening condition
   - Perceived effect of training on localization accuracy
   - Perceived effect on response time for each listening condition

Figure 148. Experimental design for Phase III, with independent variables, experimental order, participant assignment, and dependent measures listed.

A Microsoft® Excel random number generator was used to assign 24 participant numbers to an arrival order. Participants who were assigned numbers 1 to 12 were assigned to the trained group, and participants 13 to 24 were assigned to the untrained group. Two participants were replaced during the experiment, for reasons explained later. The replacement participants were assigned the participant number of the participant who they replaced. In order to generalize to the U.S. Military population of 84% male and 16% female, participants were limited to 18 males and six females with nine males and three females randomly assigned to both the trained and untrained group (U.S. Department of Defense, 2018). Four sets of an identical 3 x 6 Latin square were repeated to counterbalance the listening condition order for each participant. The participant listening condition order was maintained throughout the study. Table 101 displays the participant order for the Phase III experiment by sex, group assignment, and listening condition order.

Table 101. Participant study order by sex, group assignment (random assignment based on arrival order), and listening condition (counterbalanced using a repeating 3 x 6 Latin square).

			Listening Condition order		
Arrival order	Participant Number	Group Assignment	1	2	3
M1	P20	Untrained	ITE	OTE	Open
M2	P13	Untrained	Open	ITE	OTE
M3	P21	Untrained	OTE	Open	ITE
M4	P2	Trained	ITE	OTE	Open
M5	P4	Trained	OTE	ITE	Open
M6	P18	Untrained	ITE	Open	OTE
M7	P12	Trained	ITE	Open	OTE
M8	P7	Trained	Open	ITE	OTE
M9	P19	Untrained	Open	ITE	OTE
M10	P15	Untrained	OTE	Open	ITE
M11	P1	Trained	Open	ITE	OTE
M12	P3	Trained	OTE	Open	ITE
M13	P8	Trained	ITE	OTE	Open
M14	P10	Trained	OTE	ITE	Open
M15	P24	Untrained	ITE	Open	OTE
M16	P16	Untrained	OTE	ITE	Open
M17	P23	Untrained	Open	OTE	ITE
M18	P11	Trained	Open	OTE	ITE
F1	P6	Trained	ITE	Open	OTE
F2	P14	Untrained	ITE	OTE	Open
F3	P22	Untrained	OTE	ITE	Open
F4	P17	Untrained	Open	OTE	ITE
F5	P5	Trained	Open	OTE	ITE
F6	P9	Trained	OTE	Open	ITE

4.3.1 Independent Variables (IVs)

Independent Variable – Group

Per Figure 148, two between-subjects group levels were used in this investigation: trained (experimental) and untrained (control). Participant age range, mean and median were very similar between the two levels (Table 102). Participants were informed which group they were assigned to prior to the pretest in accordance with the VT IRB requirements.

Table 102. Participant demographics by group.

	Trained (n=12)	Untrained (n=12)
Age (years)		
Range	19 - 30	19 - 34
Median	26	27
Mean	25.6	26.9
SD	2.9	4.6

Open ear

The open ear listening condition was included in this investigation for several reasons. First, testing the open ear condition established a baseline performance, enabling a within-subjects comparison of training effect for each TCAPS device. Secondly, the open ear condition is the most commonly-encountered listening condition for U.S. Service Members in training and combat environments where hazardous noise exposure is not imminent or expected, but threat or hazard localization remains paramount. Lastly, several studies have identified barriers to HPDs and TCAPS compliance. Abel (2008) and Bevis et al. (2014) specifically described discomfort and a perceived loss of auditory situation awareness as reasons for non-compliance by U.S. Service Members. In Bevis et al. (2014), all 16 focus groups mentioned that auditory localization was negatively affected by hearing protection devices. One British Army Soldier stated, “If you can’t locate that position then you’re redundant” (Bevis et al., 2019, p131). Therefore, by examining localization performance in the open ear, the influence of device-imposed changes to environmental cues and comfort could be eliminated. Furthermore, the open ear condition addressed the secondary objective of assessing the validity of using PALAT system-obtained results to generalize to auditory localization in the field.

In-the-ear TCAPS

The earplug-style 3M™ PELTOR™ TEP-100 Tactical Earplug is an active, or powered electronic sound transmission, in-the-ear hearing protection device, shown in Figure 149. The TEP-100 Tactical Earplugs are issued as a set of two identical, rechargeable electronic earplugs with a recharging case. For testing purposes, the investigator designated a right and left ear device in each set according to serial numbers. The right and left device designations were maintained throughout the study to reduce confounding effects of differences between earplugs. The 3M™ PELTOR™ level-dependent technology is advertised to "provide hearing protection, and helps improve situational awareness and communication" (3M, 2016a, p1). As a passive earplug, the TEP-100 is advertised to provide a mean attenuation of 23 NRR according to the EPA-required labeling on the device (3M, 2016a). The TEP-100 is compatible with several styles of eartips including the 3M™ PELTOR™ Ultrafit eartips shown in Figure 149 which are the standard issue version for the U.S. Military. As a result, each participant in this experiment were fitted with one of the three sizes of Ultrafit eartips with the TEP-100. A professional U.S. military audiologist, K. Cave, conducted a visual inspection of each participant's ear canal and ensured the participant was fitted with the proper Ultrafit eartip size.

Figure 149. 3M™ PELTOR™ TEP-100 electronic earplug-style TCAPS device.

The TEP-100 tactical earplug is equipped with two volume settings, “normal” and “high,” that is operated by a single button. The investigator tested the TEP-100 volume settings to identify the unity gain setting. Unity gain was previously defined by Casali and Lee (2016a) as the state where the electronic gain control is set to overcome or offset the passive attenuation of the earplug and provide as close to natural hearing as possible. Four TEP-100 devices loaned to the Virginia Tech Auditory Systems Laboratory, two devices from U.S. Army PEO Soldier and two devices from 3M™, were tested in a reverberation chamber to identify the unity gain setting during Phase II of the overarching experiment. One of the TEP-100 devices was found to have significant differences in sound pressure level measurements and was not used during this investigation. The remaining three TEP-100 devices were evenly assigned between the trained and untrained groups so that participants from each group used all three devices.

The following steps were performed to identify the unity gain setting for the TEP-100. A ½ inch Larson-Davis 2575 measurement microphone (SN: 2559) and Larson-Davis 9000C Preamp (SN: 0521) were placed in the center of the reverberation chamber and connected to a Larson-Davis 2900 Model Spectrum Analyzer (SN: A0280) at an investigator table located outside of the chamber. The microphone was calibrated at 94.0 dBA (1000 Hz tone) using a Quest QC-20 Calibrator (SN: QOA070051). A pink noise signal was generated via a MATLAB® program and measured at 70 dBA, 10 second Leq, fast time constant. Next, an acoustical test manikin, known as KEMAR (Knowles Electronics Manikin for Acoustic Research by GRAS), was positioned in the center of the reverberation chamber and the measurement microphone was fitted inside the left ear canal of the KEMAR and the right ear canal was occluded with putty. The pink noise signal was measured in the open ear listening condition at 77.6 dBA which served as the reference level for unity gain. Each TEP-100 earplug was then fitted in the left ear of the KEMAR and the sound pressure level of the pink noise signal was measured three times at each volume setting: off, or passive, setting, normal volume, and high volume.

The “normal” volume setting provided the closest unity gain for the TEP-100 and thus was the setting used for this experiment (Table 103). Figure 150 displays the sound pressure level measurements of the 70 dBA pink noise signal at each 1/3 octave-band frequency for the open ear and TEP-100 at normal volume setting on the KEMAR manikin. The sound pressure levels measured under the TEP-100 are noticeably lower from 100 Hz to 315 Hz than the open ear levels. The TEP-100 also did not transmit the pink noise at the 10000 Hz 1/3 octave-band frequency.

Table 103. Sound pressure level (SPL) of 70 dBA pink noise measured using KEMAR manikin for three sets of TEP-100 devices and mean (by left and right ear designation). Levels compared with the open ear 77.6 dBA reference level displayed as the delta (Δ), open ear – TEP-100.

		Device 1		Device 2		Device 3		Mean	
Listening Condition	**Gain level**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**	**SPL (dBA)**	**Δ**
Open ear (Reference Level)		77.6		77.6		77.6		77.6	
		Left (SN: 64174)		Left (SN: 61389)		Left (SN: 36576)		Left ear	
TEP-100	Off (passive)	31.3	(-46.3)	31.8	(-45.8)	29.1	(-48.5)	30.7	(-46.9)
	Normal	**77.4**	**(-0.2)**	**76.9**	**(-0.7)**	**77.2**	**(-0.4)**	**77.2**	**(-0.4)**
	High	88.1	(10.5)	87.4	(9.8)	88.0	(10.4)	87.8	(10.2)
		Right (SN: 64517)		Right (SN: 64343)		Right (SN: 36173)		Right ear	
	Off (passive)	31.6	(-46)	32.0	(-45.6)	29.8	(-47.8)	31.1	(-46.5)
	Normal	**79.7**	**(2.1)**	**78.2**	**(0.6)**	**75.2**	**(-2.4)**	**77.7**	**(0.1)**
	High	90.3	(12.7)	89.1	(11.5)	86.0	(8.4)	88.5	(10.9)

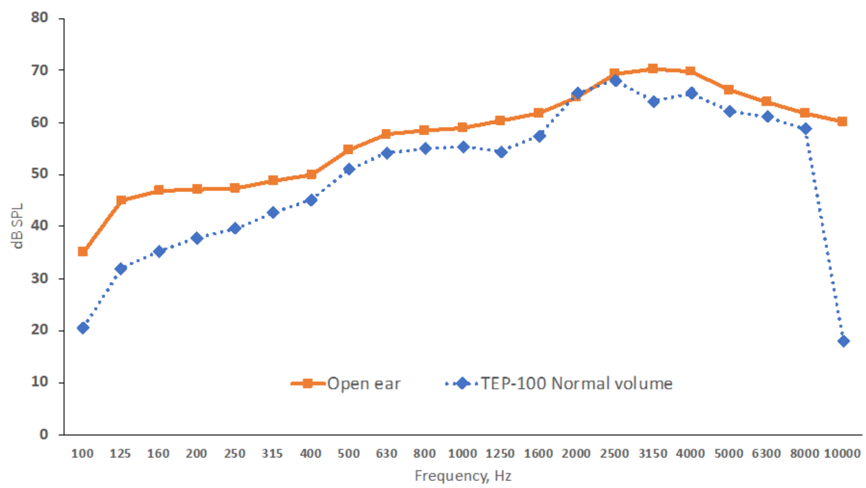


Figure 150. Mean frequency response of 70 dBA pink noise measured using KEMAR under open ear and TEP-100 devices on normal volume setting by 1/3 octave-band frequencies.

Over-the-ear TCAPS

The earmuff-style 3M™ PELTOR™ ComTac™ III headset is an active, or electronic sound transmission, over-the-ear hearing protection device, shown in Figure 151. This battery-powered TCAPS is equipped with four volume settings and an additional boost mode to amplify low level external sounds to audible, but not hazardous levels, and pass them through the muff. According to the manufacturer's literature, the 3M™ PELTOR™ ComTac™ III utilizes a proprietary digital audio circuit to compress hazardous noise to a permissible safe exposure level of less than 82 dBA (3M, 2016b). As a passive headset, the 3M™ PELTOR™ ComTac™ III is advertised to provide a NRR of 23 (3M, 2016b).

Figure 151. 3M™ PELTOR™ ComTac™ III electronic earmuff-style TCAPS device.

Three ComTac™ III headsets were loaned to the Virginia Tech Auditory Systems Laboratory, one headset from U.S. Army PEO Soldier and two headsets from 3M™, for the study. All three headsets were tested to identify the unity gain setting using the same procedure described above. The highest volume setting, or fourth increase from default, provides the closest unity gain for the ComTac™ III and was used for this experiment (Table 104). Figure 152 displays the sound pressure level measurements of the 70 dBA pink noise signal at each 1/3 octave-band frequency for the open ear and ComTac™ III at the high volume setting on the KEMAR manikin. The sound pressure levels measured under the TEP-100 are noticeably lower from 4000 Hz to 10000 Hz than the open ear levels.

Table 104. Sound pressure level (SPL) of 70 dBA pink noise measured using KEMAR manikin for three sets of ComTac™ III devices and mean. Levels compared with the open ear 77.6 dBA reference level displayed as the delta (Δ), open ear – ComTac™ III.

Listening Condition	Gain level	Device 1 (SN: 7500) SPL (dBA)	Δ	Device 2 (SN: 7607) SPL (dBA)	Δ	Device 3 (SN: 1099) SPL (dBA)	Δ	Mean SPL (dBA)	Δ
Open ear (Reference Level)		77.6		77.6		77.6		77.6	
ComTac™ III	Off (passive)	38.6	(-39.0)	38.0	(-39.6)	40.1	(-37.5)	38.9	(-38.7)
	1 (Low)	56.4	(-21.2)	57.5	(-20.1)	57.4	(-20.2)	57.1	(-20.5)
	2	62.2	(-15.4)	63.4	(-14.2)	63.3	(-14.3)	63.0	(-14.6)
	3	68.2	(-9.4)	69.4	(-8.2)	69.3	(-8.3)	69.0	(-8.6)
	4 (High)	**74.2**	**(-3.4)**	**75.4**	**(-2.2)**	**75.2**	**(-2.4)**	**74.9**	**(-2.7)**

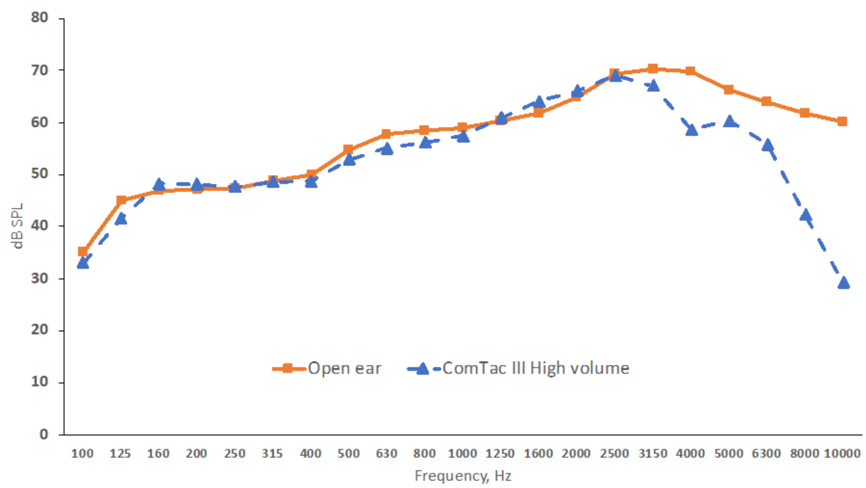


Figure 152. Mean frequency response of 70 dBA pink noise measured using KEMAR under open ear and ComTac™ III devices at high volume setting by 1/3 octave-band frequency.

Independent Variable – Stage of Training

Two within-subjects stage of training levels were used in this experiment: pretest and posttest (Figure 148). Three sets of pretests and posttests were administered to each participant, one for each listening condition, following the prescribed counterbalanced order for each participant. Each pretest and posttest for a listening condition consisted of a total of 36-presentations, three presentations from each of the 12 loudspeaker or 12 remote firing device locations. For all participants, the pretests were completed in one, approximately 30-minute testing session and the posttests were completed in one, approximately one-hour testing session.

As noted earlier, posttests were administered within one to three days of the pretest for the untrained group (mean=1.8 days, *SD*=0.8 days) and within one to three days of the final training session for the trained group (mean=1.7 days, *SD*=0.9 days).

Pretests were conducted in-office using the PALAT system. The pretest employed a dissonant, non-harmonically related, tonal signal comprised of 104, 295, 450, 737, 2967, 4959, 7025, and 7880 Hz (Casali & Lee, 2019). A series of auditory localization studies successfully demonstrated that the dissonant signal provides binaural and monaural cues necessary to test and train localization (Casali & Robinette, 2014; Casali & Lee, 2016a; Casali & Lee, 2016b; Cave et al., 2019). The dissonant tone frequencies were selected to provide the predominant localization cues accessible using the following mechanisms: interaural timing differences (ITDs), interaural level differences (ILDs), and pinnae spectral cues. ITD cues dominate localization of sounds below 1500 Hz as the wavelength must be able to "bend around" the diameter of the head to render timing cues (Moore B. C., Space perception, 1997). ILDs occur when the near ear receives a more intense signal than the far ear (Emanuel, Maroonroge, & Letowski, 2009). ILD cues are dominated by higher frequencies with frequencies above 2000-3000 Hz providing the most information (Moore B. C., 1997). Lastly, the highly contoured surface of the pinnae and successive funneling into the ear canal resonates higher frequencies (Emanuel et al., 2009). This contouring creates spectral changes in the signal even with small changes in sound location, particularly in the 3000-8000 Hz region (Pickles, 1988). Figure 153 displays the spectral content of the 70 dBA dissonant signal presented by the PALAT system.

Posttests were conducted in-field on a rural, lightly wooded farm using a military relevant auditory stimulus in the form of a .22 caliber blank gunshot. Both the experimental site and stimulus were previously used in a sound localization study using active HPDs (Talcott et al., 2012). The investigator measured the spectral content of the .22 blank gunshot in situ and verified that the stimulus contains frequency content necessary to provide interaural timing differences (ITDs) and interaural level differences (ILDs) (Figure 153).

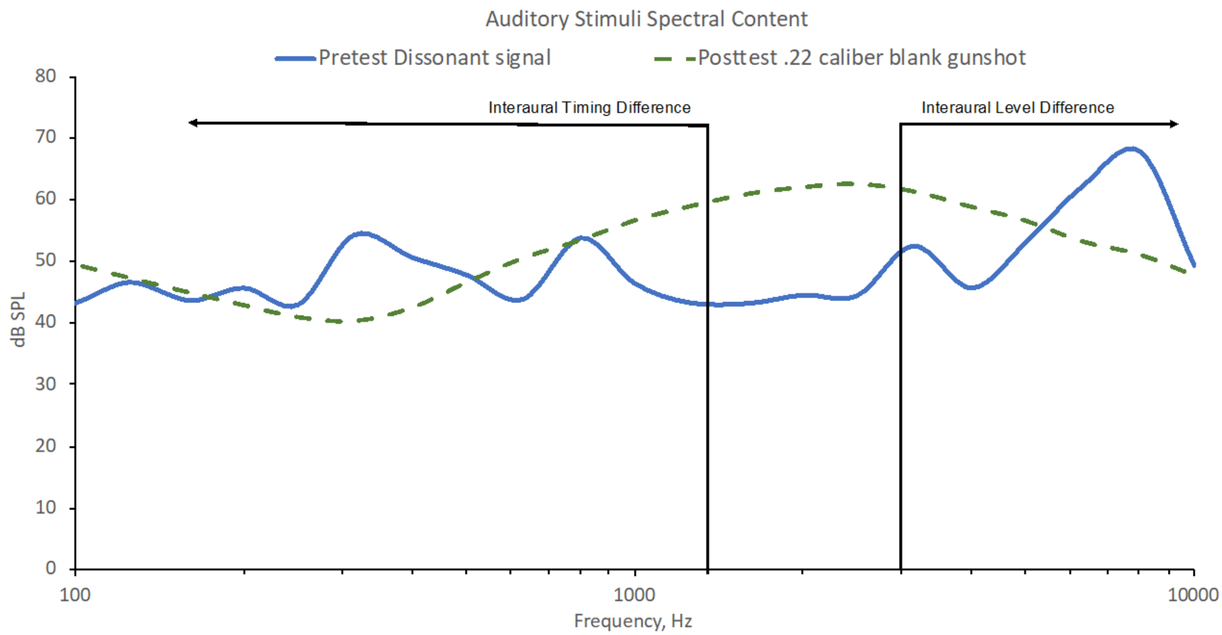


Figure 153. 1/3 octave-band spectral content of pretest dissonant training signal (played on PALAT system) and posttest .22 caliber blank gunshot by 1/3 octave-band frequency. Recorded at the participant's ear in office environment (for PALAT) or at outdoor field site (for transfer-of-training test). Overall sound pressure level of 70 dBA for both signals.

4.3.2 Dependent Measures

Three classes of dependent measures were used to test localization performance: 1) localization accuracy, 2) response time, and 3) subjective ratings (listed in Figure 148). The following sections describe each dependent measure in detail.

Localization accuracy

Three measures of localization accuracy were recorded and analyzed: 1) absolute correct responses, 2) ballpark correct responses, and 3) number of front-back errors. Each test in this investigation presented three signals (dissonant tone or gunshot) from 12 locations in random order for a possible maximum score of 36 correct on each test. The 12 signal locations were separated azimuthally by 30° resembling the 12-hour positions on an analog clock face. U.S. Military Service Members are trained to identify and communicate threat direction or points-of-interest using the 12 clock face number positions with 12 o'clock serving as the frontal midline reference (Department of the Army, 2017). For example, if a military unit were on a patrol walking through the woods in a northerly direction and heard gunshots from an enemy located directly to the east, the members of the unit would yell, "contact, enemy three o'clock." Thus, the investigator decided to present signals from the 12 clock face azimuthal locations. However, 24 response locations were provided to allow the participant to select a direction between two adjacent clock face positions if they were unsure of the exact signal location. Figure 154 shows the test screen from the Surface Pro computer tablet that the participant used during the pretest and posttest, displaying 24 response options (black circles) and 12 signal locations (black circles marked with yellow numbers).

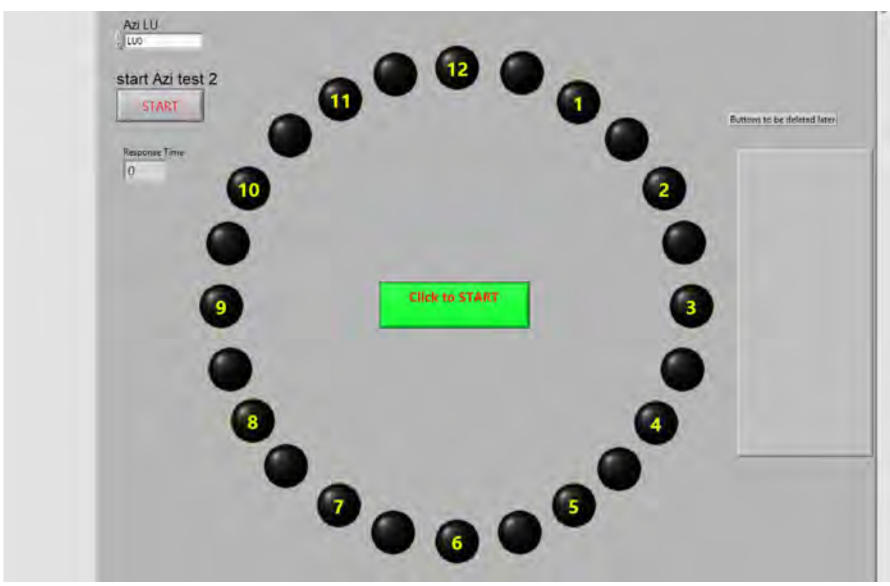


Figure 154. Participant pretest and posttest screen displaying 24 response options (black circles) and 12-signal locations (black circles with yellow numbers).

1. *Absolute correct responses:* the total number of occurrences in which the participant responded with the exact azimuthal location of the signal location. Figure 80 displays an example of an absolute correct response indicated by the arrow if the signal originated from the one o'clock position.

2. *Ballpark correct responses:* the total number of occurrences in which the participant responded with an azimuthal location within ±15° of the location of the presented signal. A ballpark score is achieved if the participant response matches the exact location of the signal location (absolute correct) or if the response identifies a speaker location directly adjacent, left or right, to the presented signal location. Figure 155 displays an example of the range for a ballpark correct response indicated by the grey shaded region of a signal originating from one o'clock.

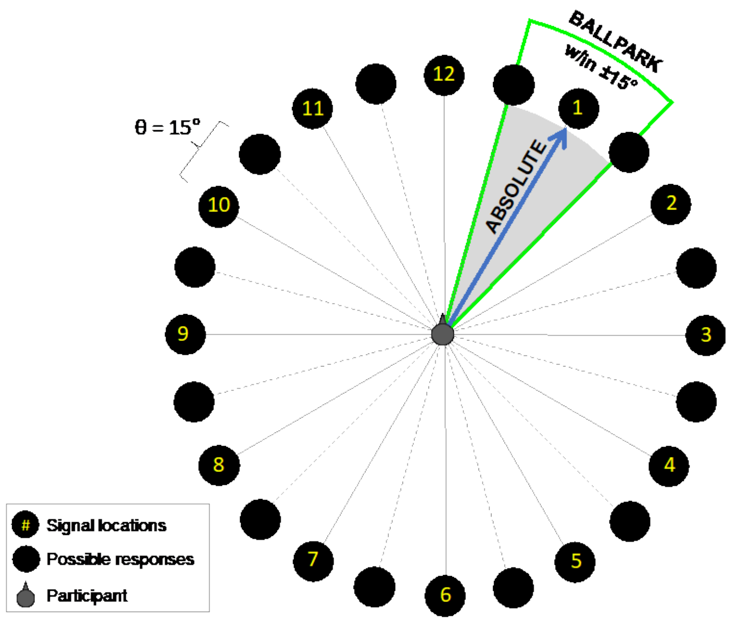


Figure 155. Absolute correct response (arrow) and ballpark correct response region (grey shaded region) when the signal emanates from the one o'clock position.

3. *Front-back reversal errors:* the total number of occurrences in which the participant responded with an azimuthal location in the back (to the rear of participant) from four o'clock to eight o'clock (120-degees to 240-degrees) when the signal was presented in the front from ten

o’clock to two o’clock (300-degrees to 60-degrees) and vice-versa. This window for front and back reversals is consistent with the new ANSI 3.71 standard window from 290-degrees to 70-degrees in front of the participant and 110-degrees to 250-degrees behind the participant (American National Standards Institute (ANSI), 2019). However, this experiment’s operational definition of front-back reversal differs from the ANSI standard by allowing front-back reversals to occur if the difference between the source and response crosses the median plane. For example, a front-back reversal occurs in this experiment if a sound originates from the seven o’clock position and the participant responds with the one o’clock position. The investigator felt this offered a more realistic operational definition of front-back reversals for auditory situation awareness in U.S. Military operations. If a U.S. service member perceived a gunshot from the one o’clock position (in front of them) that actually originated from seven o’clock (behind them), then the service member would have made a front-back reversal that could be detrimental to survivability. Figure 156 displays the front and back regions where either the signal originated and the response was selected to constitute a front-back reversal error if the signal and response were in opposite regions.

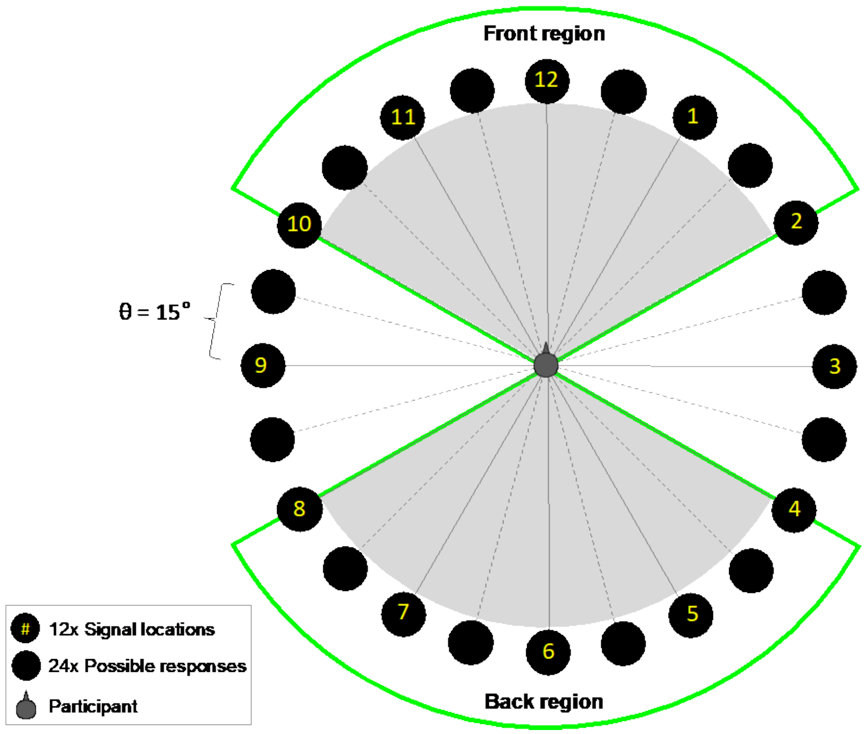


Figure 156. Front and back regions (shaded regions) depicting the range of signal locations and response locations for possible front-back reversal errors.

Response time

Response time was measured as the duration of time occurring from signal onset, dissonant tone (PALAT) or gunshot (field test), to the participant response selection on the computer tablet. Response time was automatically calculated via the LabView computer program on the Surface Pro computer tablet. The response time clock onset was triggered by the participant selecting the green “Click to START” icon located in the center of the test screen (Figure 154). Selecting the “Click to START” icon simultaneously presented the dissonant tone or the gunshot for the tests conducted in the PALAT or field, respectively. The response time clock offset occurred when the participant selected a speaker icon on the response display. A window located on the left side of the test screen displayed the running clock. After response selection, the display showed the most recent response time, allowing the participant to view their response time. Response times were recorded in 100 millisecond resolution. The maximum allowable response time was set at 10 seconds. Mean response times were calculated for each pretest and posttest and used as the dependent measure score.

Subjective ratings

Participants completed a questionnaire at the conclusion of every pretest and posttest for every listening condition (Appendix K). Every questionnaire included the same six questions so that comparisons could be made between tests. Participants in the trained group were asked an additional question at the conclusion of the posttest related to the effect of training on localization performance. All questions used a semantic differential, bipolar rating scale with seven discrete choices (example shown in Figure 157). Question 7 was used only for trained group after each listening condition during the posttest.

Every questionnaire included the same six questions so that comparisons could be made between tests. Participants in the trained group were asked an additional question at the conclusion of the posttest related to the effect of training on localization performance. All questions used a semantic differential, bipolar rating scale with seven discrete choices (example shown in Figure 82). Question 7 was used only for trained group after each listening condition during the posttest.

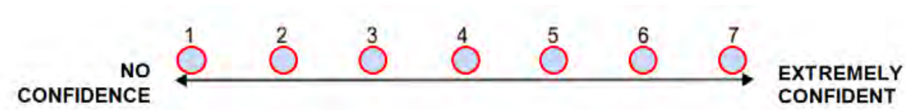


Figure 157. Example of semantic differential rating scale.

1. *Confidence in ability to localize*:

   "Rate how **confident you were** in your ability to locate sounds under this listening condition" from 1 (no confidence) to 7 (extremely confident).

2. *Perceived localization accuracy*:

   "Rate your perceived **accuracy** to determine sound location under this listening condition" from 1 (highly inaccurate) to 7 (highly accurate).

3. *Difficulty in judging location of sound*:

   "Rate how **difficult** it was to judge the **location** of the sounds under this listening condition" from 1 (extremely difficult) to 7 (extremely easy).

4. *Perceived reaction time*:

   "Rate your perceived **reaction time** in determining the sound location under this listening condition" from 1 (extremely slow) to 7 (extremely fast).

5. *Comfort of TCAPS device or open ear*:

   "Please rate how **comfortable this hearing protection device condition** (or open ear) was while wearing it during the experiment" from 1 (extremely uncomfortable) to 7 (extremely comfortable).

6. *Likelihood to use TCAPS device*:

   "How likely would you be to wear this hearing protection device during a task similar to this experiment that required sound localization if you had access to this hearing protection device" from 1 (extremely unlikely) to 7 (extremely likely).

7. *Preparedness from training*:

   "Rate the degree of preparedness you felt as a result of the training on the localization system (ring of loudspeakers) compared to the task of localizing .22 blank gunshot sounds" from 1 (extremely unprepared) to 7 (extremely prepared).

4.3.3 Participants

This human-subjects experiment was approved by the Western Institutional Review Board (WIRB protocol #20190789, VT-IRB #19-176, Appendix J) which acted as the assigned review board for Virginia Tech as of the date of this research. Participants were required to be between the ages of 18 to 45 years with up to 25% females in order to generalize to the U.S. Military population (Defense Manpower Data Center (DMDC), 2017). The study sample consisted of 24 participants: 18 males and 6 females, age 19 to 34 years with a mean age of 26.3

years (*SD*=3.8). Two additional male participants were involved in the study but were replaced due to one failing to complete the study (illness) and one participant's performance resulting in a statistical outlier on two performance measures (discussed later in section 3.4 Results). Participants were recruited from Virginia Tech and the surrounding communities. Participants were compensated $10 per hour and received a $30 bonus upon completion of the study. Transportation was provided to the field site for all but one participant who was reimbursed for mileage at the Virginia Tech-assigned rate of $0.58 per mile.

Participants were required to have normal hearing and no previous experience with localization studies or training. All participants were screened for hearing thresholds not to exceed 25 dB HL at 250, 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz, with no threshold difference between each ear to exceed 15 dB (bilaterally-symmetrical). Following the participant's informed consent, they were otoscopically inspected to check for ear canal obstructions, irritation, or infections that could affect localization performance. One of the investigators, an Active Duty Army Audiologist, performed the otoscopic inspections and administered the hearing tests. If the participant passed otoscopic inspection, a manual pure-tone audiogram using a standard Hughson-Westlake procedure was conducted using a Beltone Electronics Corporation Model 119 Audiometer (SN: 10B0561, calibrated 26 December 2019). The test was performed in the VT-ASL portable test booth located in the same room as the PALAT system (Figure 158). Table 105 displays the mean pure-tone hearing level thresholds (dBHL) for all participants and by group. Following the audiogram, participants were screened to ensure no prior experience with localization training or TCAPS devices (Appendix L).

Table 105. Mean pure-tone hearing level thresholds (dBHL) for all participants and by group.

		Frequency (Hz)							
	Ear	250	500	1000	2000	3000	4000	6000	8000
All participants	Right	10.2	6.3	4.6	0.0	4.2	2.3	5.4	9.2
	Left	8.5	6.3	5.2	-0.2	2.1	4.6	5.0	8.5
Trained	Right	10.0	5.8	2.5	-0.8	3.3	1.7	6.7	7.5
	Left	9.2	7.9	5.8	1.3	3.8	5.0	6.3	12.9
Untrained	Right	10.4	6.7	6.7	0.8	5.0	2.9	4.2	10.8
	Left	7.9	4.6	4.6	-1.7	0.4	4.2	3.8	4.2

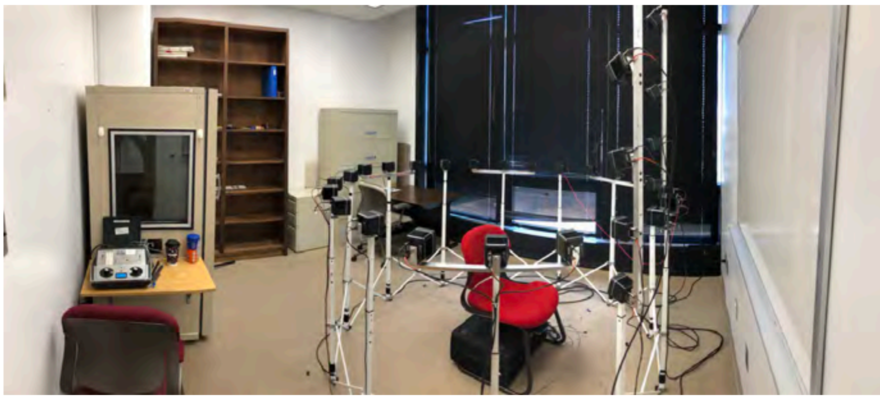


Figure 158. Portable audiometric booth (left) co-located with the PALAT system in the office environment used for training and pretesting.

4.4 Phase III Apparatus

The Phase III investigation was conducted in two locations: an office environment on the campus of Virginia Tech where training and pretesting occurred and an outdoor field-conducted posttest on a rural farm in Pulaski County, Virginia.

4.4.1 In-Office: PALAT System

The pretest and localization training were conducted using the PALAT system apparatus. The PALAT system was located in a small office room on the fifth floor of Whittemore Hall at Virginia Tech. The PALAT system apparatus was operated in the same office during Phase II and was found to provide a similar localization testing and training experience and performance results as the full-scale, laboratory grade DRILCOM system used in previous localization experiments (Casali & Robinette, 2014; Casali & Lee, 2016a; Casali & Lee, 2016b). The

PALAT room was approximately 13.5 feet by 12.5 feet and contained typical office furniture including a desk, chairs, wooden bookshelf, metal storage cabinets, dry-erase board, metal window blinds, carpeted floor, and dropped panel ceiling. In addition, a metal portable audiometric booth was located in the corner of the room (Figure 158). The small office space was selected due to its semi-reverberant environment that represents a typical setting where the military (or civilian industry) would employ the PALAT system. Likewise, the investigator left the acoustically reflective furniture inside the office assuming that users of the PALAT system would have the constraint of using the system in rooms designated for other purposes; thus, there was no attempt to "e" the office for uniformity of reflections or other acoustic considerations. In other words, the office environment used to evaluate the PALAT system was believed to be as realistic as possible, representative of that which would be typically encountered in actual military training practice. The PALAT system was positioned in the room so that no speaker was within two feet of any reflective surface but the system was not centered in the room. Centering the PALAT system in the room would be preferred in order produce a more uniform reflective surface. Hartmann (1983) found that early reflections from side walls created the largest decremental effect on localization due to spectral information that conflicted with the direct sound wave. The investigator decided against centering the portable system assuming that future users of the PALAT system may have similar limitations due to varying room sizes and shapes.

The small semi-reverberant office used for PALAT testing and the hemi-anechoic laboratory room housing the DRILCOM system (for comparison) were tested to find the ambient noise floor and reverberation time (RT60). Measurements were made with a Larson-Davis Model 831 Sound Level Meter (SN: 0002486) with a ½-inch Larson-Davis 2575 measurement microphone (SN:LW131180) and Larson-Davis PRM831 Preamp (SN: 017153). The microphone was calibrated at 94.0 dBA (1000 Hz tone) using a Quest QC-20 Calibrator (SN: QOA070051). The investigator performed five measurements, in the center of the room and approximately one-meter from each room corner. RT60 measurements were taken using an impulse noise at approximately 120 dBA produced by hitting together two-wooden 2x4 blocks. The RT60 measurements were calculated using a 30 dB decrease in level to avoid limitations posed by the noise floor. The calculation to extrapolate the RT30 values to RT60 values was performed automatically by the sound level meter. Noise floor and RT60 values are shown in Table 106 for both the DRILCOM and PALAT rooms.

Table 106. Mean noise floor and reverberation time (RT60) measurements of the DRILCOM and PALAT rooms, as measured in octave bands.

		Frequency (Hz)					
		250	500	1000	2000	4000	8000
DRILCOM Room	Noise Floor (dB SPL)	34.0	30.5	31.3	33.9	36.6	40.0
	RT60 (ms)	408	272	182	144	119	110
PALAT Room	Noise Floor (dB SPL)	42.1	33.1	31.5	33.5	36.6	40.0
	RT60 (ms)	407	402	348	339	410	396

The PALAT system is a 2-meter diameter circular, horizontal and vertical (front) localization apparatus consisting of 32 loudspeakers with 24 loudspeakers (horizontal) and eight additional vertical loudspeakers housed in a semi-reverberant room (Figure 159). Two of the horizontal loudspeakers are used during elevation testing to provide 10 vertical loudspeakers. Only azimuth speakers were used in this study. All loudspeakers are separated by an angle of 15° from the center of the apparatus, or center head position of the participant. The horizontal loudspeakers are located one meter from the participant. The loudspeakers are mounted on a portable, collapsible frame consisting of 12 telescopic poles. The telescopic poles allow for horizontal loudspeaker heights of 43.5, 45.5, and 47.5 inches above the floor, to accommodate seating heights of different individuals. The speaker heights were set to 45.5 inches above the floor for the duration of the in-office experiment. The PALAT system is controlled by the user seated in the middle of the loudspeaker array via a Microsoft® Surface Pro running a LabView software program. The system uses Cambridge Audio Minx Min 12 loudspeakers with a 2.25-inch single cone driver, a Stewart Audio AV30MX-2 two channel stereo mixer amplifier, and a Numato 32 channel USB relay module. Under the participant chair, a pink noise generator mounted inside the equipment case emits 55 dBA of pink noise. To generate the pink noise signal, a Mystic Marvels LLC. PNG-400 Pink Noise Generator routed through a Stewart Audio AV30MX-2 amplifier and Minx Min 12 loudspeaker is used (Figure 159). The 55 dBA pink noise masks extraneous sounds during the experiment while maintaining a +15 dB signal to noise ratio given the 70 dBA dissonant signal. Figure 160 displays the spectral content of the dissonant training signal and the pink noise as measured in the PALAT system. The microphone was placed in the center of the PALAT system at the approximate height of the participant's head. The two poles housing the elevation speakers were removed during all in-office testing and training in order to present a more uniform apparatus for azimuthal testing (all 24 speakers aligned on one horizontal plane).

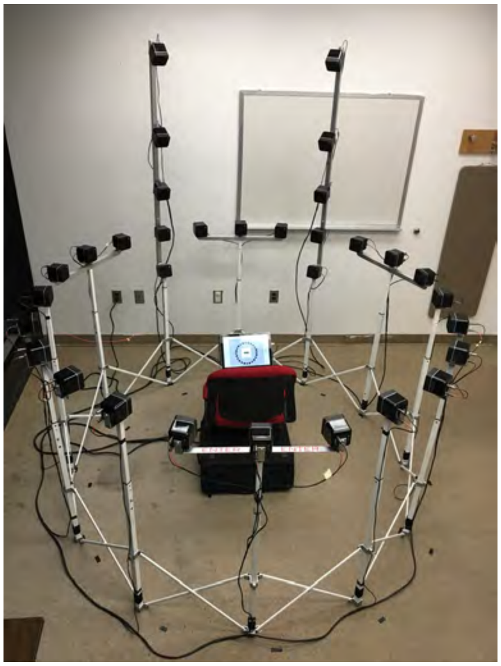


Figure 159. PALAT system apparatus located in a semi-reverberant office room at Virginia Tech.

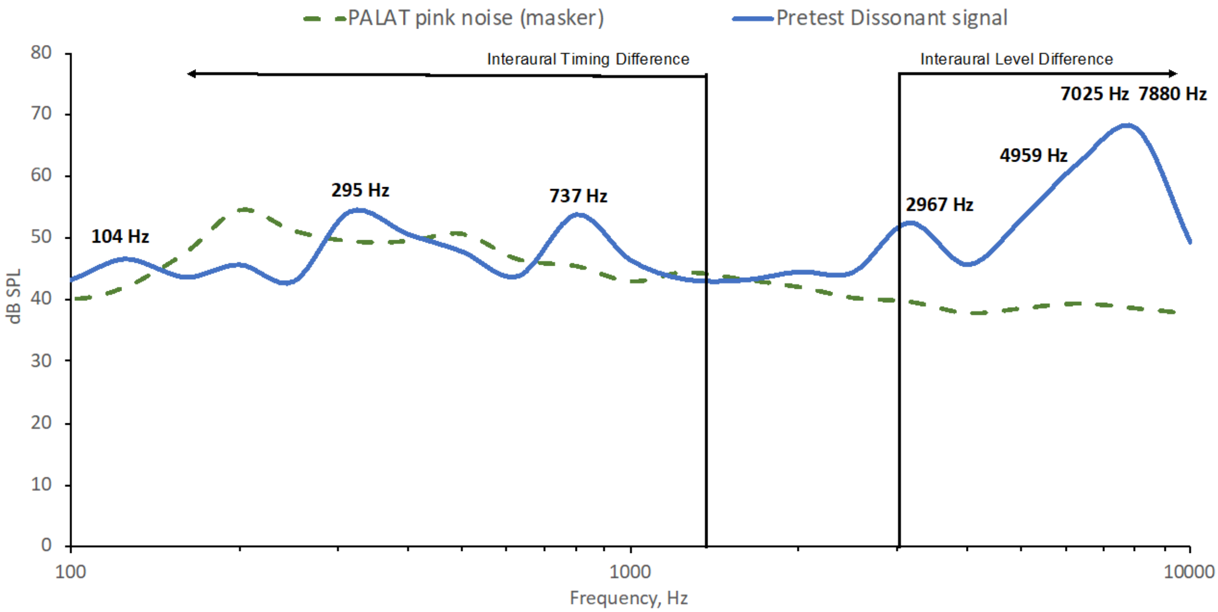


Figure 160. 1/3-octave band spectral content of PALAT system pink noise (green dashed line) and dissonant signal (blue solid line) in 1/3 octave-band frequency. Eight pure tones comprising the dissonant tone are labeled above the respective frequency.

The participant-controlled computer tablet was used to initiate the dissonant signal presentation and record the user response, azimuth and time. The participant used a stylus pen to operate the controls on the computer tablet. Prior to testing, the participant was given instructions and received a demonstration on how to control the computer tablet and software program. When ready, the participant selected the “Click to START” icon on the computer tablet. The LABVIEW program then sent a signal to the USB-controlled relay switch to close the relay corresponding to the presentation speaker. Simultaneously, LABVIEW transmitted the dissonant audio signal through a 3.5 mm audio cable through the amplifier and relay switch to the presentation speaker. Signal transmission from LABVIEW also triggered the onset of the response timer. To randomize the speaker location, the software program used a random number generator to select a loudspeaker position with the constraint of requiring three presentations from each loudspeaker location during each test. The participant registered a response by selecting one of the 24 loudspeaker locations by touching the stylus to one of the black circles corresponding to the loudspeaker locations, as accurately and quickly as possible. Upon response, the LABVIEW software stopped the response timer and recorded the response to a Comma Separated Values (CSV) file. The software program recorded the participant number, listening condition, test type (azimuth or elevation), loudspeaker source location, participant response location, response time, date, time, and stage of training for each trial. The software program also calculated and displayed a running total absolute score and running total ballpark score for the test (example shown in Figure 161). The CSV file was saved to a shared folder so participant scores could be accessed after each test.

Subject	Condition	Type	Source	Response	Absolute	Ballpark	Response Time	Date	Hour	Stage of training
S20	COMTAC III Unity	Azimuth	18	18	1	1	2.2	4/15/2019	10:17 AM	LU0
S20	COMTAC III Unity	Azimuth	16	22	1	1	2.6	4/15/2019	10:17 AM	LU0
S20	COMTAC III Unity	Azimuth	0	4	1	1	3.9	4/15/2019	10:17 AM	LU0
S20	COMTAC III Unity	Azimuth	20	21	1	2	4.4	4/15/2019	10:17 AM	LU0
S20	COMTAC III Unity	Azimuth	6	6	2	3	1.7	4/15/2019	10:17 AM	LU0
S20	COMTAC III Unity	Azimuth	4	4	3	4	2.5	4/15/2019	10:17 AM	LU0

Figure 161. Example CSV file output of PALAT system.

4.4.2 In-Field Site

The field-conducted posttest was conducted outdoors on a rural farm located in Pulaski County, Virginia. The participant stood in the center of 50-foot circular clearing, surrounded by a lightly wooded forest of relatively mature trees in which twelve hard-wired, but remotely operated blank-firing devices (Figure 162). The site was previously used for a sound localization study where eight firing positions surrounded the participant (Talcott et al., 2012). Due to the extended time period between the 2012 and current experiments, and the addition of four firing positions, the field site was re-cleared of obstructions and reoriented to align the twelve firing positions so that no large trees were in the direct line of sight (or direct sound ray) of each remote firing device. The clearing in the woods was situated at the top of a small rolling hill. The hill rolled off in a nonuniform pattern with the steepest roll-off in the northeast (12 o'clock), fairly flat terrain to the southeast (three o'clock) and northwest (nine o'clock), and a gradual roll-off to the southwest (six o'clock). Based on the terrain features, the investigator established the 12 o'clock remote firing position to the northeast at a magnetic azimuth of 32°. The investigator used an optical level and U.S. Military tritium lensatic compass to mark the 12 firing position azimuths separated by 30° angles measured from the center point. Table 107 displays the magnetic azimuth of 12 firing device positions measured from the center point where the participant stood.

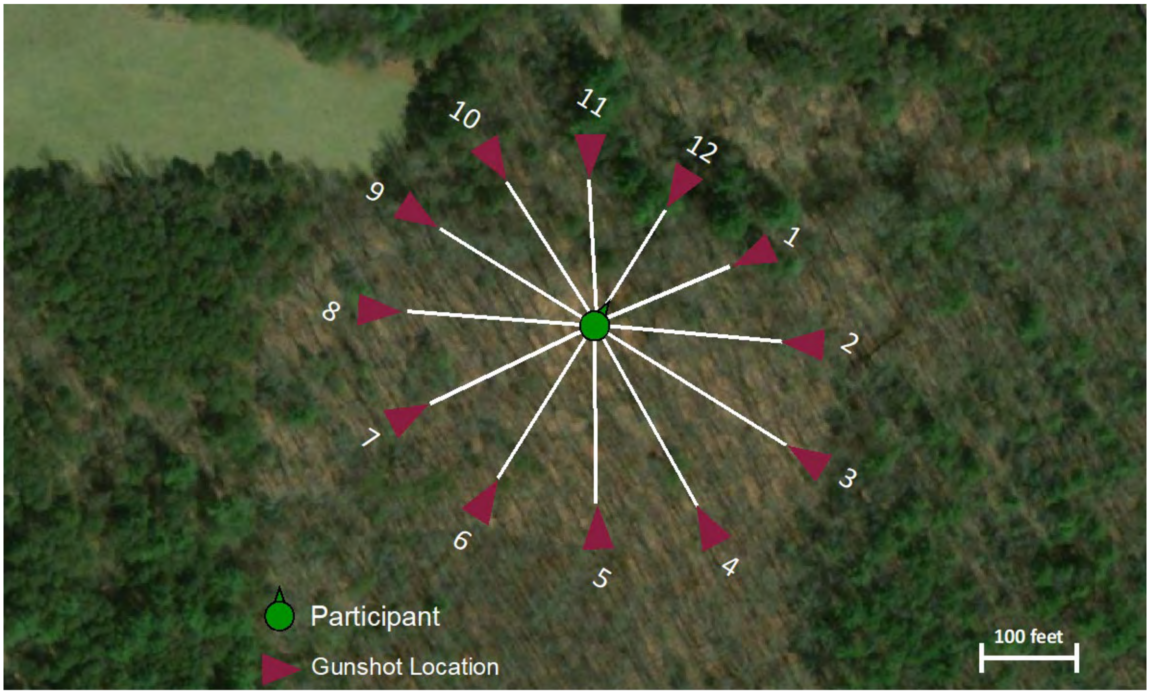


Figure 162. Aerial view of field site layout with 12-remote firing device positions, located around the participant.

Table 107. Magnetic azimuth (degrees), radial distance (feet), and sound pressure level (dBA max) at center position of each remote firing device location.

Remote device location	12	1	2	3	4	5	6	7	8	9	10	11
Magnetic azimuth (degrees)	32	62	92	122	152	182	212	242	272	302	332	2
Radial distance (feet)	150	160	200	250	215	210	220	200	200	192	175	160
SPL (dBA max)	69.4	71.7	70.2	72.1	67.7	70.5	70.8	70.9	72.5	70.9	71.6	72.0

The investigator matched the blank gunshot signal in the posttest with the 70 dBA-max signal used in the pretest. Sound levels were controlled by adjusting the distance of the remote firing devices from the center to achieve the target level. The 12 distances to the firing devices were adjusted until the mean sound pressure level from three .22 caliber blank gunshots resulted in approximately 70 dBA-max at each location. The ambient noise floor was measured to ensure that the 70 dBA-max gunshot signal was easily detectable and provided interaural timing and interaural level cues used in sound localization (Figure 163). Table 107 displays the radial distance and mean sound pressure level measured for each firing position. Of note, the investigator discovered slight inconsistencies in sound pressure levels produced by blank rounds fired from the same remote firing device. These variations were possibly due to variations in the manufacturing process or an effect of how the gunpowder was ignited in each blank round, but were obviously not under experimental control and thus constituted a small random variance source. These fluctuations were randomly distributed throughout all experiments effecting both groups and each listening condition.

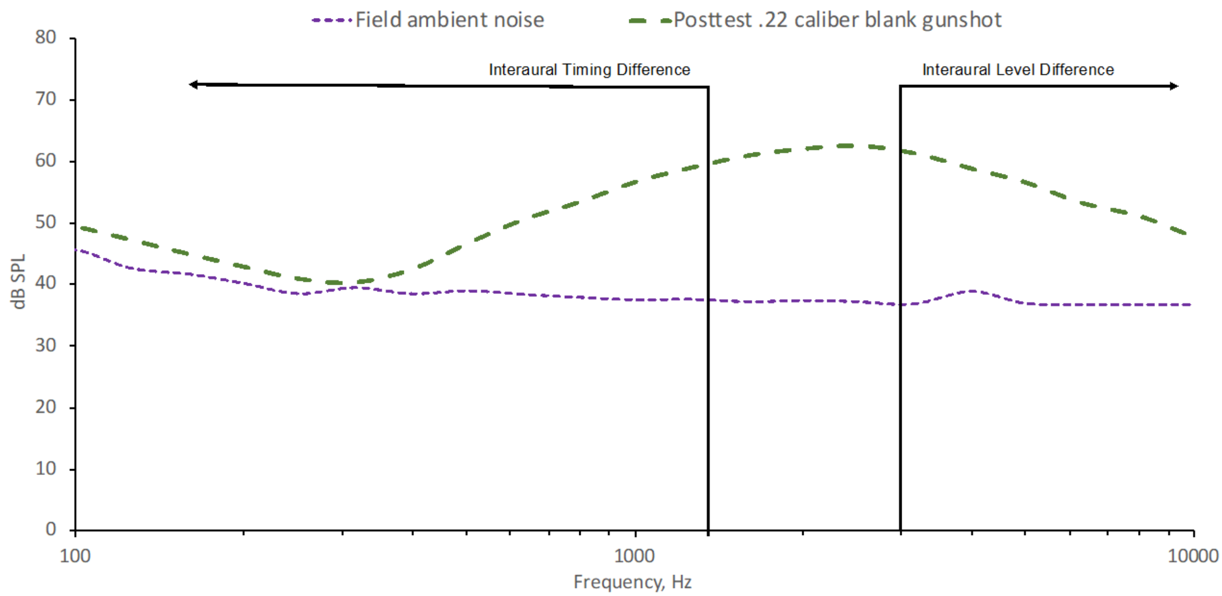


Figure 163. Mean 1/3-octave band spectral content of .22 caliber blank gunshot from all 12 firing device positions (dashed green line) versus ambient noise floor (dotted purple line) measured in 1/3 octave-band frequency at the participant center head position.

Figure 164 shows a panoramic photograph of the field site from the participant's perspective marked with each firing device location by clock face number. A sign displaying the number 12 was placed approximately 50 feet directly in front of the participant along the 12 o'clock azimuth (and only at that position) to help orient the participant and give them a visual reference point to focus on prior to firing each blank gunshot (Figure 165).

Figure 164. Field site panoramic picture (from left at 7 o'clock to right at 6 o'clock), with clock face positions identified. Only the 12 o'clock position actually included a sign during the conduct of the field experiment.

Figure 165. Participant view of computer tablet used to operate the field posttest and 12 o'clock reference sign to orient participant.

Outdoor Azimuthal Gunshot Presentation System

In Talcott et al.'s (2012) study at the same field site, .22 caliber blanks were fired from revolver pistols by the investigators. Three investigators walked to each of the eight firing positions located at a radial distance of 150 feet in a predetermined order. Each participant spent approximately 4 hours at the field site conducting sound localization testing in the Talcott et al. (2012) protocol. In the present study, the number of firing locations increased to 12, and the radial distance increased by 100 to 150 feet at most firing locations. Due to the labor and time-intensive requirements of the field experiment, the investigator automated the .22 caliber blank gunshot delivery. The investigator designed a remote firing device that consisted of three electric magnetic locks, three spring loaded firing pins, and a bar to hold three .22 caliber blanks (Figure

166). The electric magnetic lock contained a solenoid attached to a small lever with a hook that released a U-shaped lock when supplied with approximately 12 Volts. A key ring attached to the top of the spring-loaded firing pin inserted into the U-shaped lock. The firing pin was held under tension above the .22 caliber blank rounds. The three firing pins were aligned directly over a ¾ by 1-inch aluminum bar that held three .22 caliber blanks. When the firing pin released, the pin struck the rim of the rim-fired .22 caliber blank.

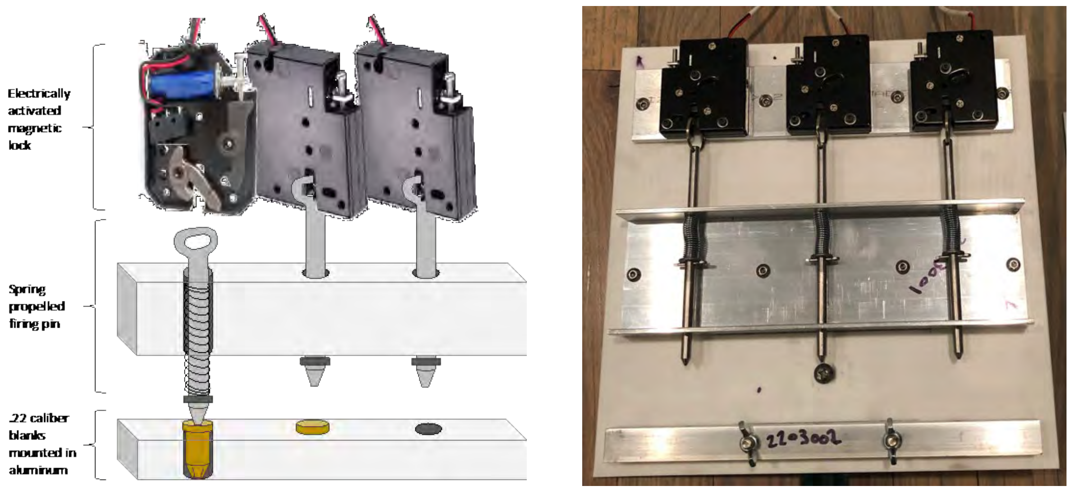


Figure 166. Remote firing device design concept sketch and final product.

The field localization test employed 12 remote firing devices each containing three separately controlled firing mechanisms (electric magnetic lock, spring loaded firing pin, and .22 caliber blank). An additional firing mechanism was built on the reverse side of the four firing devices that were placed at the 12 o'clock, 3 o'clock, 6 o'clock, and 9 o'clock positions. These additional devices served as a single familiarization round prior to each posttest. The remote firing devices were mounted on a steel u-post at a height of approximately 4 feet above ground (Figure 167). The 12 o'clock firing device was adjusted to a height of 6 feet above the ground to compensate for the steep roll-off directly in front of the participant. Sound measurements were recorded and analyzed to determine the optimal orientation of the remote firing device to present a relatively consistent sound signature from each firing position. The investigator found that orienting the firing device perpendicular to the participant reduced the visual signature and resulted in the most consistent sound levels (Figure 167).

Figure 167. One remote firing device containing three remote firing mechanisms mounted on a steel u-post located in the wooded forest at the field localization site (Left: Front view of remote firing device, Right: Profile (side) view as seen from direction of the participant), which reduced the visual signature.

The remote devices were hard-wired to a control box containing 12, four-position rotary switches, an LED power indicator light, a safety toggle switch, 8-ampere fuse, and Numato® USB 16-channel relay switch (Figure 168). The remote firing system was initiated by a participant-controlled Microsoft® Surface Pro computer tablet running a localization testing LabView code, almost identical to the interface used to control the PALAT system. The remote firing devices were located between 150 to 300 feet away from the participant and connected to the control box by 18-gauge wires. Due to the voltage drop across the long distance of electrical wire, the investigator used two, 12-Volt gel car batteries connected in series to supply 24 Volts from the control box. The resulting voltage at each firing device measured between 10 to 15 Volts, depending upon the radial distance of each remote firing position. One of the car batteries powered the 12-Volt relay switch. The control box was operated by an investigator seated approximately 10 feet behind the participant. The investigator ensured the correct firing device and firing mechanism were selected on the control box prior to the participant initiating the blank round by clicking on the “Click to START” button on the computer tablet. Figure 169 displays the components and general wiring of the remote firing system (Appendix M).

Figure 168. Remote firing device control box.

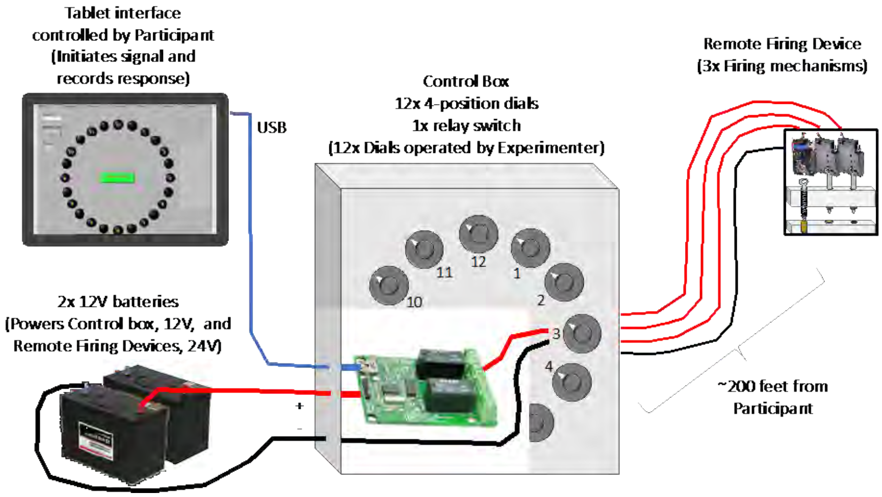


Figure 169. Remote firing device block diagram with all major components.

Outdoor Auditory Localization Data Capture System

The same participant-controlled computer tablet used in the in-office pretest was used to initiate the .22 blank gunshot and record the user response, response azimuth and time. The computer tablet was placed on a music stand located directly in front of the standing participant, but well below the participant's head to prevent interference with gunshot's direct sound wave (Figure 170). The participant was given instructions and received a demonstration on how to control the computer tablet and software program prior to testing. Prior to each trial, or blank gunshot, one of the investigators seated behind the participant set the rotary dial to the proper firing position and turned on the control box switch. The investigator then informed the participant that the system was armed by saying "READY." When ready, the participant selected the "Click to START" button on the computer tablet. The LabView program then signaled to the USB-controlled relay switch to close the corresponding switch located inside the control box. Closing the switch allowed 24 Volts, supplied by the two car batteries, to be routed through relay switch, to the investigator-activated rotary dial, and to the desired firing mechanism. Six randomly-generated firing sequences were preprogramed into the LabView software. One sequence was randomly assigned to each posttest with the constraint that a unique sequence was used for all three listening conditions. The investigator used a sequence order checklist to ensure the correct rotary dial selection. The hard copy sequence checklist was synchronized with the order in LabView to enable automated scoring according to absolute and ballpark criteria. Just as in the in-lab study, after signal presentation the participant selected their response on the 24-icon display via stylus. The participant was then prompted via display to speak their response. A second investigator seated behind the participant recorded the verbalized response as a backup data source. The participant response on the computer tablet triggered response timer offset and recorded the time to a Comma CSV file. The software program recorded the participant number, listening condition, test type (azimuth or elevation), loudspeaker source location, participant response location, response time, date, time, and stage of training for each trial. The software program also calculated and displayed a running total using the absolute and ballpark criteria for the test for the investigator (example shown in Figure 161). The CSV file was saved to a Dropbox folder and shared through a mobile hotspot so that participant scores could be accessed after each test by the investigator.

Figure 170. Participant-controlled computer tablet placed on music stand at the center of the field experiment site.

4.5 Phase III Experimental Procedures

The experimental procedure for this investigation involved four distinct phases: 1) recruitment and screening, 2) pretest, 3) training (experimental condition only), and 4) posttest. The following sections detail the experimental procedures for each of these phases.

4.5.1 Recruitment and screening

Participants were recruited from the Virginia Tech community and surrounding areas via posted flyers (Appendix N), emails, and word of mouth. Potential participants contacted one of the investigators through email and a screening date was scheduled. A copy of the Phase III informed consent (Appendix O) was sent by email to the interested participant prior to their scheduled screening date for review. The potential participants were notified that a hard copy of the informed consent would also be provided at the time of their screening session. At the time of the screening, one of the investigators read aloud through the informed consent with the participant and answered all of the participant's questions. The participant was then given as much time as needed to read through the informed consent before signing. After signing the informed consent, an otoscopic inspection and an audiogram were administered (discussed in section 4.3.3). Following the audiogram, the participant was screened to ensure adherence with demographic requirements (Appendix L). The participant was then scheduled for the remainder of their tests and training sessions.

4.5.2 In-office pretest

On each day and prior to participant arrival, the investigator conducted a calibration of the PALAT system to ensure signal delivery of approximately 70 dBA. The investigator then populated the LabView fields with the participant number, listening condition, and auditory stimulus (Figure 171).

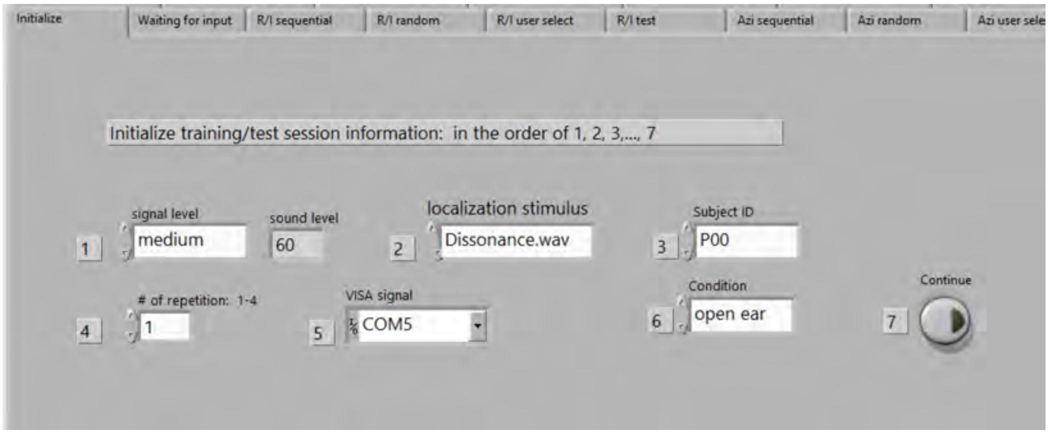


Figure 171. PALAT system initialization screen.

Upon arrival, the participant was given an overview of the PALAT system. Participant instructions appear in Appendix P. The participant was informed of the purpose of the study and given a demonstration of how to operate the computer tablet user interface and localization software program. The investigator ensured the participant was seated in the center of the loudspeaker array. For TCAPS listening conditions, the investigator ensured the TCAPS were turned on and set to the unity gain prior to fitting the participant. Then, to ensure consistency of proper fit, the participant was fitted with their assigned TCAPS device by the investigator. The investigator ensured the TCAPS devices were comfortable and informed the participant to notify the investigator if they experienced any discomfort or acoustic feedback from the TCAPS. Figure 172 shows how a participant was seated and operated the PALAT system for each pretest and training session. The investigator was present in the room during every session. The investigator sat outside of the loudspeaker array behind or to the side of the participant and guided the participant through the testing and training.

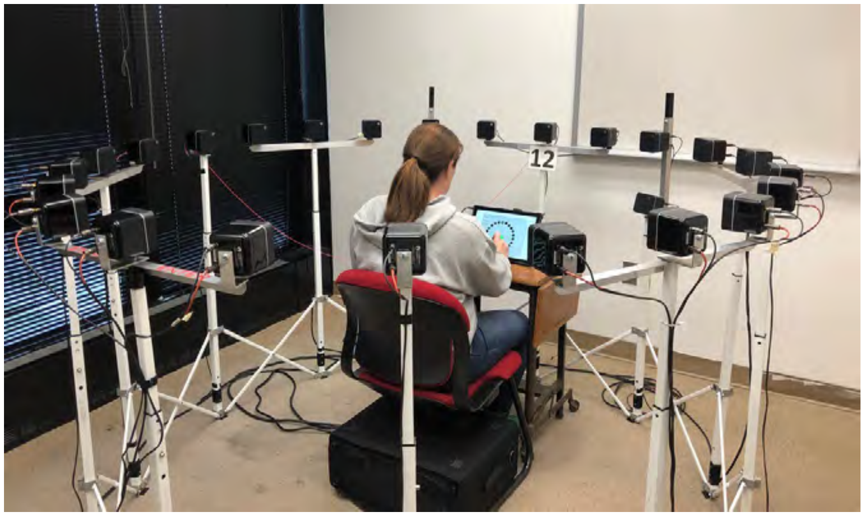


Figure 172. Participant operating the PALAT system.

Prior to each pretest, the participant received a familiarization unit consisting of a total of four presentations of the dissonant signal from the 12, 3, 6, and 9 o’clock loudspeaker positions. The familiarization unit was conducted to orient the participant to the PALAT system, familiarize the participant with the dissonant tonal signal, and allow the participant to practice operating the computer tablet. To perform the familiarization, the investigator selected the “H sequential training” button from the main menu screen (Figure 173). The investigator selected Learning Unit 0 (LU0) from the dropdown menu on the sequential training screen and then selected the “Start” button (Figure 174). The participant was then handed the stylus pen and instructed to wait for the investigator to position themselves outside of the loudspeaker array. Once in position, the investigator instructed the participant to begin the familiarization whenever they were ready by clicking on the “Click to Sound” button located in the center of the sequential training screen and to respond by touching one of the 24 response buttons represented by the black circles on the sequential training screen. Prior to starting the familiarization unit, the participant was informed that the sequence of signal presentation would be emitted from 12 o’clock, then 3 o’clock, then 6 o’clock, and lastly 9 o’clock. However, the participant was allowed to respond on the computer tablet with any of the 24 response options. The participant was instructed to face forward and look at the white sign with the black number "12" prior to clicking on the "Click to Sound" button. The participant was informed that they were free to move their head and rotate their body at the onset of the dissonant signal to aid in localization and target identification. Head movement can be used to overcome a lack of localization cues by creating momentary changes in the sound spectrum at each ear. As a result, localization errors are reduced when the listener is allowed to move their head (Muller & Bovet, 1999; Thurlow & Mergener, 1970). Head movements were allowed in this experiment in order to more closely replicate a U.S. service member operational situation. The participant was reminded to "respond as accurately and as quickly as possible." During the familiarization unit, the participant was allowed to ask the investigator questions. The investigator assisted the participant if they were having trouble operating the system. The PALAT system automatically returned to the main menu screen at the completion of the familiarization unit.

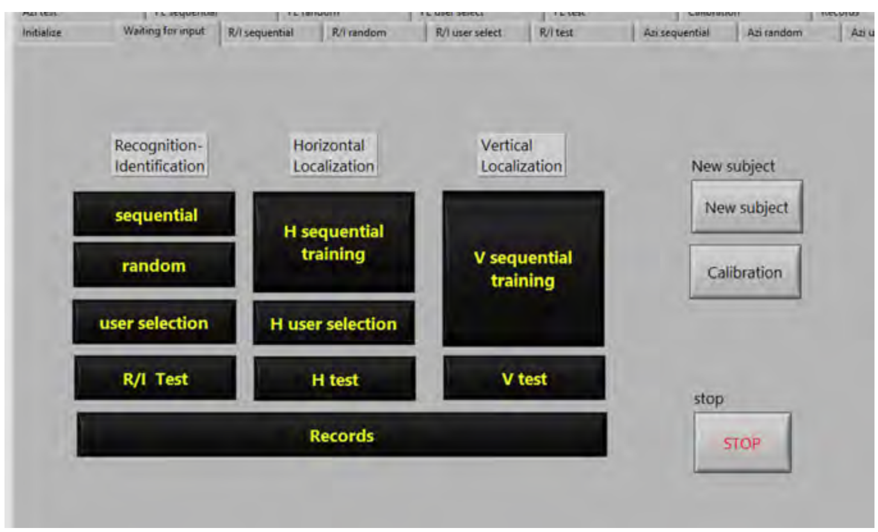


Figure 173. PALAT system main menu screen on computer tablet displaying training and testing protocol options.

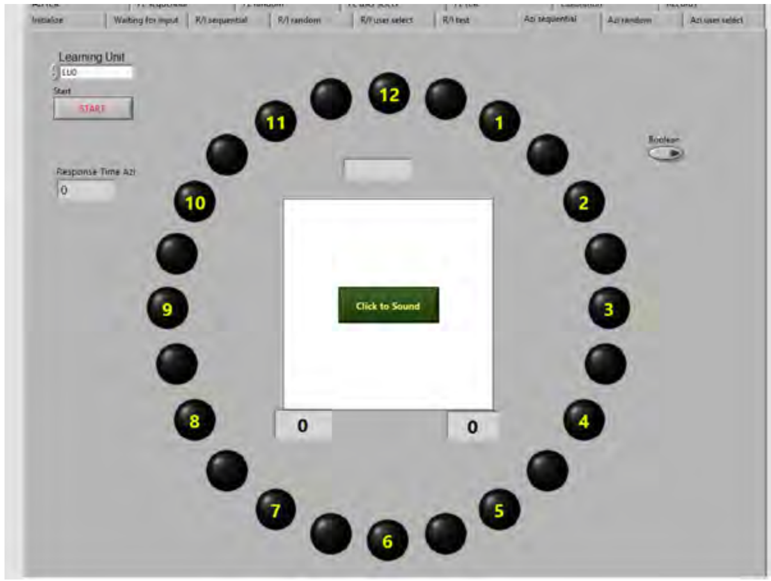


Figure 174. PALAT system sequential training screen on computer tablet displaying initiation training trial.

Following the familiarization, the participant was asked if they had any questions about the task or how to operate the computer tablet. The participant was not allowed to retake the familiarization but the investigator would answer any questions and re-demonstrate how to use the computer tablet if the participant was confused. Once the participant was ready, the investigator would select the “H test” button from main menu to navigate to the PALAT system testing screen (Figure 175). The testing screen was built to look very similar to the training screen to reduce operating errors. The main difference between the training and testing screens was the removal of the white box in the middle of the training screen that provided feedback to the participant. The investigator then selected Learning Unit 0 (LU0), or pretest, from the dropdown menu and clicked on the “Start” button to initialize the system (Figure 175). The participant was then handed the stylus pen and instructed to wait for the investigator to get setup outside of the loudspeaker array. The participant was reminded that the pretest consisted of 36 random presentations, or trials, with each of the 12 numbered loudspeakers playing three times during the test. The participant was instructed to select on the touchscreen where they perceived the dissonant tone originated by clicking on one of the 24 black circles representing the 12 loudspeaker locations and 12 positions between each loudspeaker. The participant was instructed to face forward and look at the white sign with the black number “12” prior to clicking on the “Click to START” button. The participant was informed that they were free to move their head and rotate their body at the onset of the dissonant signal to aid in localization. The participant was then instructed to “respond as accurately and as quickly as possible.”

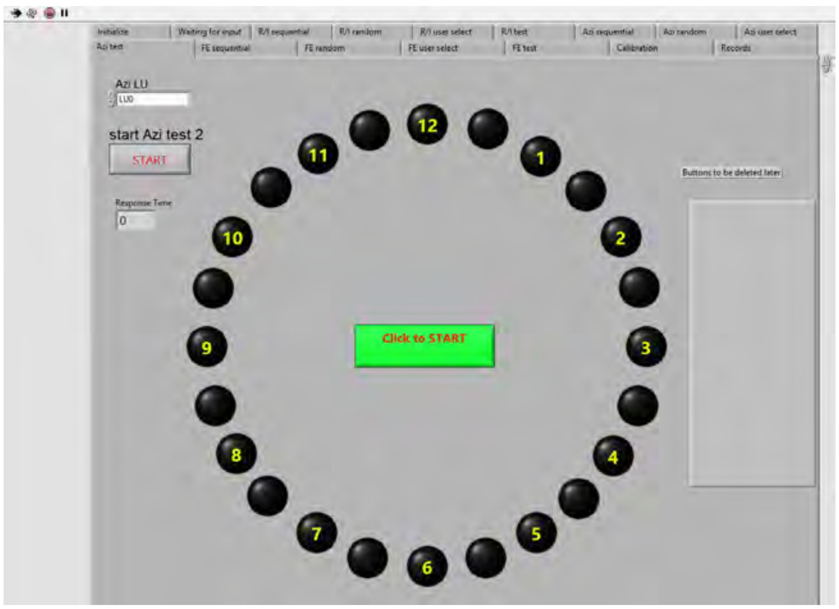


Figure 175. PALAT system test screen on computer tablet displaying initiation and response options.

Once in position, the investigator instructed the participant to begin the pretest whenever they were ready by clicking on the “Click to START” button located in the center of the testing screen and to respond by touching one of the 24 response buttons represented by the black circles on the sequential training screen. Each time the participant selected the “Click to START” button, a dissonant signal was emitted from one of the 12 loudspeakers located at the 12 clock face positions while triggering the response timer onset. Upon selecting one of the 24 response buttons, the response timer stopped and a new row of data was stored in the CSV file. At the completion of the 36 presentations and participant responses, the PALAT system informed the participant that the test was completed by a pop-up window stating, “This completes the test.” The system returned to the main screen after the participant clicked the “Ok” button.

Following the pretest under a TCAPS listening condition, the investigator removed and turned off the TCAPS device. The investigator then accessed the pretest score from the Dropbox file and informed the participant of their absolute and ballpark score. The participant was then asked to complete a questionnaire on the computer tablet. The investigator was available to answer any questions concerning the questionnaire. After completing the questionnaire, the investigator changed the PALAT system listening condition in the computer tablet and prepared the participant for the next pretest. The familiarization unit and pretest were repeated for the remaining two listening conditions. Following the completion of all three pretests, the investigator confirmed to the participant the schedule for the next training session or field posttest.

4.5.3 Training Session (Experimental group only)

The auditory localization training employed in this study was originally designed by Lee and Casali (2017) and modified during Phase I of the overarching experiment. Participants randomly assigned to the experimental group conducted three, one-hour localization training sessions consisting of five learning units (LUs). The participant underwent training with one listening condition per session. Each LU consisted of the following subunits:

1) Sequential- For LU 1, the dissonant signal played in sequential order around the 12-loudspeaker array for four “laps,” for a total of 48 presentations. The progression of the sequential presentations was as follows:

a) starting at 12 o’clock and moving clockwise through 11 o’clock,

b) starting at 9 o’clock and moving counterclockwise through 10 o’clock

c) starting at 3 o’clock and moving clockwise through 2 o’clock, and

d) starting at 6 o’clock and moving counterclockwise through 5 o’clock

For LU 2-5, the system delivered only one “lap” around the clock face, totaling 12 presentations, with a randomly-assigned starting position and direction of progression. As in the familiarization, the participant had prior knowledge of the signal location via computer tablet display. Figure 176 displays the feedback for an absolute correct response. Figure 177 displays the feedback for an incorrect response.

2) Random- The participant did not have prior knowledge of which loudspeaker would present the signal, but knew immediately if he/she answered correctly. The signal was played from each signal location three times in random order, for a total of 36 presentations.

3) User-select- The participant had 18 trials to choose loudspeaker locations from which they wished to hear additional presentations (Figure 178).

4) Test- A signal played from each of the 12 loudspeaker locations three times in random order, for a total of 36 presentations. The participant did not have prior knowledge of the sound location nor did they receive immediate feedback regarding their response.

At the end of each session, the participant completed the same questionnaire administered after the pretest.

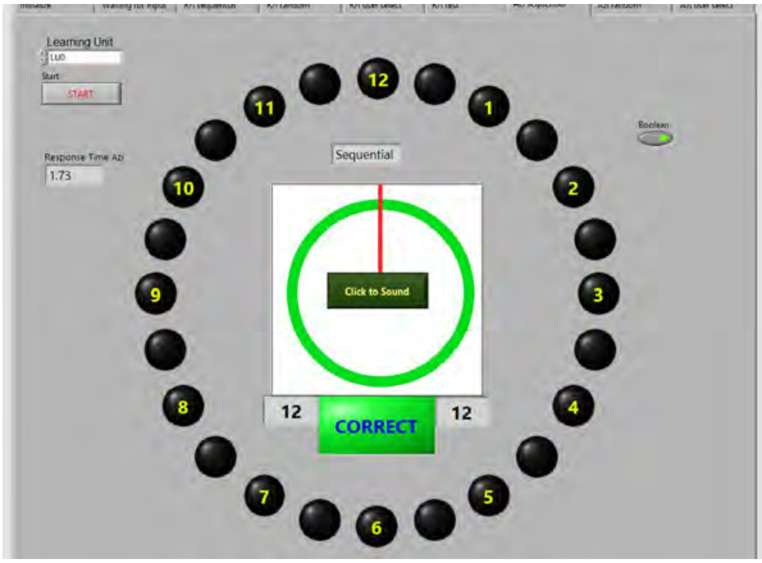


Figure 176. Participant interface on the computer tablet displaying 24 response location options and system-generated feedback for an absolute correct response.

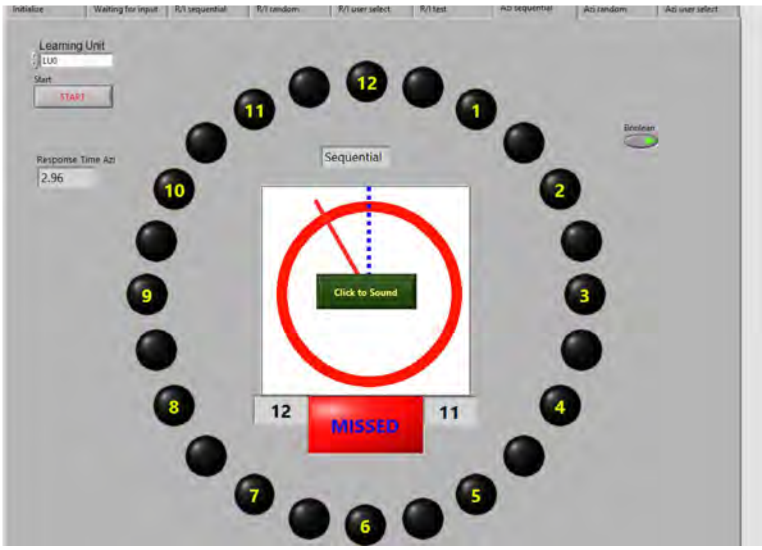


Figure 177. Participant interface on the computer tablet displaying feedback provided for an incorrect response.

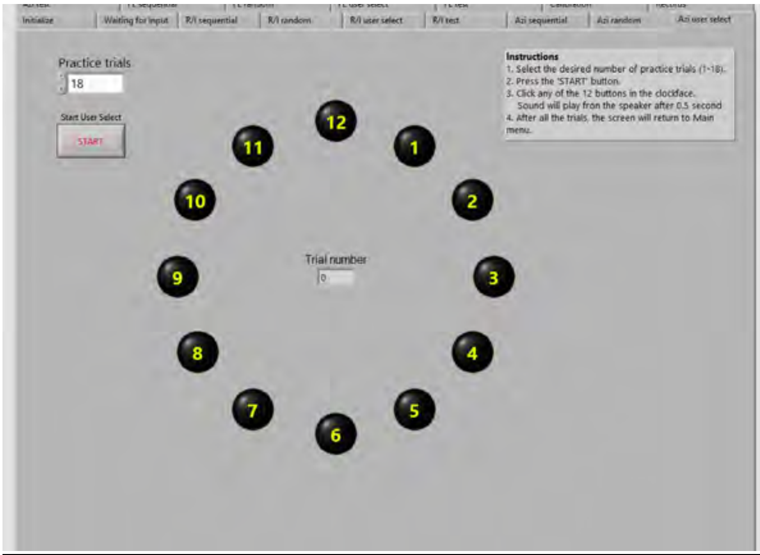


Figure 178. Example of the user-select display interface on the computer tablet displaying loudspeaker locations used for practice.

4.5.4 In-field posttest

The participant met the investigator at Whittemore Hall and were driven about 45 minutes to the field site. Upon arrival at the field site, the participant was offered bottled water, insect repellant, and sunscreen. The participant was then escorted on foot approximately 150 feet uphill to the 9 o'clock firing position. The investigator showed the participant the blank rounds and component parts of the remote firing device, for full disclosure purposes. The participant then walked approximately 200 feet up the hill to the center of the remote-firing device array. A sign with the number "12" was placed 50 feet from the center point to orient the participant and provide a visual reference point. The participant stood facing the 12 o'clock target, but was allowed to move his/her head to aid in localization. An investigator oriented the participant to the field site layout, the posttest procedure, and the response procedures. The participant the signal and registered their response using the same computer tablet as used during the pretest and localization training. After responding on the computer tablet, the participant was instructed to verbalize their response location so that the investigator could record the response as a backup data source. Following instructions for TCAPS posttests, the investigator fitted the participant with the TCAPS device. The two investigators were seated behind the participant between the 6 o'clock and 7 o'clock positions (Figure 179). One investigator observed the participant and recorded the response locations. The other investigator operated the control box used to route the electrical signal to the firing mechanism containing the blank .22 caliber rounds.

Figure 179. Field posttest site layout with participant standing in center of remote firing devices facing 12 o'clock target position and two investigators seated behind participant between the 6 o'clock and 7 o'clock positions.

Prior to each posttest, the participant received a familiarization unit similar to the pretest. The participant was instructed prior to starting the familiarization unit that a blank gunshot would be initiated from 12 o'clock, 3 o'clock, 6 o'clock, and lastly 9 o'clock. The participant was instructed to respond to the signal as accurately and quickly as possible. As in the training and pretest, the participant had the option of selecting one of 24 possible blank gunshot locations, 12 active gunshot locations and 12 inactive. Following the familiarization unit, the participant was administered a posttest. As in the pretest and LU-generated tests, the participant heard three presentations, or gunshots, from each of the 12 locations in a randomized order for a total of 36 presentations. The participant conducted one posttest under each of the three listening conditions in the same counterbalanced order as their pretest.

After each listening condition and at the conclusion of the posttest, the participant completed the same questionnaire used during the pretest and localization training. The trained group had one additional question that queried their perceived degree of preparedness as a result of lab-conducted training.

The wind speed was measured and recorded at the start of every posttest. Mean, minimum and maximum wind speeds for all posttests are shown in Table 108. No posttests were conducted if wind gusts measured above 8 miles per hour (mph). In addition, the posttests were suspended in inclement weather more than a very light rain mist. The temperature and humidity were measured and recorded at the start of each posttest. During the posttest, the investigator monitored the wind speeds and weather conditions to ensure gusts did not exceed 8 mph. In order to mitigate weather delays during testing, the investigator monitored the weather forecast prior to scheduling the posttest. Testing ceased only for changes in wind versus those in temperature or humidity as wind caused masking effects. In other words, wind creates noise unrelated to the original sound. Whereas variation in humidity and temperature contributed to the external validity of the experiment without drastic change in the overall sound source.

Table 108. Weather conditions during the field test.

	Mean	Minimum	Maximum
Wind Speed	0.7 mph	0 mph	2 mph
Temperature	69°	50°	81°
Humidity	62%	38%	87%

4.6 Phase III Results

Data reduction and calculations were performed using Microsoft® Excel. Statistical analysis was performed with JMP® 14 software, IBM® SPSS® Statistics, and Excel v16.16.10.

4.6.1 Outlier Analysis

After running 24 participants through the full experiment, a Dixon *Q*-test was performed on all dependent measures to identify outliers. The outlier analysis was performed separately each of the three quantitative dependent measures for each listening condition for pretest and posttest. The resulting sample size for each Dixon *Q*-test was *n*=24, 12 untrained and 12 trained participants. To perform the Dixon *Q*-test, the subset of data for each test was arranged sequentially from lowest to highest value. A Dixon *Q*-test was then performed manually using one of the following formulae:

(3) $$Q = \frac{|x_n - x_{n-1}|}{|x_n - x_1|} \quad \text{or} \quad Q = \frac{|x_2 - x_1|}{|x_n - x_1|}$$

where *n* is the sample size and the *x* represents the ordered values from lowest to highest, $x_1 < x_2 < \ldots < x_n$ (Dixon, 1951). The numerator in equation (3) represents the gap between the two highest values, $|x_n\text{-}x_{n-1}|$, or the gap between the two lowest values, $|x_2\text{-}x_1|$. The denominator in equation (3) represents the range of the data, $|x_n\text{-}x_1|$. The equation that resulted in the largest gap was used to identify the existence of a single outlier for each data subset. The calculated *Q*-value for each data subset was compared to Dixon's r_{10} table for *n*=24 using a 95% confidence interval (Dixon, 1951). If $Q \geq 0.34$, then an outlier was deemed present.

Two significant outlier data points were found using the Dixon *Q*-test. A significant outlier was present for the absolute correct score on the pretest under the TEP-100 listening condition (*Q*=0.37). The outlier data point of an absolute correct score=36 was associated with participant 23 from the untrained group. Figure 180 displays the absolute scores on the pretest under the TEP-100 listening condition by group.

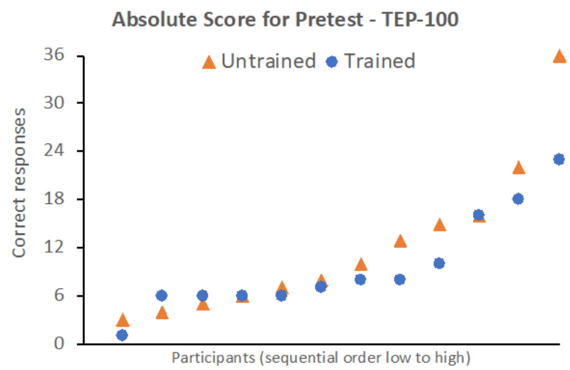


Figure 180. Absolute correct score on pretest under TEP-100 listening condition for all participants. Values ordered and plotted from lowest to highest score by group.

A significant outlier was present for response time on the posttest under the open ear listening condition (*Q*=0.37). The outlier data point of a response time=4.86 seconds was associated with participant 23 from the untrained group. Figure 181 displays the response times on the posttest under the open ear listening condition by group.

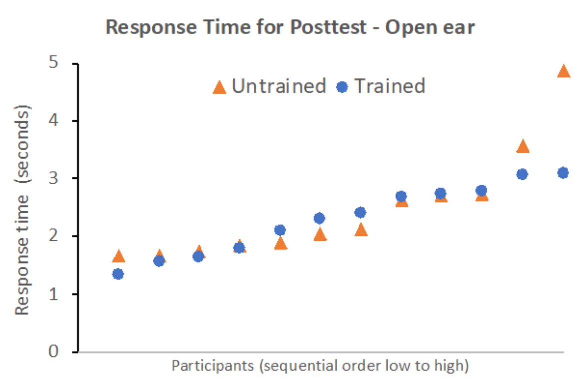


Figure 181. Response time on posttest under open ear listening condition for all participants. Values ordered and plotted from lowest to highest response time by group.

Both outlier data points were associated with participant 23 who was assigned to the untrained group. In both instances, participant 23's absolute correct score and response time would have biased the mean data in favor of *better* performance for the trained group. Additionally, the outlier would have increased the pretest absolute correct score for the untrained group, and thus, exaggerating the mean difference between pretest and posttest. Inclusion of participant 23's response time would have resulted in an exaggerated increase in response time from the pretest to the posttest. As a result, the investigator decided to replace outlier participant 23 with a new participant. This replacement was assigned to the same group (untrained) and counterbalancing scheme for listening condition (open ear, TEP-100, and ComTac™ III) as the outlier. All data points and analyses were then performed using the new participant 23's results. With the new data set inclusive of the outlier replacement, another Dixon *Q*-test was performed on all dependent measure data and no significant outliers were detected.

4.6.2 Objective performance

Analysis technique overview

Auditory localization performance on a set of measures (absolute correct, ballpark correct, and response time) according to training condition (untrained versus trained), listening condition (open ear, TEP-100, and ComTac™ III), and stage of training (pretest versus posttest) was first examined using a multivariate analysis of variance (MANOVA). The MANOVA evaluated the presence of significant mean differences according to experimental manipulation on a composite set of dependent measures. The test statistic employed in the MANOVA was Wilk's Lambda (λ), which represents the percentage of variance unexplained by the manipulation of the independent variables (Pituch & Stevens, 2016). Generally, the analysis stayed within the significance level for the mixed factor MANOVA using α=0.05. However, given the lack of experimental control in the field study, the α level was set to a less stringent α=0.10 for follow-up univariate tests in order to further analyze results.

Another statistic which was applied was partial eta squared (η^2_p), which represents the ratio of total variance in the sample that can be explained by the factor (Pituch & Stevens, 2016). Some of the effect sizes exceeded an η^2_p criterion for medium or large effect sizes of .06 and .14, respectively (Cohen, 1988), implying a meaningful effect. In regards to power, given that complement of the power value is the likelihood of making a Type II error, a power value of 0.8 is generally considered acceptable (Cohen, 2013). However, given the nature of the field study, 0.7 was considered as an acceptable power value.

The degree of auditory localization needed to perform ground combat-related tasks has yet to be validated or quantified. As such, the ballpark correct localization criterion (i.e., a response within ±15° of the location of the actual presented signal) was included in the analysis to describe performance using a less stringent standard should this requirement become known. The intent of applying different criteria to results of this study was to enable interpretation and relevancy according to differing standards. However, the investigator maintains that the 30° (±15°) accuracy criterion enables vision to better overcome localization blur. Azimuthal separation factor has implications for point of sound origin separation distance at effective range of military relevant signals. Auditory localization orients the listener to the direction of the sound and cues the visual modality, effectively reducing the response time in target identification (Wickens et al., 2013). A wider azimuthal separation angle at the listener translates to a broader field of view search as the distance increases from the listener. Larger separation angles become problematic in ground combat scenarios where military threats originate from long distances. Table 109 shows a comparison of the resulting visual field search distances associated with 30°

and 45° azimuthal separation for military weapons originating from their effective range distances.

Table 109. Visual field of view search distances associated with 30° and 45° azimuthal separation for military threats originating from their effective range distances (USMC, 2017)

Military Threat	Effective Range (m)	Field of view search distance at effective range (m)		
		30°	45°	Difference (Δ)
AK-47 gunshot	300	155	229	74
Rocket Propelled Grenade (RPG)	500	258	383	125
AK-74 Sniper rifle	800	414	612	198
PKM machine gun	1000	517	756	239
82mm mortar launch	3000	1553	2296	743
107mm rocket launch	> 5000	2588	3826	1238

The analysis included a within-subjects factor with more than two levels. Therefore, Mauchly's Test of Sphericity evaluated the assumption of homogeneity of variance between related of group comparisons (Portney & Watkins, 2009). This test is not applied to the between-subjects variable or a variable with only two levels, such as stage of training in this study. *F*-tests on a source of variance that is associated with violations of the homogeneity of variance assumption can underestimate the likelihood of making a Type I error (Portney & Watkins, 2009). Mauchly's Test of Sphericity evaluates the need for adjusting the *p* value when violations are detected (Portney & Watkins, 2009). A significant *p*-value in the Mauchly's Test indicates a violation occurred and an adjustment to reduce the degrees of freedom is made (Portney & Watkins, 2009). As a result, the critical value needed to achieve a significant finding is greater, correcting for the greater likelihood of making a Type I error (Portney & Watkins, 2009). Two estimators, the Greenhouse-Geisser and Huynh-Feldt, provide measures of Epsilon (ϵ) that describes the deviation of the covariance matrix from sphericity, and both were applied with the results in Table 110. A value of 1 indicates no deviation, and thus adherence to the assumption of sphericity (Pituch & Stevens, 2016). The Greenhouse-Geisser and Huynh-Feldt can underestimate and overestimate ϵ, respectively (Pituch & Stevens, 2016). As such the Greenhouse-Geisser is usually the first estimate applied if the correction results in a significant finding (Portney & Watkins, 2009). If results are not significant, the Huynh-Feldt is typically applied (Portney & Watkins, 2009). Mauchly's Test of Sphericity (Table 110) did not result in any significant values; therefore, the assumption of sphericity was met for the MANOVA and no corrections were required.

Following the MANOVA, univariate ANOVAs were conducted. Main effects were followed up with pairwise comparisons using a Bonferroni adjustment. Simple-effects *F*-tests followed up any interactions in order to partition the data according to the sources of most interest. Given that each factor analyzed in the simple-effects *F*-tests procedure only had two levels, follow-up pairwise comparisons were not conducted.

Finally, in graphing the data, for most graphs that follow, arithmetic mean values are plotted in bar graph form, with 95% confidence limits shown about the mean, and means with different letters indicative of statistical significance. In addition, in all tables, statistically-significant effects are denoted by boldface font.

Results: Evaluation of transfer-of-training effects from the in-lab to in-field localization performance

The mixed-factors MANOVA (Table 110) for the effect of training group, listening condition, and stage of training on the set of dependent measures showed a significant effect for listening condition, Wilk's λ (6, 17=0.04, F=1.79, $p <$ 0.000). Follow-up univariate ANOVAs were conducted for each dependent measure (Table 111). For the between-subjects variable of group, a significant difference was found using the measure of ballpark score, $F(1, 22)$=3.05, p=0.095 , η_p^2=0.12. For the within-subjects variable of listening condition, significant differences existed using the measures of absolute correct score, $F(2, 44)=120.44$, $p<0.000$, $\eta_p^2=0.85$, ballpark correct score, $F(2,44)=143.94$, $p<0.000$, $\eta_p^2=0.87$, and response time $F(2, 44)=11.35$, $p<0.000$, $\eta_p^2=0.34$.

Table 110. Mauchly's Test of Sphericity and MANOVA results evaluating the effects of training group, listening condition, and stage of training on absolute score, ballpark score, and response time (bolded text indicates a significant test result at the α=0.05 significance level).

Mauchly's Test of Sphericity on Within-Subjects Variables						Epsilon (ϵ)	
Variables	Measure	Mauchly's Criterion	Chi-Square	df	p	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	Absolute	0.91	1.92	2	0.384	0.92	1
	Ballpark	0.77	5.47	2	0.065	0.81	0.91
	RT	0.97	0.58	2	0.750	0.97	1
Listening Condition x Stage of Training	Absolute	0.99	0.25	2	0.882	0.99	1
	Ballpark	0.99	0.17	2	0.920	0.99	1
	RT	0.92	1.74	2	0.419	0.93	1

Source	Wilk's λ	df	F value	p	η_p^2
Between-Subjects					
Training Group (Trained; Untrained)	0.88	(3, 20)	0.92	0.447	0.12
Within-Subjects					
Listening Condition (Open; TEP-100; ComTac™ III)	**0.04**	**(6, 17)**	**1.79**	**< 0.000**	**0.96**
Stage of Training (Pretest; Posttest)	0.73	(3, 20)	2.51	0.088	0.27
Stage of Training x Training Group	0.83	(3, 20)	1.41	0.270	0.17
Listening condition x Training Group	0.77	(6, 17)	0.83	0.566	0.23
Stage of Training x Listening condition	0.82	(6, 17)	0.61	0.720	0.12
Listening condition x Stage of Training x Training Group	0.61	(6, 17)	1.79	0.161	0.39

Table 111. Univariate ANOVA results for each dependent measure for each independent variable (bold text indicates significant values at the α=0.10 significance level).

Source		df	Mean Square	*F value*	*p*	η_p^2
Between Subjects						
Group (G)						
	Absolute	1	100.00	1.39	0.252	0.06
	Ballpark	**1**	**142.01**	**3.05**	**0.095**	**0.12**
	Response Time	1	0.60	0.26	0.618	0.01
Error (S/G)						
	Absolute	22	72.10			
	Ballpark	22	46.62			
	Response Time	22	2.34			
Within Subjects						
Listening Condition (C)						
	Absolute	**2**	**3295.26**	**120.44**	**<0.000**	**0.85**
	Ballpark	**2**	**3516.90**	**143.94**	**<0.000**	**0.87**
	Response Time	**2**	**3.19**	**11.35**	**<0.000**	**0.34**
C x G						
	Absolute	2	10.65	0.39	0.680	0.02
	Ballpark	2	0.76	0.03	0.970	0.00
	Response Time	2	0.33	1.19	0.310	0.05
Error (*C x S/G*)						
	Absolute	44	27.36			
	Ballpark	44	24.43			
	Response Time	44	0.28			
Stage of training (T)						
	Absolute	1	0.03	0.00	0.975	0.00
	Ballpark	1	15.34	0.77	0.390	0.03
	Response Time	**1**	**1.96**	**6.10**	**0.022**	**0.22**
T x G						
	Absolute	**1**	**110.25**	**4.05**	**0.057**	**0.16**
	Ballpark	1	47.84	2.39	0.136	0.10
	Response Time	1	0.30	0.94	0.344	0.04
Error (*T x S/G*)						
	Absolute	22	27.25			
	Ballpark	22	19.98			
	Response Time	22	0.32			
C x T						
	Absolute	2	28.38	1.55	0.223	0.07
	Ballpark	2	28.38	1.46	0.244	0.06
	Response Time	2	0.06	0.60	0.553	0.03
C x T x G						
	Absolute	**2**	**52.94**	**2.90**	**0.066**	**0.12**
	Ballpark	2	21.05	1.08	0.348	0.05
	Response Time	2	0.03	0.28	0.758	0.01
Error (C x T x S/G)						
	Absolute	44	18.27			
	Ballpark	44	19.47			
	Response Time	44	0.09			
Total		429	7634.77			

<u>Group Main Effect: Post hoc test for Ballpark Correct Score</u>

Pairwise comparisons for the effect of group was not conducted given that only two levels of the independent variable, trained and untrained, were used in this experiment. Instead, mean ballpark scores were examined and showed that the trained group (*M*=20.39, *SD*=8.96) scored higher than the untrained group (*M*=18.40, *SD*=8.30). The means for each group and 95% confidence intervals are plotted below in Figure 182.

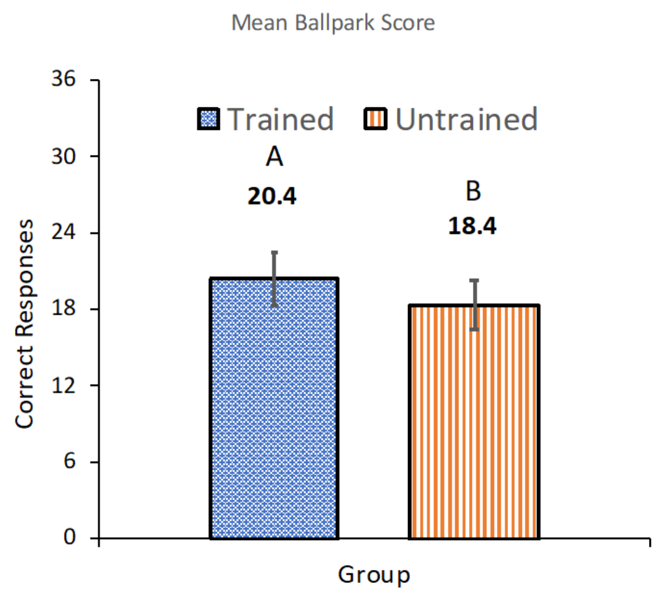


Figure 182. Mean ballpark correct scores with 95% confidence intervals plotted. Different letters indicate significant differences at $p \leq 0.10$.

<u>Listening Condition Main Effect: Post hoc test for Absolute Correct Score</u>

Pairwise comparisons were conducted for each listening condition (within the main effect of listening condition) using the measure of absolute correct score Table 112. All pairwise comparisons used a Bonferroni adjustment, which results in α=0.167 given that three comparisons were made (α=0.05/3). The mean absolute correct score for the open ear condition

(*M*=26) differed significantly from mean scores obtained in the TEP-100 (*M*=10.9) and ComTac™ III (*M*=12.6) conditions (Figure 183).

Table 112. Pairwise comparisons for listening condition using the absolute correct score.

Listening Condition		*M*	*SE*	*p*
Open ear	**TEP-100**	**15.15**	**1.01**	**<0.000**
Open ear	**ComTac™ III**	**13.40**	**0.91**	**<0.000**
ComTac™ III	TEP-100	-1.75	1.19	0.462

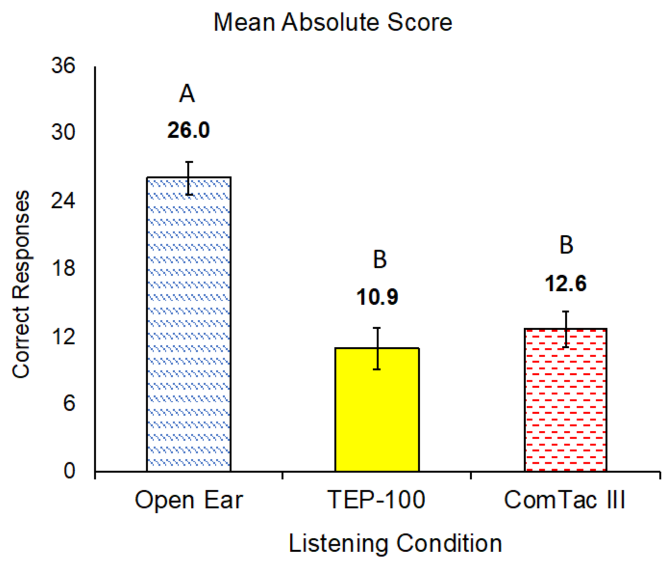


Figure 183. Mean absolute correct scores for each listening condition with 95% confidence intervals plotted. Different letters indicate significant differences at $p \leq 0.10$.

Listening Condition Main Effect: Post hoc test for Ballpark Correct Score

Pairwise comparisons conducted using the ballpark correct measure, shown in Table 113, also resulted in significant differences between the open (*M*=29.2) and TEP-100 (*M*=13.3) conditions and between the open and ComTac™ III (*M*=15.7) conditions. Mean ballpark scores within the main effect of listening condition are shown in Figure 184.

Table 113. Listening condition significant pairwise comparisons using the **ballpark** correct score.

Listening Condition		*M*	*SE*	*p*
Open ear	**TEP-100**	**15.85**	**1.04**	**<0.000**
Open ear	**ComTac™ III**	**13.52**	**0.76**	**<0.000**
ComTac™ III	TEP-100	2.33	1.19	0.186

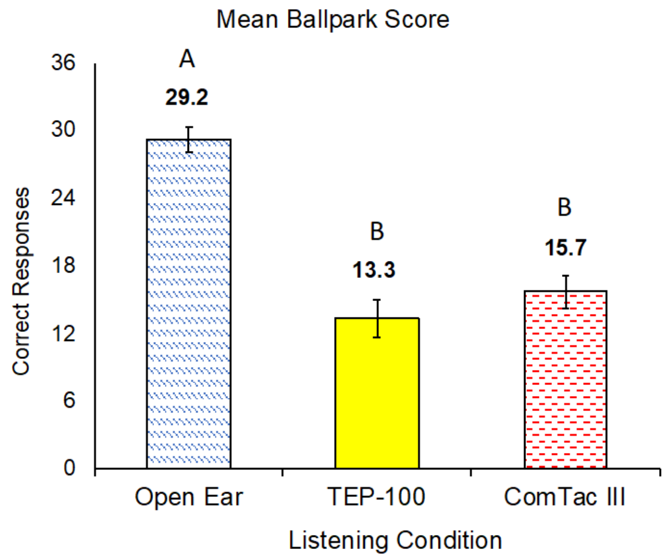


Figure 184. Mean ballpark correct scores for each listening condition with 95% confidence intervals plotted. Different letters indicate significant differences at $p \leq 0.10$.

Listening Condition Main Effect: Post hoc test for Response Time

Pairwise comparisons on the measure of response time in seconds, shown in Table 114, showed significant differences between the mean response time for open ear (*M*=2.2) and TEP-100 (*M*=2.6) and between the open ear and the ComTac™ III (*M*=2.6). Mean response times within the main effect of listening condition are shown in Figure 185.

Table 114. Listening condition significant pairwise comparisons using the **response time** score.

Listening Condition		*M*	*SE*	*p*
Open ear	**TEP-100**	**-0.48**	**0.12**	**<0.000**
Open ear	**ComTac™ III**	**-0.41**	**0.11**	**0.001**
ComTac™ III	TEP-100	-0.07	0.10	1.000

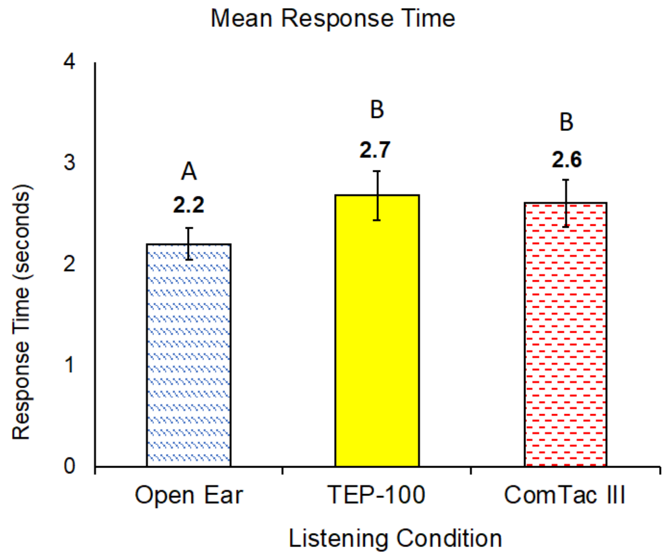


Figure 185. Mean response times for each listening condition with 95% confidence intervals plotted. Different letters indicate significant differences at $p \leq 0.10$.

Stage of Training Main Effect: Post hoc test for Response Time

The repeated-measures ANOVA conducted for the stage of training main effect on the measure of response time was significant, $F(1,22)=6.10$, $p=0.022$, $\eta_p^2=0.2$. The mean response time at pretest ($M=2.4$) was significantly lower than the posttest ($M=2.6$) (Figure 186). Given only two levels of the independent variable, pairwise comparisons were not conducted.

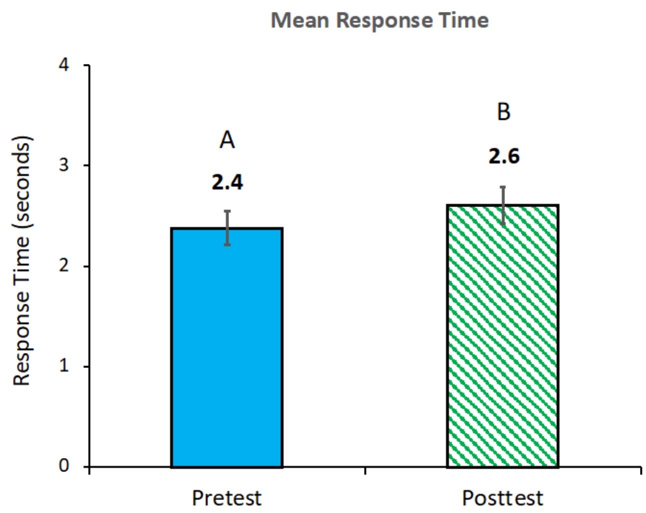


Figure 186. Mean response times from pretest to posttest. Different letters indicate significant differences at $p \leq 0.10$.

<u>Stage of Training x Group Interaction: Post hoc test for Absolute Correct Score</u>

Simple-effects *F*-tests further analyzed the significant interaction, at $\alpha=0.10$, for stage of training (pretest versus posttest) by group (trained versus untrained) using the absolute correct localization measure. Specifically, to determine if the groups significantly differed at pretest and posttest, two between-subjects ANOVAs were conducted at each training stage. To evaluate the assumption of homogeneity of variances, Levene's tests were conducted. The Levene's test calculates the deviation scores of the participants within each group from the group mean and then coverts the scores to absolute values (Pituch & Stevens, 2016). An ANOVA is then conducted comparing the mean absolute deviation scores between groups (Pituch & Stevens, 2016). A result of $p<0.05$ for the Levene's test indicates a violation to the assumption of homogeneity of the variance. In this analysis, the Levene's tests supported equality of the variances for the ANOVA conducted at the pretest stage examining group differences, $F(1,70)=1.06$, $p=0.297$, and at the posttest stage, $F(1,70)=1.45$, $p=0.233$; therefore, no corrections for heterogeneity were necessary. Simple-effects *F*-tests, shown in Table 115, ensued, yielding no significant differences between groups at pretest, $F(1,22)=0.00$, $p=0.946$, but exhibiting significance at posttest, $F(1,22)=7.17$, $p=0.011$. Mean absolute correct scores for each group at pretest and posttest are displayed in Figure 187. Results supported group equivalence at pretest collapsed across listening conditions, while showing significantly higher scores for the trained group over the untrained group at posttest.

Table 115. Simple-effect *F*-tests for trained group versus untrained group at each stage of training using the absolute correct score.

Source	*SS*	*df*	*MS*	*F*	*p*
G at Pretest	0.13	1	0.13	0.00	0.946
G at Posttest	**210.13**	**1**	**210.13**	**7.71**	**0.011**
Error (T x S/G)	599.40	22	27.24		
Total	809.65	24			

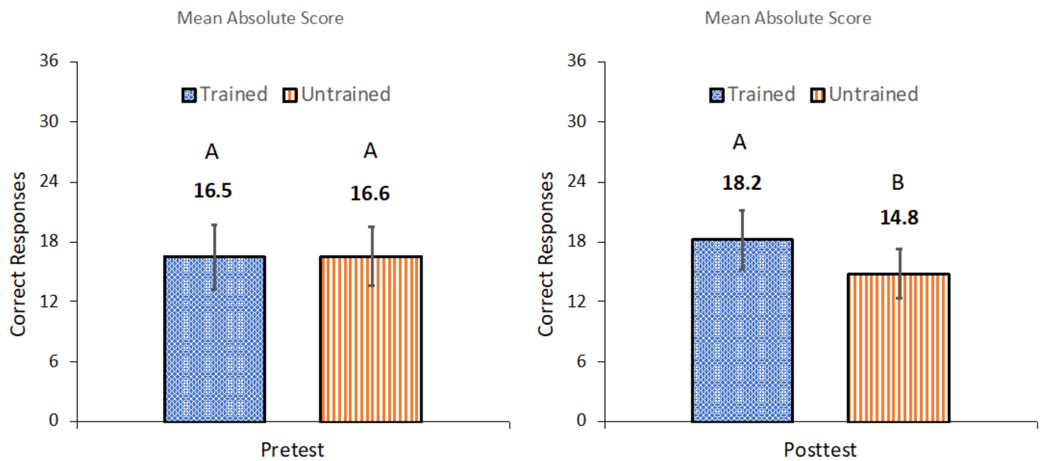


Figure 187. Mean absolute correct scores for each group at pretest and posttest. Different letters indicate significant differences at $p \leq 0.10$.

To further analyze the significant interaction of stage of training and group on the measure of absolute correct, simple-effects *F*-tests ANOVA were conducted. Separate ANOVAs were run for each group comparing performance between pretest and posttest. Analyzing only two levels of the repeated measures precluded sphericity testing. No significant differences were found between pretest and posttest absolute scores in the untrained group, *F*(1, 22)=2.09, *p*=0.163, or the trained group, *F*(1, 22)=1.96, *p*=0.175. ANOVA results are listed in Table 116 and means are displayed in Figure 188.

Table 116. Simple-effects *F*-tests for each group examining pretest versus posttest performance collapsed across listening conditions using the absolute correct score.

Source	*df*	*SS*	*MS*	*F*	*p*
Untrained	1	56.89	56.89	2.09	0.163
Trained	1	53.39	53.39	1.96	0.175
Error (T x S/G)	22	599.40	27.24		

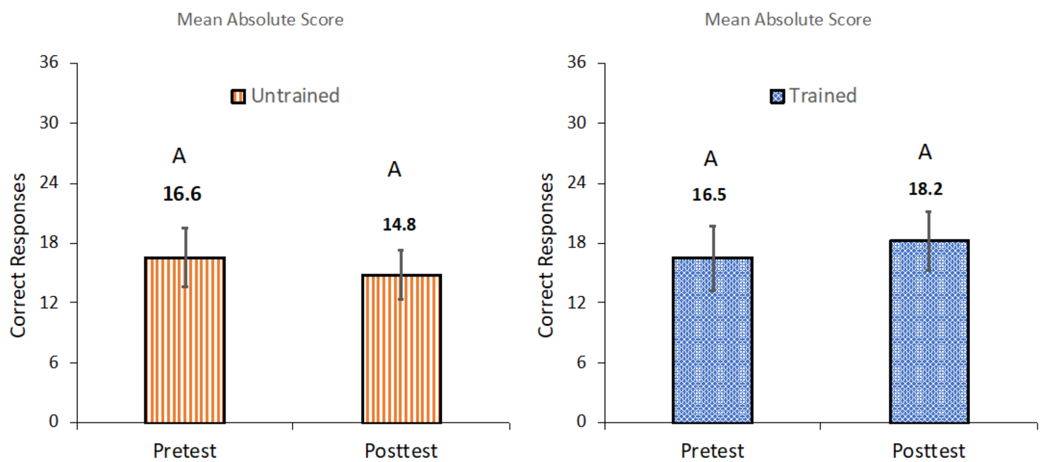


Figure 188. Mean absolute correct scores for each training group comparing pretest and posttest performance. Different letters indicate significant differences at $p \leq 0.10$.

__Listening Condition x Stage of Training x Group Interaction: Post hoc test for Absolute Correct Score__

To analyze the significant interaction of listening condition, stage of training, and group on the absolute correct measure, simple-effects *F*-tests were conducted. Separate ANOVAs conducted for each listening condition (open ear, TEP-100, and ComTac™ III) evaluated differences in scores for the trained versus untrained group at each stage of training. The Levene test supported equality of variances at pretest for the open ear, *F*(1,22)=0.13, *p*=0.726 and TEP-

100, $F(1,22)=0.09$, $p=0.771$, but not for the ComTac™ III $F(1,22)=6.10$, $p=0.022$. However, ANOVA F-tests tend to be robust to violations of variance equality given equal group sizes, as was the case in this study. Simple-effects results, listed in Table 117 and Figure 189, show no significant differences between trained and untrained participants at pretest for the open ear, $F(1, 44)=0.23$, $p=0.634$, the TEP-100, $F(1, 44)=0.33$, $p=0.569$, and the ComTac™ III, $F(1, 44)=0.82$, $p=0.370$. Thus, these results indicated group equivalence at pretest in each listening condition. The simple-effects test above was then repeated, comparing the trained versus untrained group for each listening condition, but for the posttest stage of training. Levene's statistic supported equality of variances for the open ear, $F(1,22)=0.003$, $p=0.959$, TEP-100, $F(1,22)=0.427$, $p=0.520$, and ComTac™ III, $F(1,22)=2.46$, $p=0.131$ (Table pp). The F-tests showed significant differences between the trained and untrained groups for the open ear, $F(1, 44)=13.18$, $p=0.001$, and the TEP-100 conditions, $F(1, 44)=3.83$, $p=0.057$. In all listening conditions, the mean absolute correct score was higher for the trained group when tested in the field, as shown in Table 118 and Figure 190.

Table 117. Simple-effect F-tests examining absolute correct score differences at **pretest** in trained versus untrained participants for each listening condition (open ear, TEP-100, and ComTac™ III)

Source	*df*	*SS*	*MS*	*F*	*p*
Open ear	1	4.17	4.17	0.23	0.634
TEP-100	1	6.00	6.00	0.33	0.569
ComTac™ III	1	15.04	15.04	0.82	0.370
Error (C x T x S/G)	44	817.08	18.27		
Total	47	842.29	43.48		

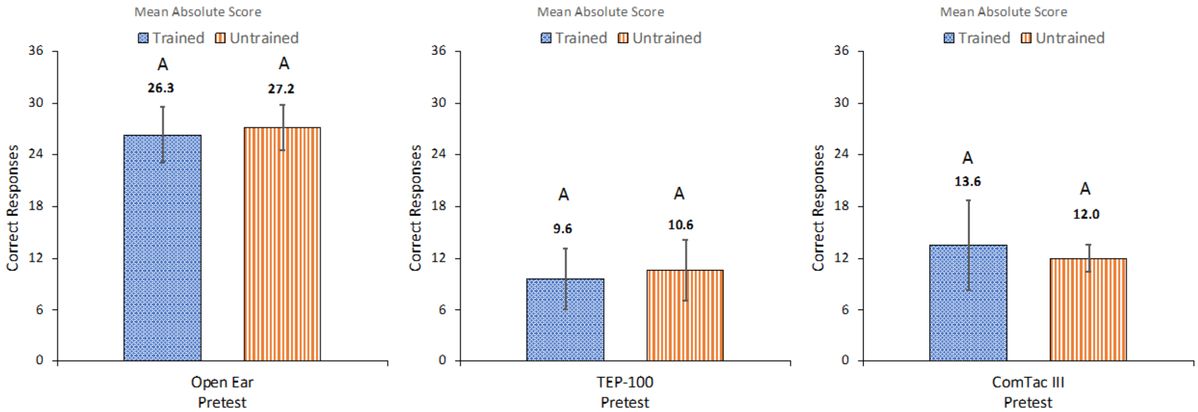


Figure 189. Mean absolute correct scores for the trained versus untrained groups at pretest for each listening condition. Different letters indicate significant differences at $p \leq 0.10$.

Table 118. Simple-effect *F*-tests examining absolute correct score differences at **posttest** in trained versus untrained participants for each listening condition (open ear, TEP-100, and ComTac™ III)

Source	*df*	*SS*	*MS*	*F*	*p*
Open ear	**1**	**240.67**	**240.67**	**13.18**	**0.001**
TEP-100	**1**	**70.04**	**70.04**	**3.83**	**0.057**
ComTac™ III	1	1.50	1.50	0.08	0.227
Error (Listening Condition x Stage of training/Group)	44	817.08	18.27		
Total	47	1129.29	330.48		

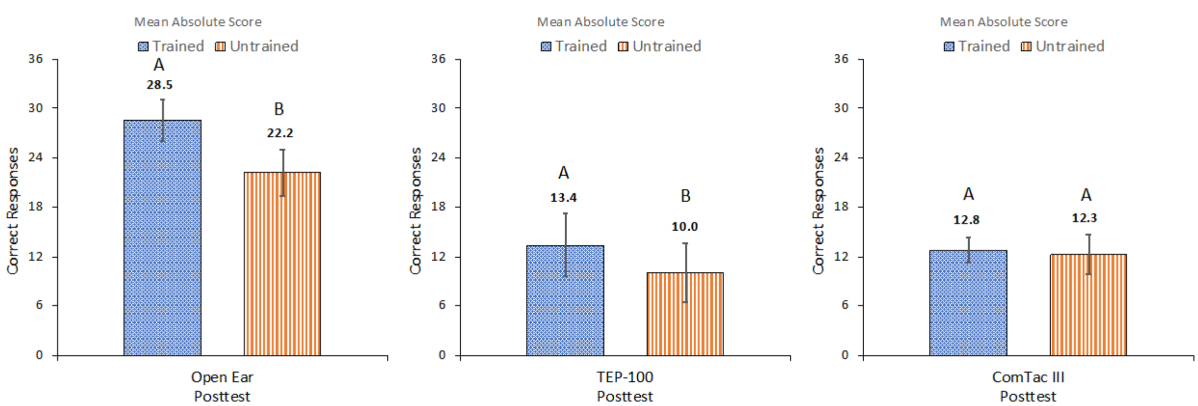


Figure 190. Mean absolute correct scores at posttest for each listening condition comparing the trained and untrained groups. Different letters indicate significant differences at $p \leq 0.10$.

Additional simple-effects testing was conducted on the 3-way interaction of listening condition, stage of training, and group. Simple-effects *F*-tests examined pretest versus posttest performance for each group and for each listening condition. The untrained group demonstrated significantly lower mean performance (Figure 191) in the open condition in the posttest versus pretest, $F(1, 44)=8.21$, $p=0.006$ (Table 119). Significant differences between pretest and posttest were not found for the TEP-100, $F(1, 44)=0.021$, $p=0.885$, or ComTac™ III conditions, $F(1, 44)=0.021$, $p=0.885$. For the trained group, the TEP-100 condition resulted in significantly higher scores in the posttest versus the pretest, $F(1, 44)=4.83$, $p=0.033$. No significant differences were found in the open ear, $F(1, 44)=1.54$, $p=0.221$, or ComTac™ III conditions, $F(1, 44)=0.23$, $p=0.634$. Results of the trained group simple-effects *F*-tests and means are provided in Table 120 and Figure 192, respectively.

Table 119. Simple-effect *F*-tests examining absolute correct score differences at pretest versus posttest for the **untrained** group. Different letters indicate significant differences at $p \leq 0.10$.

Source	*df*	*SS*	*MS*	*F*	*p*
Open ear	**1**	**150.00**	**150.00**	**8.21**	**0.006**
TEP-100	1	2.04	2.04	0.11	0.742
ComTac™ III	1	0.38	0.38	0.02	0.885
Error (C x T x S/G)	44	803.88	18.27		
Total	47	956.3	170.69		

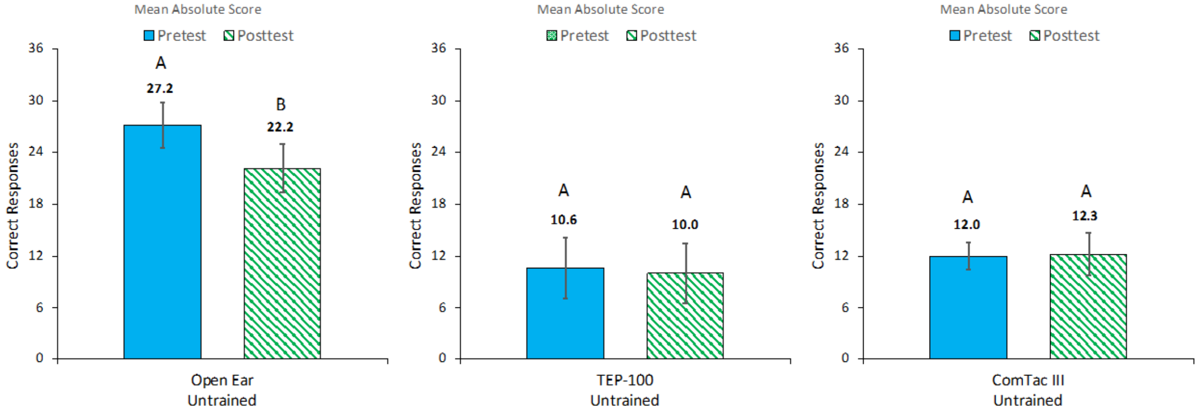


Figure 191. Mean absolute correct scores for each listening condition for the **untrained** group **comparing pretest and posttest** performance. Different letters indicate significant differences at $p \leq 0.10$.

Table 120.Simple-effect *F*-tests examining absolute correct score differences at pretest versus posttest for the trained group.

Source	*df*	*SS*	*MS*	*F*	*p*
Open ear	1	28.17	28.17	1.54	0.221
TEP-100	**1**	**88.16**	**88.16**	**4.83**	**0.033**
ComTac™ III	1	4.17	4.17	0.23	0.634
Error (C x T x S/G)	44	817.08	18.27		
Total	47	937.58	138.77		

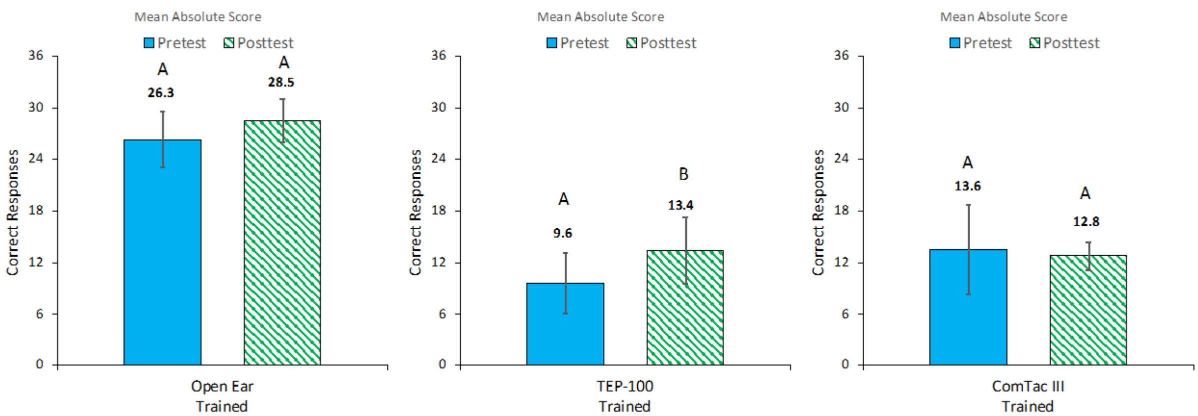


Figure 192. Mean absolute correct scores for each listening condition for the **trained** group **comparing pretest and posttest** performance. Different letters indicate significant differences at $p \leq 0.10$.

Front-back Reversal Errors

In addition to analyses using number of absolute and ballpark correct responses and response time measures, analyses were conducted using the dependent variable of the number of front-back reversal errors out of 36 trials. A front-back reversal occurred when the participant responded that a sound originating from 4 o'clock through 8 o'clock positions was located in the 10 o'clock through 2 o'clock positions, and vice-versa. As such, this type of error is known as a 120° arc front-back reversal. A mixed-factors ANOVA was conducted in order to examine the effect of group (between-subjects), listening condition (within-subjects), and stage of training (within-subjects) on the mean number of 120° arc front-back reversals. Mauchly's test of sphericity showed that the within-subjects factors met the assumption of equality of the variances (Table 121). The post hoc testing was not performed on stage of training given that only two levels were used in the analysis. Results from the mixed-factors ANOVA, using $\alpha=0.10$, showed only a main effect for listening condition was significant, $F(2, 44)=78.9$, $p<0.000$. All ANOVA results are provided in Table 122. Plotted means for each listening condition (Figure 193) showed that the highest number of front-back reversal errors occurred in the ComTac™ III condition ($M=8.4$), followed by the TEP-100 ($M=7.9$), and then the open-ear (1.1). Follow-up pairwise comparisons for the main effect for listening condition using a Bonferroni correction showed that the mean errors for open ear condition differed significantly from mean errors obtained in the TEP-100 and ComTac™ III conditions (Table 123).

Table 121. Mauchly's test of sphericity for mixed ANOVA for the effect of group, listening condition, and stage of training on **front-back errors** using the **120-degree arc criterion**.

Mauchly's Test of Sphericity					Epsilon (ϵ)	
Variables	Mauchly's Criterion	Chi-Square	df	*p*	Greenhouse-Geisser	Huynh-Feldt
Listening Condition	0.98	0.41	2	0.82	0.98	1
Listening Condition x Stage of training	0.81	4.40	2	0.11	0.84	1

Table 122. Mixed-factor ANOVA table evaluating differences in front-back reversal errors using the 120-degree arc criterion according to group, listening condition, and stage of training.

Source	df	Mean Square	*F value*	*p*	η_p^2
Between Subjects					
Group (G)	1	8.51	0.49	0.49	0.02
Error(S/G)	22	17.27			
Within Subjects					
Listening Condition (C)	**2**	**784.92**	**78.90**	**<0.000**	**0.78**
C x G	2	0.34	0.03	0.97	0.002
Error (C x S/G)	44	9.95			
Stage of Training (T)	1	31.17	2.65	0.12	0.11
T x G	1	2.01	0.17	0.68	0.008
Error (T x S/G)	22	11.76			
C x T	2	3.34	0.64	0.53	0.03
C x T x G	2	0.34	0.07	0.94	0.003
Error (C x T x S/G)	44	5.23			
Total	143	874.84			

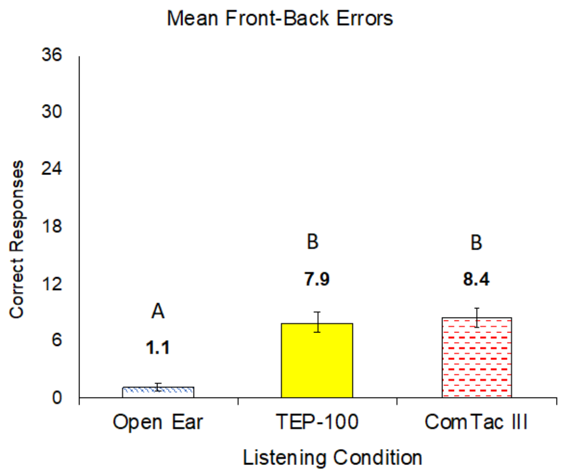


Figure 193. Mean front-back reversal errors using the 120-degree arc criterion for each listening condition. Different letters indicate significant differences at $p \leq 0.10$.

Table 123. Significant pairwise comparisons between listening conditions for front-back reversal errors using the 120° arc criterion with a Bonferroni adjustment

Listening Condition		*M*	*SE*	*p*
Open ear	**TEP-100**	**-6.73**	**0.67**	**<0.000**
Open ear	**ComTac™ III**	**-7.25**	**0.60**	**<0.000**
TEP-100	ComTac™ III	-0.52	0.66	1.000

Regression Analysis

In order to determine if performance on the portable system predicted in-field localization accuracy, especially in those who conducted training, regressions analysis was conducted. Specifically, the absolute score correct score was used in-office to predict in-field performance to assess the validity of using the in-office environment localization as a means to improve in-field performance. Therefore, post hoc regression was calculated to predict in-field performance based on in-lab results using the absolute correct score. In general, regression describes the magnitude of the relationship between the independent and dependent variables (Portney & Watkins, 2009). The regression line can be used to predict values of the dependent variable given a value of independent variable (Portney & Watkins, 2009). The null hypothesis for linear regression analyses is that the slope of the regression line is equal to zero (Portney & Watkins, 2009). In other words, a change in the independent variable results in no change in the dependent variable. As part of regression analyses, the *r*, or correlation coefficient is calculated. The *r* value reflects how closely the data matches the predicted values of the regression line, or goodness of fit (Portney & Watkins, 2009). Squaring the correlation coefficient, known as r^2, reflects the percentage of variance of the dependent variable accounted for by the independent variable. An α=0.10 value was used as the criterion for a significant linear regression in these analyses. A significant finding would indicate that given an absolute pretest score obtained in-office given a certain group membership (trained versus untrained) and listening condition (open ear, ComTac™ III, and TEP-100), the change in posttest score obtained in-field could be predicted.

Therefore, linear regression was conducted to examine the predictive value of pretest score for each combination of group and listening condition on posttest score. Regression analyses did not result in significant values for the following conditions: open ear condition for the trained group (F[1, 10]=2.51, p=0.144), open ear condition for the untrained group (F[1, 10]=0.41, p=0.536), TEP-100 condition for the untrained group (F[1, 10]=2.12, p=0.168), ComTac™ III condition for the trained group (F[1, 10]=2.12, p=0.176), and ComTac™ III condition for the untrained group (F[1, 10]=0.556, p=0.473).

A significant regression was found for the trained group using the TEP-100, F(1, 10)=3.71, p=0.083. Trained participants' posttest scores on the TEP-100 increased by 0.57 in the field for each correct answer on the pretest score (Figure 194). Table 124 shows the results of the ANOVA table. The resulting prediction equation is as follows:

Trained, TEP-100 Posttest Absolute Correct Score=7.06 + 0.57(Trained, TEP-100 Pretest Absolute Correct Score)

Table 124. ANOVA table for pretest prediction of posttest absolute correct scores in the trained group, TEP-100 condition.

Source	*SS*	*df*	*MS*	*F*	R^2	*p*
Model	**138.90**	**1**	**138.90**	**3.71**	**0.27**	**0.083**
Error	374.02	10	37.40			

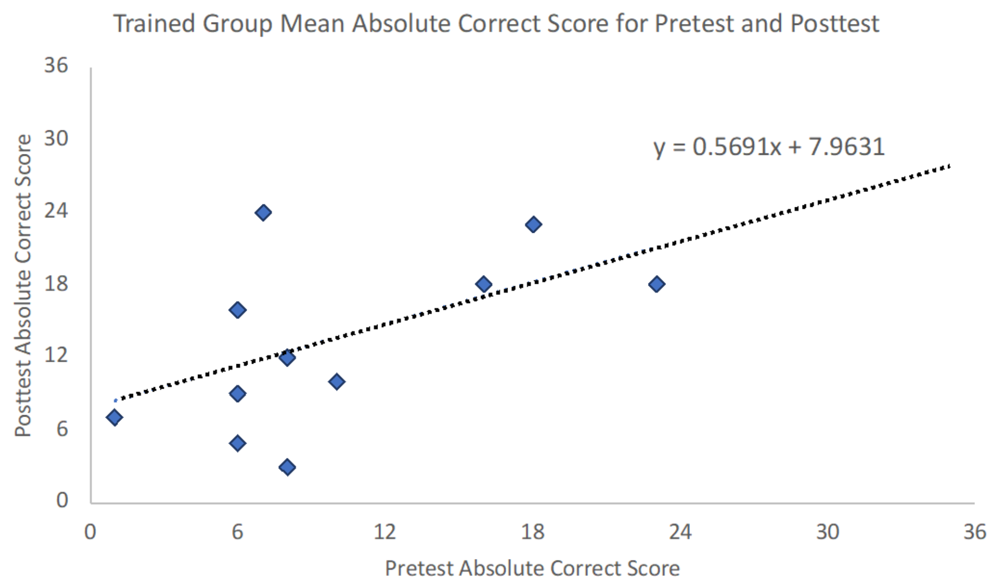


Figure 194. Mean absolute correct score on pretest and posttest for trained group. Regression line and equation plotted.

Independent-samples *t*-tests were then conducted to assess if the slope of the regression lines differed significantly between the trained and untrained groups from pretest to posttest for *each* listening condition, using the absolute correct scores. The α level was set to 0.10 and divided by the number of planned comparisons (3), resulting in a significance criterion value of 0.033. Levene's tests showed no violations to equality of the variance assumptions occurred for the open ear, $F(1, 22)=0.29$, $p=0.595$, TEP-100, $F(1, 22)=0.37$, $p=0.551$ and ComTac™ III, $F(1,22)=2.36$, $p=0.139$. *T*-test results, provided in Table 125, showed that significant group differences existed between the trained and untrained groups in the open ear condition, $t(22)=3.20$, $p=0.004$ (Figure 195), and the TEP-100 condition, $t(22)=2.49$, $p=0.021$ (Figure 196).

No significant difference existed between the trained and untrained group in the ComTac™ III condition. Examining the means for each group at for the open ear and TEP-100 conditions showed that training improved the participants' absolute correct scores from pretest to posttest, but performance declined in the posttest without training.

Table 125. Descriptive statistics and independent-samples *t*-tests, using the 0.033 corrected alpha level comparing group differences within each listening condition measured by the slope of the regression line from pretest to posttest for absolute correct score.

Variable	*n*	*M*	*SD*	*df*	*t*	*p*
Open ear				**22**	**3.20**	**0.004**
Trained	12	0.35	0.76			
Untrained	12	-0.83	1.03			
TEP-100						
Trained	12	1.00	1.07	**22**	**2.49**	**0.021**
Untrained	12	-0.10	1.10			
ComTac™ III						
Trained	12	0.38	1.36	22	0.75	0.139
Untrained	12	0.04	0.76			

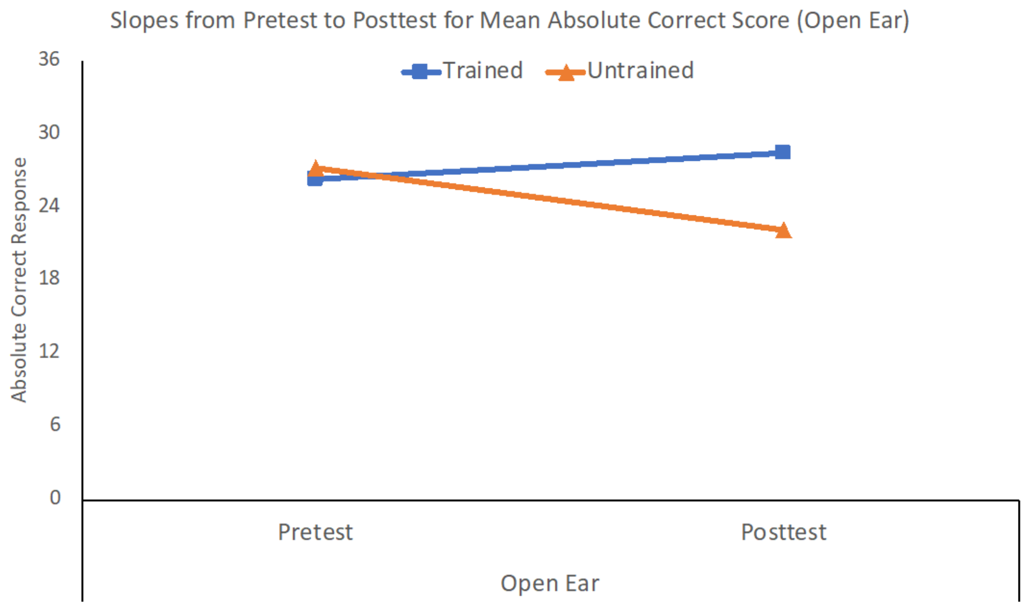


Figure 195. Slopes from pretest to posttest for mean absolute correct score for open ear by group.

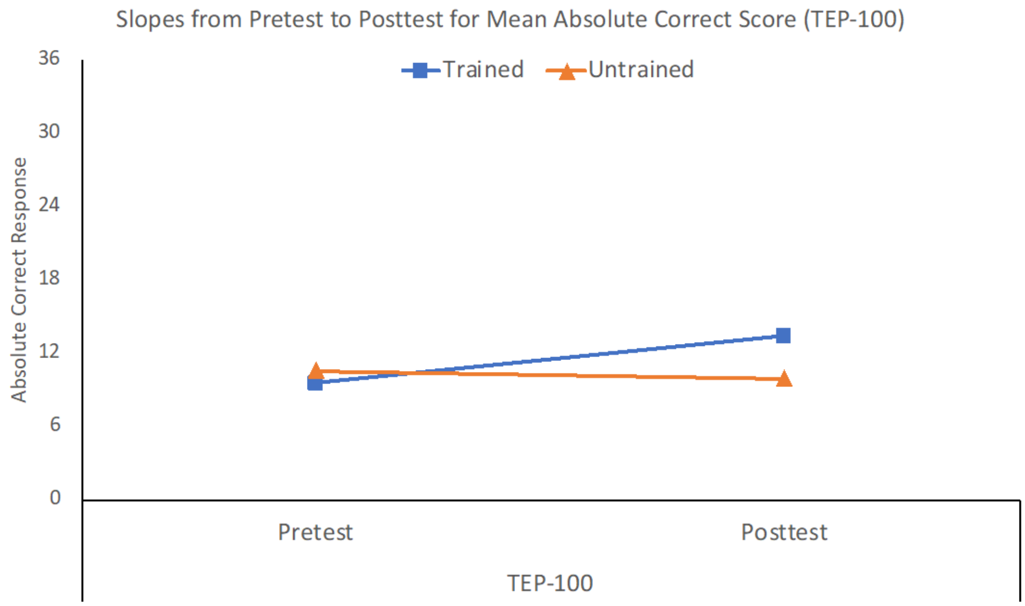


Figure 196. Slopes from pretest to posttest for mean absolute correct score for open ear by group.

4.6.3 Subjective ratings

Following the pretest and posttest, all participants answered a questionnaire which consisted of various bipolar rating scales. Participants in the trained condition also completed a questionnaire after the test in the LU5 subunit for each listening condition. The questionnaires are provided in Appendix K. To assess paired differences from pretest to posttest for each group and each training condition, Wilcoxon signed-rank tests were performed. This test procedure is the non-parametric equivalent to dependent samples *t*-tests. The Wilcoxon signed-ranks test assesses the direction and magnitude of differences of paired scores (Portney & Watkins, 2009). In this procedure, difference scores are ranked, disregarding the +/-sign and eliminating pairs with difference scores equal to zero (Portney & Watkins, 2009). Then, respective signs are assigned to the ranks (Portney & Watkins, 2009). If a participant's difference scores result in a tie, a mean rank is assigned (Portney & Watkins, 2009). Rejecting the null hypothesis for a Wilcoxon signed-ranks test means an unequal number of positive and negative ranks existed

(Portney & Watkins, 2009). Conversely, supporting the null hypothesis indicates that an equal number of positive and negative ranks existed (Portney & Watkins, 2009). In this study, a significant finding indicated ratings were significantly different at $\alpha=0.05$, given a condition, from pretest to posttest.

In order to assess group differences for each listening condition given a certain stage of training, Mann-Whitney *U* tests were performed. This test is the non-parametric counterpart to an independent samples *t*-test. The testing procedure involves ranking all of the scores, regardless of group membership, in ascending order (Portney & Watkins, 2009). The ranks for each group are then summed, with equal sums for groups supporting the null hypothesis (Portney & Watkins, 2009). Adequately large differences between the sums for each group results in rejecting the null hypothesis (Portney & Watkins, 2009). Given that non-parametric tests do not have a parallel procedure for mixed-factors ANOVAs, between-subjects and within-subjects testing was evaluated separately. An α level of 0.05 was adopted for all non-parametric tests.

Question 1. Perceived Confidence

Participants were asked to respond to the following: Rate how **confident you were** in your ability to locate sounds under this listening condition from 1 (no confidence) to 7 (extremely confident). Wilcoxon signed-ranks tests showed no significant differences from pretest compared to posttest for each group and listening condition combination (Table 126). Mean ratings for each listening condition, group, and stage of training are plotted in Figure 197.

Table 126. Wilcoxon results comparing ratings for pretest versus posttest for each listening condition and group for Question 1, Perceived Confidence.

Listening Condition	Group	Z	*p*
Open ear	Trained	-0.58	0.564
Open ear	Untrained	-0.71	0.480
TEP-100	Trained	-0.50	0.620
TEP-100	Untrained	-1.31	0.190
ComTac™ III	Trained	-0.92	0.357
ComTac™ III	Untrained	-0.14	0.887

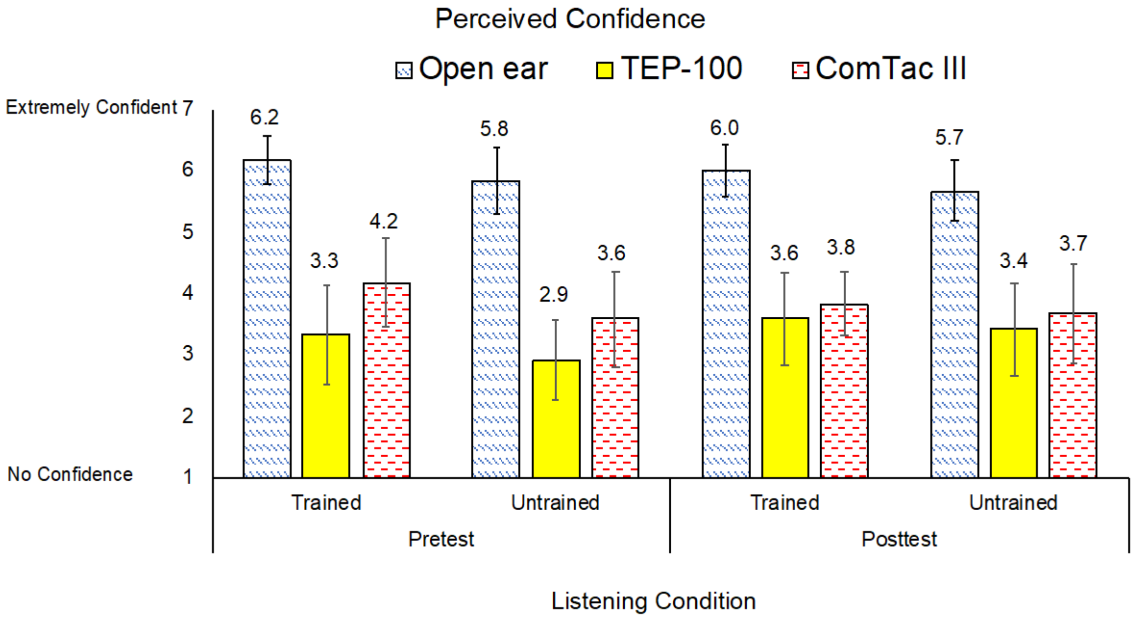


Figure 197. Plotted means and 95% confidence intervals for ratings at pretest and posttest for each group and listening condition for Question 1, Perceived Confidence.

To evaluate differences in ratings of confidence in the trained versus untrained groups, Mann-Whitney *U* tests were conducted at pretest (Table 127) and posttest (Table 128) collapsed across all listening conditions. Results were not statistically significant for group differences at pretest or posttest, across listening conditions.

Table 127. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions on ratings of confidence for Question 1, Perceived Confidence.

Training Condition	N	Mean Rank	*U*	*p*
Trained	36	39.21	550.5	0.265
Untrained	36	33.79		

Table 128. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 1, Perceived Confidence.

Training Condition	N	Mean Rank	*U*	*p*
Trained	36	37.86	599	0.574
Untrained	36	35.14		

Mann-Whitney *U* tests were conducted to evaluate group differences at pretest (Table 129) and then at posttest (Table 130) for each listening condition on ratings of confidence. Results showed no significant differences in confidence ratings at pretest between the trained and untrained group for each device. Figures displaying non-significant findings are provided in Appendix Q given that the added volume and complexity of such figures would not add to the main body of the document. As such, results of the analyses are graphed in Figure 221, Appendix Q. For the posttest ratings, Mann-Whitney U tests conducted for each listening condition comparing groups, shown in Table 130 and Figure 222, showed no significant differences in ratings of confidence between trained and untrained groups for each listening condition.

Table 129. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at pretest for Question 1, Perceived Confidence.

Training Condition	N	Mean Rank	*U*	*p*
Open				
Trained	12	13.67	58	0.387
Untrained	12	11.33		
TEP-100				
Trained	12	13.42	61	0.51
Untrained	12	11.58		
ComTac™ III				
Trained	12	13.92	55	0.312
Untrained	12	11.08		

Table 130. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at posttest for Question 1, Perceived Confidence.

Training Condition	N	Mean Rank	*U*	*p*
Open				
Trained	12	13.75	57	0.354
Untrained	12	11.25		
TEP-100				
Trained	12	13.33	62	0.552
Untrained	12	11.67		
ComTac™ III				
Trained	12	12.88	67.5	0.788
Untrained	12	12.13		

Within-subjects non-parametric analyses were conducted to compare confidence ratings among devices, i.e., listening conditions, for each group at pretest and then for each group at posttest. In order to compare ratings of the perceived confidence across listening conditions for the trained group in the field, a Friedman two-way analysis of variance by ranks was performed. The Friedman test is the non-parametric counterpart to the repeated-measures ANOVA. In the test procedure, participants are treated as an independent variable (Portney & Watkins, 2009). Data are then organized with participants arranged in rows and levels of conditions in columns (Portney & Watkins, 2009). Ranks are then assigned for each participant, ranking results across the row (e.g., three treatments would result in three rankings for each participant) (Portney &

Watkins, 2009). Ties are handled by assigning an average value for the row. Ranks are then generated for each column, or treatment level (Portney & Watkins, 2009). The null hypothesis supports that the ranks for the columns are equal (Portney & Watkins, 2009). The alternative hypothesis is that at least one pair of treatment levels are different (Portney & Watkins, 2009). At the pretest, Friedman tests showed significant differences in ratings of confidence among the listening conditions for the trained ($\chi^2[2]=19.86$, $p< 0.00$) and untrained ($\chi^2[2]=18.67$, $p< 0.00$) groups. Results are provided in Table 131 and Figure 198. Follow-up pairwise comparisons used a criterion α level of 0.016, given that the overall α level was set to 0.05, but three comparisons (0.05/3=0.016) were conducted for each Friedman's test. For the trained condition, follow-up pairwise comparisons using the Wilcoxon tests, Table 132, at pretest showed significant differences in the open, M=6.2, versus TEP-100 condition, M=3.3, Z=2.96, p=0.003 , open M=6.2, versus ComTac™ III, M=4.2, Z=2.96, p=0.003, and TEP, M=3.3, versus ComTac™ III condition, M=4.2, Z=2.49, p=0.013. For the untrained condition, Wilcoxon signed-rank tests, Table 133, showed significant differences at pretest in ratings of confidence between the open ear, M=5.8 and TEP-100, M=2.9, Z=3.08, p=0.002, and between the open ear and ComTac™ III, M=3.6, Z=3.08, p=0.002. Significant differences in confidence ratings were also found at posttest among listening conditions for trained ($\chi^2[2]=18.47$, $p< 0.002$) and untrained ($\chi^2[2]=16.33$, $p< 0.000$) groups (Table 134 and Figure 199). For the trained group at posttest, Wilcoxon signed-rank tests, Table 135, showed significant differences in ratings of confidence between the open ear, M=6.0, and TEP-100, M=3.6, Z=2.95, p=0.003, and between the open ear and ComTac™ III, M=3.8, Z=3.09, p =0.002. For the untrained group at posttest, Wilcoxon signed-rank tests, Table 136 showed significant differences at posttest in ratings of confidence between the open ear, *M*=5.7, and TEP-100, *M*=3.8, *Z*=2.96, *p*=0.003 and between the open ear and ComTac™ III, *M*=3.7, *Z*=2.82, *p*=0.005.

Table 131. Friedman test results demonstrating significant differences among listening conditions for the trained and untrained groups at pretest for Question 1, Perceived Confidence.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**19.86**	**12**	**2**	**<0.000**
Untrained	**18.67**	**12**	**2**	**<0.000**

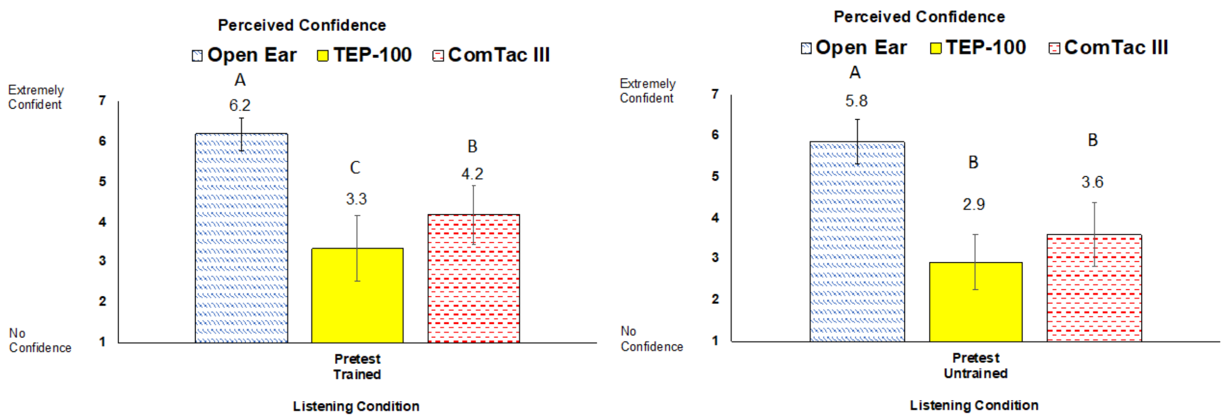


Figure 198. Mean ratings for each group at pretest for Question 1, Perceived Confidence.

Table 132. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at pretest for Question 1, Perceived Confidence.

Listening Condition	Z	*p*
Open - TEP 100	**2.96**	**0.003**
Open - ComTac™ III	**2.96**	**0.003**
TEP-100 – ComTac™ III	**2.49**	**0.013**

Table 133. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at pretest for Question 1, Perceived Confidence.

Listening Condition	Z	*p*
Open - TEP 100	**3.08**	**0.002**
Open - ComTac™ III	**3.08**	**0.002**
TEP-100 – ComTac™ III	1.12	0.263

Table 134. Friedman test results demonstrating significant differences among listening conditions for the trained and untrained groups at posttest for Question 1, Perceived Confidence.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**18.47**	**12**	**2**	**<0.002**
Untrained	**16.33**	**12**	**2**	**<0.000**

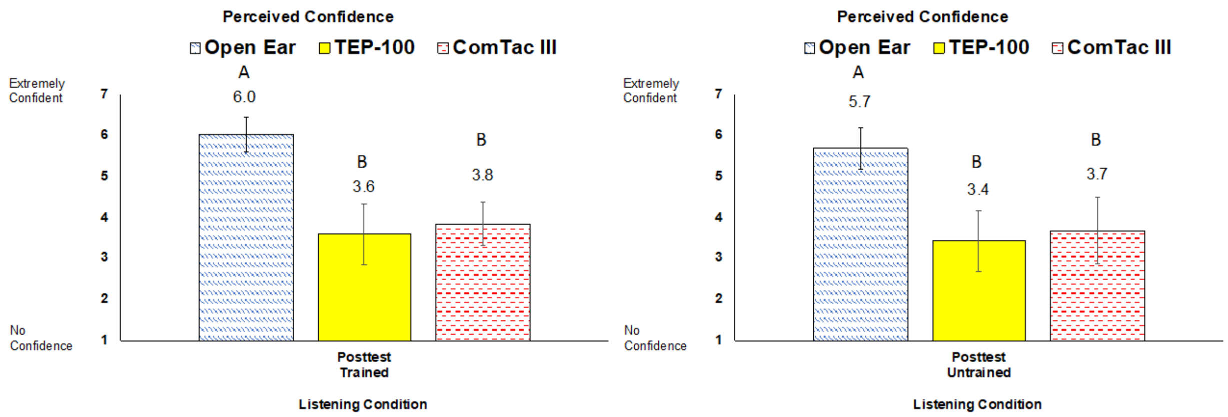


Figure 199. Mean ratings for each group at posttest for Question 1, Perceived Confidence.

Table 135. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at posttest for Question 1, Perceived Confidence.

Listening Condition	*Z*	*p*
Open - TEP 100	**2.95**	**0.003**
Open - ComTac™ III	**3.09**	**0.002**
TEP-100 – ComTac™ III	0.30	0.763

Table 136. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at posttest for Question 1, Perceived Confidence.

Listening Condition	*Z*	*p*
Open - TEP 100	**2.96**	**0.003**
Open - ComTac™ III	**2.82**	**0.005**
TEP-100 – ComTac™ III	0.37	0.714

Question 2. Perceived Accuracy

Participants were asked the following: Rate your perceived **accuracy** to determine sound location under this listening condition from 1 (highly inaccurate) to 7 (highly accurate). On ratings of perceived accuracy, Wilcoxon signed-ranks tests showed no significant differences from pretest compared to posttest that were evaluated for each group and listening condition combination (Table 137). Mean ratings of confidence for each listening condition, group, and stage of training are provided in Figure 200. To evaluate differences in perceived accuracy in the experimental versus control groups, separate Mann-Whitney *U* tests were conducted at pretest (Table 138) and posttest (Table 139) collapsed across all listening conditions. Results were not significant for group differences in ratings of perceived accuracy at pretest or posttest, collapsed across listening conditions for perceived accuracy.

Table 137. Wilcoxon signed-ranks results comparing confidence ratings for pretest versus posttest for each listening condition and group for Question 2, Perceived Accuracy.

Listening Condition	Group	*Z*	*p*
Open ear	Trained	-0.63	0.527
Open ear	Untrained	-0.58	0.564
TEP-100	Trained	0	1
TEP-100	Untrained	0	1
ComTac™ III	Trained	-1.61	0.107
ComTac™ III	Untrained	-0.75	0.454

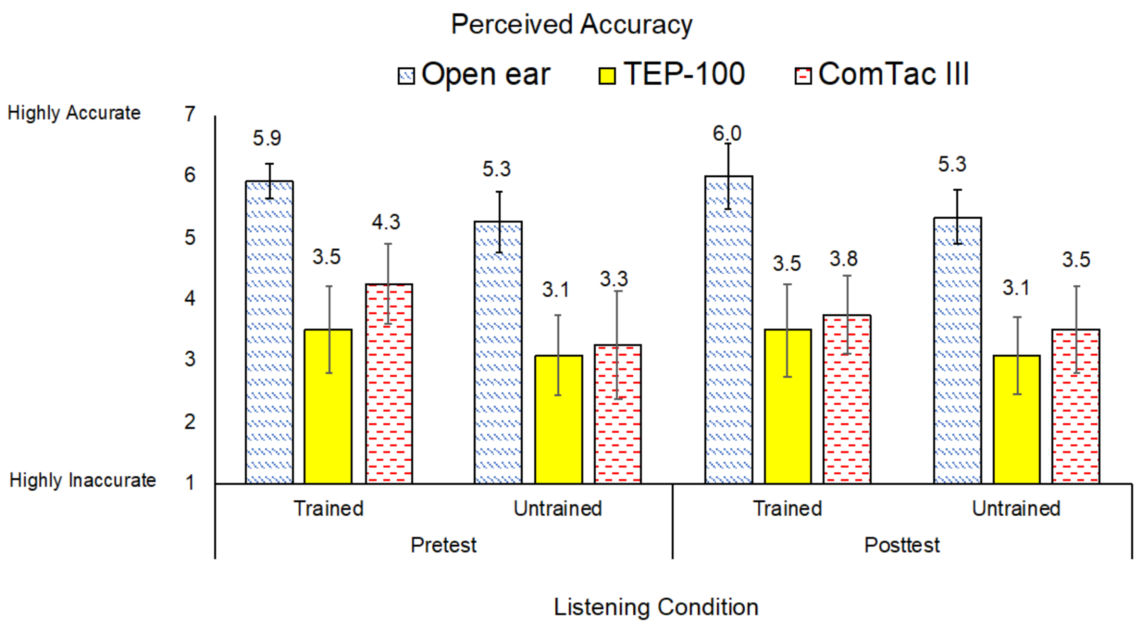


Figure 200. Plotted means and 95% confidence intervals for ratings at pretest and posttest for each group and listening condition for Question 2, Perceived Accuracy.

Table 138. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 2, Perceived Accuracy.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	36	41.08	483	0.058
Untrained	36	31.92		

Table 139. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 2, Perceived Accuracy.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	36	39.15	552.5	0.274
Untrained	36	33.85		

Mann-Whitney *U* tests were conducted to evaluate group differences in ratings of perceived accuracy at pretest (Table 140) and posttest (Table 141 and Figure 201) for each listening condition. At pretest, the trained (*M*=5.9) versus untrained group (*M*=5.3) showed a significant difference in the open ear condition, *U*=37, *p*=0.027. Figures of non-significant findings are displayed in Appendix Q, Figure 223. At posttest, groups showed no significant differences in ratings of perceived accuracy according to listening condition (Table 141).

Table 140. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at pretest for Question 2, Perceived Accuracy.

Training Condition	N	Mean Rank	*U*	*p*
Open				
Trained	**12**	**15.42**	**37.0**	**0.027**
Untrained	**12**	**9.58**		
TEP-100				
Trained		13.63	58.5	0.413
Untrained		11.38		
ComTac™ III				
Trained	12	14.92	43.0	0.088
Untrained	12	10.08		

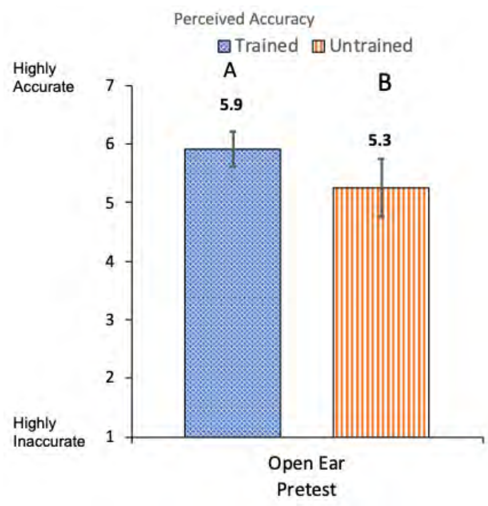


Figure 201. Mann-Whitney *U* results comparing trained and untrained ratings at pretest in the open ear condition for Question 2, Perceived Accuracy.

Table 141. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at posttest for Question 2, Perceived Accuracy.

Training Condition	*n*	Mean Rank	*U*	*p*
Open				
Trained	12	15.08	41.0	0.060
Untrained	12	9.92		
TEP-100				
Trained	12	13.67	58.0	0.400
Untrained	12	11.33		
ComTac™ III				
Trained	12	13.04	65.5	0.699
Untrained	12	11.96		

Within-subjects non-parametric analyses were conducted to compare perceived accuracy ratings among devices for each group at pretest and then for each group at posttest. At pretest, Friedman tests showed significant differences in ratings of confidence among the listening conditions for the trained ($\chi^2[2]=18.73$, $p<0.000$) and untrained ($\chi^2[2]=17.30$, $p<0.000$) groups. Results are provided in Table 142 and Figure 202. Follow-up pairwise comparisons used a criterion significance of α=0.016, for reasons discussed earlier. For the trained condition, follow-up pairwise comparisons using Wilcoxon tests, Table 143, at pretest showed significant differences in the open, *M*=5.9, versus TEP-100 condition, *M*=3.5, *Z*=3.10, *p*=0.002, and between open versus ComTac™ III, M= 4.3, Z=2.84, *p*=0.005. For the untrained condition, Wilcoxon signed-rank tests, Table 144, showed significant differences at pretest in ratings of confidence between the open ear, *M*=5.3 and TEP-100, *M*=3.1, *Z*=3.10, *p*=0.003, and between the open ear and ComTac™ III, *M*=3.3, *Z*=2.96, *p*=0.003. Significant differences in confidence ratings were also found at posttest among listening conditions for trained ($\chi^2[2]=14.68$, $p<0.001$) and untrained ($\chi^2[2]=18.53$, $p<0.000$) groups (Table 145 and Figure 203). For the trained group at posttest, Wilcoxon signed-rank tests, Table 146, showed significant differences in ratings of perceived accuracy between the open ear, *M*=6.0, and TEP-100, *M*=3.5, *Z*=2.87, *p*=0.004, and between the open ear and ComTac™ III, *M*=3.8, *Z*=2.95, *p*=0.003. For the untrained group at posttest, Wilcoxon signed-rank tests, Table 147, showed significant differences in ratings of perceived accuracy between the open ear, *M*=5.3, and TEP-100, *M*=3.1, *Z*=3.11, *p*=0.002 and between the open ear and ComTac™ III, *M*=3.5, *Z*=2.97, *p*=0.003.

Table 142. Friedman test results demonstrating significant differences in perceived ratings among listening conditions for the trained and untrained groups at pretest, Perceived Accuracy.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**18.73**	**12**	**2**	**<0.000**
Untrained	**17.30**	**12**	**2**	**<0.000**

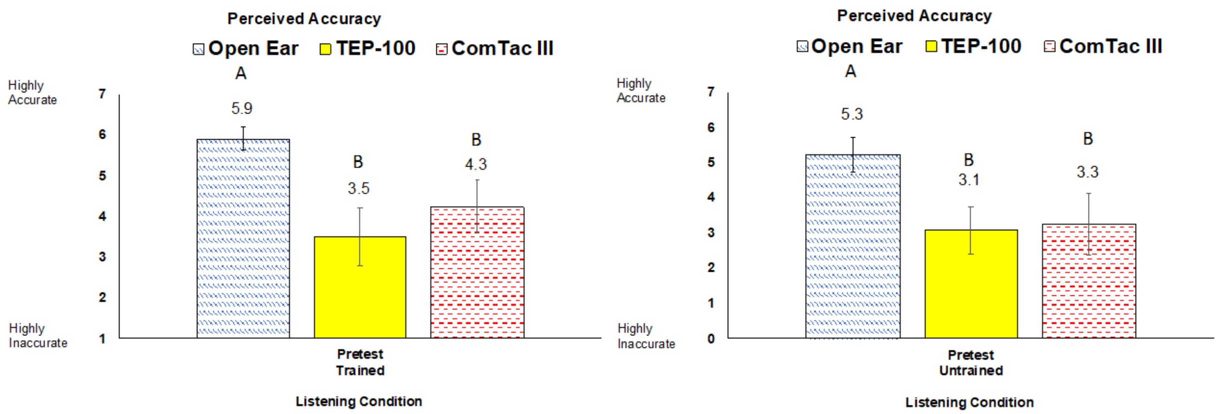


Figure 202. Mean ratings for each group at pretest for Question 2, Perceived Accuracy.

Table 143. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at pretest for Question 2, Perceived Accuracy.

Listening Condition	*Z*	*p*
Open - TEP 100	**3.10**	**0.002**
Open - ComTac™ III	**2.84**	**0.005**
TEP-100 – ComTac™ III	2.07	0.038

Table 144. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at pretest for Question 2, Perceived Accuracy.

Listening Condition	*Z*	*p*
Open - TEP 100	**3.02**	**0.003**
Open - ComTac™ III	**2.96**	**0.003**
TEP-100 – ComTac™ III	0.24	0.810

Table 145. Friedman test results demonstrating significant differences in perceived accuracy ratings among listening conditions for the trained and untrained groups at posttest for Question 2, Perceived Accuracy.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**14.68**	**12**	**2**	**0.001**
Untrained	**18.53**	**12**	**2**	**<0.000**

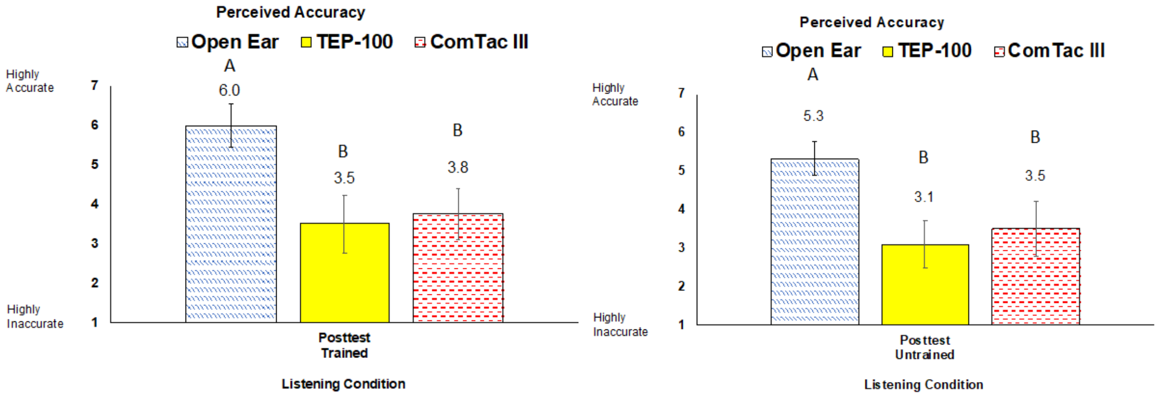


Figure 203. Mean ratings for each group at posttest for Question 2, Perceived Accuracy.

Table 146. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at posttest for Question 2, Perceived Accuracy.

Listening Condition	*Z*	*p*
Open - TEP 100	**2.87**	**0.004**
Open - ComTac™ III	**2.95**	**0.003**
TEP-100 – ComTac™ III	0.59	0.558

Table 147. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at posttest for Question 2, Perceived Accuracy.

Listening Condition	*Z*	*p*
Open - TEP 100	**3.11**	**0.002**
Open - ComTac™ III	**2.97**	**0.003**
TEP-100 – ComTac™ III	1.12	0.265

Question 3. Perceived Difficulty

Participants were asked the following: Rate how **difficult** it was to judge the **location** of the sounds under this listening condition from 1 (extremely difficult) to 7 (extremely easy). On ratings of difficulty, Wilcoxon signed-ranks tests showed no significant differences from pretest compared to posttest evaluated for each group and listening condition combination (Table 148). Mean ratings of difficulty for each listening condition, group, and stage of training are provided in Figure 204. To evaluate differences in perceived difficulty in the experimental versus control groups, separate Mann-Whitney *U* tests were conducted at pretest (Table 149) and posttest (Table 150) collapsed across all listening conditions. Results were not significant for group differences in ratings of perceived difficulty at pretest or posttest, collapsed across listening conditions. Throughout this discussion, it is important to note that *lower* ratings reflect *higher* difficulty.

Table 148. Wilcoxon signed-ranks results comparing difficulty ratings for pretest versus posttest for each listening condition and group for Question 3, Perceived Difficulty.

Listening Condition	Group	*Z*	*p*
Open ear	Trained	-1.00	0.317
Open ear	Untrained	-0.33	0.739
TEP-100	Trained	-0.26	0.792
TEP-100	Untrained	0.00	1
ComTac™ III	Trained	-1.03	0.305
ComTac™ III	Untrained	-0.91	0.366

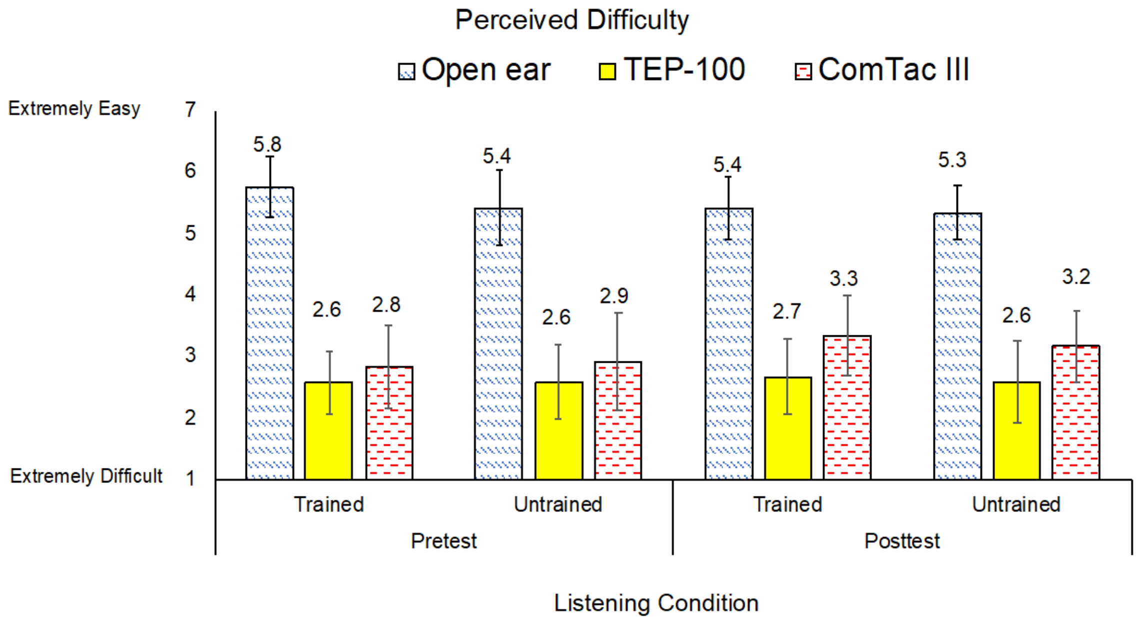


Figure 204. Plotted means and 95% confidence intervals for ratings at pretest and posttest for each group and listening condition for Question 3, Perceived Difficulty.

Table 149. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 3, Perceived Difficulty.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	36	37	630	0.836
Untrained	36	36		

Table 150. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 3, Perceived Difficulty.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	36	37.29	619.5	0.744
Untrained	36	35.71		

Mann-Whitney *U* tests were conducted to evaluate group differences in ratings of perceived difficulty at pretest (Table 151), and posttest (Table 152 and Figure 224, Appendix Q) for each listening condition. Results showed that at pretest and posttest, groups showed no significant differences in ratings of perceived difficulty by listening condition.

Table 151. Results of Mann-Whitney U tests comparing confidence ratings between training groups for each listening condition at pretest for Question 3, Perceived Difficulty.

Training Condition	*n*	Mean Rank	*U*	*p*
Open				
Trained	12	13.42	61.0	0.493
Untrained	12	11.58		
TEP-100				
Trained	12	12.79	68.5	0.829
Untrained	12	12.21		
ComTac™ III				
Trained	12	12.46	71.5	0.976
Untrained	12	12.54		

Table 152. Results of the Mann-Whitney *U* test evaluating group differences at posttest for the TEP-100 condition for Question 3, Perceived Difficulty.

Training Condition	*n*	Mean Rank	*U*	*p*
Open				
Trained	12	13.00	66.0	0.710
Untrained	12	12.00		
TEP-100				
Trained	12	12.96	66.5	0.742
Untrained	12	12.04		
ComTac™ III				
Trained	12	13.13	64.5	0.653
Untrained	12	11.88		

Within-subjects non-parametric analyses were conducted to compare perceived difficulty ratings among devices for each group at pretest and then for each group at posttest. Of note, lower ratings reflect increased difficulty. At pretest, Friedman tests showed significant differences in difficulty ratings among the listening conditions for the trained ($\chi^2[2]=19.24$, $p<0.000$) and untrained ($\chi^2[2]=17.64$, $p<0.000$) groups. Results are provided in Table 153 and Figure 205. Follow-up pairwise comparisons used $\alpha=0.016$. For the trained group at pretest, follow-up pairwise comparisons using Wilcoxon tests, Table 154, showed significant differences in the open, $M=5.8$, and TEP-100 condition, $M=2.6$, $Z=3.09$, $p=0.002$, and between open and ComTac™ III, M=2.8, Z=3.10, $p=0.002$. For the untrained condition, Wilcoxon signed-rank tests, Table 155 showed significant differences at pretest in ratings of confidence between the open ear, *M*=5.4 and TEP-100, *M*=2.6, *Z*=3.08, *p*=0.002, and between the open ear and ComTac™ III, Mean=2.9, Z=2.90, *p*=0.004. Significant differences in difficulty ratings were also found at posttest among listening conditions for trained ($\chi^2[2]=16.31$, $p<0.00$) and untrained ($\chi^2[2]=20.31$, $p<0.00$) groups (Table 156 and Figure 206). For the trained group at posttest, Wilcoxon signed-rank tests, Table 157, showed significant differences in ratings of perceived difficulty between the open ear, *M*=5.4, and TEP-100, *M*=2.7, *Z*=2.82, *p*=0.005, and between the open ear and ComTac™ III, Mean=3.3, Z=3.10, *p*=0.002. For the untrained group at posttest, Wilcoxon signed-rank tests, Table 158, showed significant differences in ratings of difficulty between the open ear, *M*=5.3, and TEP-100, *M*=2.6, *Z*=3.09, *p*=0.002 and between the open ear and ComTac™ III, *M*=3.2, *Z*=3.09, *p*=0.002.

Table 153. Friedman test results demonstrating significant differences in perceived difficulty ratings among listening conditions for the trained and untrained groups at pretest for Question 3, Perceived Difficulty.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**19.24**	**12**	**2**	**<0.000**
Untrained	**17.64**	**12**	**2**	**<0.000**

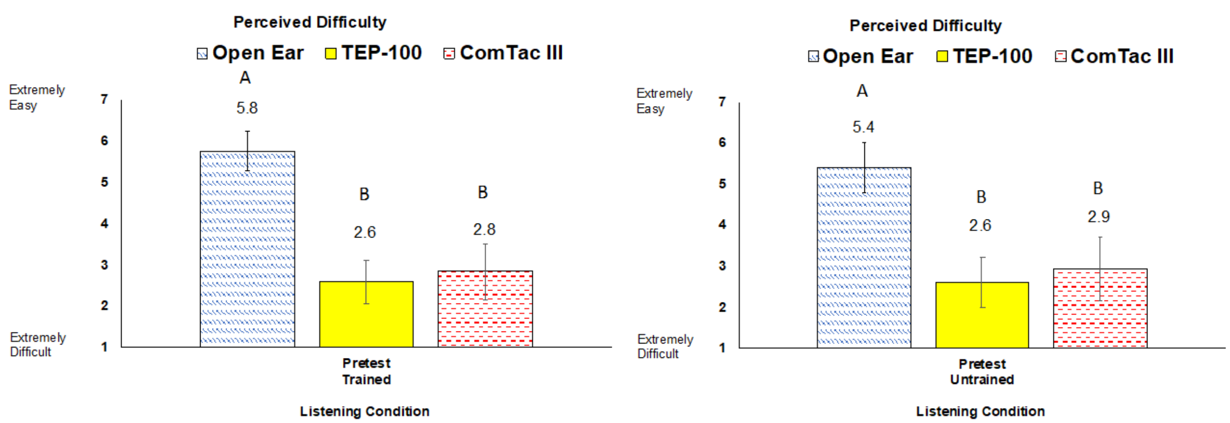


Figure 205. Mean ratings for each group at pretest for Question 3, Perceived Difficulty.

Table 154. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at pretest for Question 3, Perceived Difficulty.

Listing Condition	*Z*	*p*
Open - TEP 100	**3.09**	**0.002**
Open - ComTac™ III	**3.10**	**0.002**
TEP-100 – ComTac™ III	0.81	0.417

Table 155. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at pretest for Question 3, Perceived Difficulty.

Listening Condition	*Z*	*p*
Open - TEP 100	**3.08**	**0.002**
Open - ComTac™ III	**2.90**	**0.004**
TEP-100 – ComTac™ III	0.72	0.473

Table 156. Friedman test results demonstrating significant differences in ratings among listening conditions for the trained and untrained groups at posttest for Question 3, Perceived Difficulty.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**16.31**	**12**	**2**	**<0.000**
Untrained	**20.31**	**12**	**2**	**<0.000**

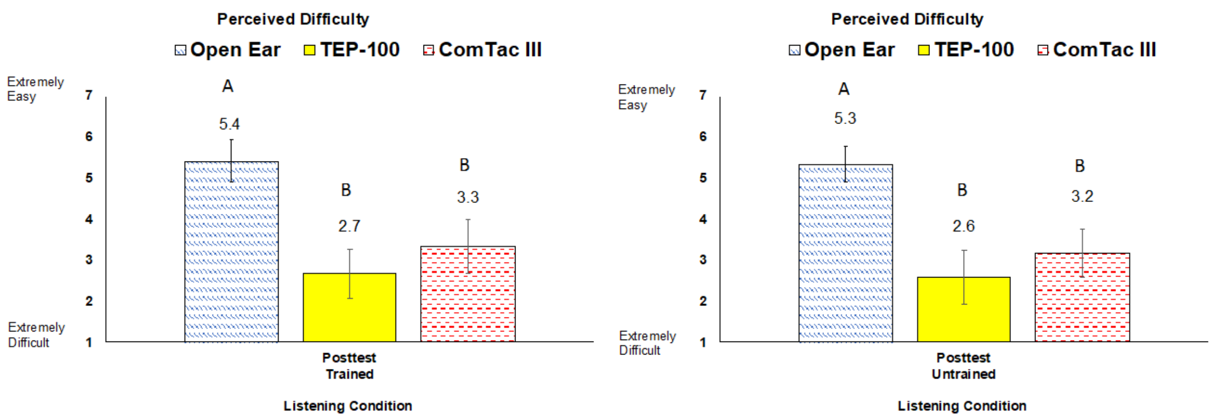


Figure 206. Mean ratings for each group at posttest for Question 3, Perceived Difficulty.

Table 157. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the trained group at posttest for Question 3, Perceived Difficulty.

Listening Condition	Z	*p*
Open - TEP 100	**2.82**	**0.005**
Open - ComTac™ III	**3.10**	**0.002**
TEP-100 – ComTac™ III	1.27	0.203

Table 158. Wilcoxon signed-rank pairwise comparisons for the open, TEP-100, and ComTac™ III for the untrained group at posttest for Question, Perceived Difficulty.

Listening Condition	Z	*p*
Open - TEP 100	**3.09**	**0.002**
Open - ComTac™ III	**3.09**	**0.002**
TEP-100 – ComTac™ III	1.41	0.159

Question 4. Perceived Reaction Time

Participants were asked the following: Rate your perceived reaction time in determining the sound location under this listening condition" from 1 (extremely slow) to 7 (extremely fast). On ratings of perceived reaction time, Wilcoxon signed-ranks tests showed significant differences from pretest, *M*=4.1, compared to posttest, *M*=3.3, in the trained group for the TEP-100 listening condition, *Z*=-2.17, *p*=0.030 (Table 159 and Figure 207). Of note, a lower rating is indicative of slower perceived reaction times. Therefore, the aforementioned significant difference between pretest and posttest is consistent with participants' perception of feeling slower during the outdoor posttest. Significant differences also occurred in the trained group for the ComTac™ III listening condition from pretest, *M*=4.3, to posttest, *M*=3.3, Z=-2.49, *p*=0.031. The ComTac™ III, trained results are also consistent with a slower perceived reaction time during the posttest. To evaluate differences in perceived reaction time in the trained versus untrained groups, separate Mann-Whitney *U* tests were conducted at pretest (Table 160) and posttest (Table 161) collapsed across all listening conditions. Results were not significant for group differences in ratings of perceived reaction time at pretest or posttest for each listening condition (Figure 225, Appendix Q) for Question 4, Perceived Reaction Time.

Table 159. Wilcoxon signed-ranks results comparing response time for pretest versus posttest for each listening condition and group for Question 4, Perceived Reaction Time.

Listening Condition	Group	*Z*	*p*
Open ear	Trained	-1.4	0.161
Open ear	Untrained	-1.51	0.132
TEP-100	**Trained**	**-2.17**	**0.030**
TEP-100	Untrained	0.00	1.000
ComTac™ III	**Trained**	**-2.49**	**0.013**
ComTac™ III	Untrained	-0.58	0.564

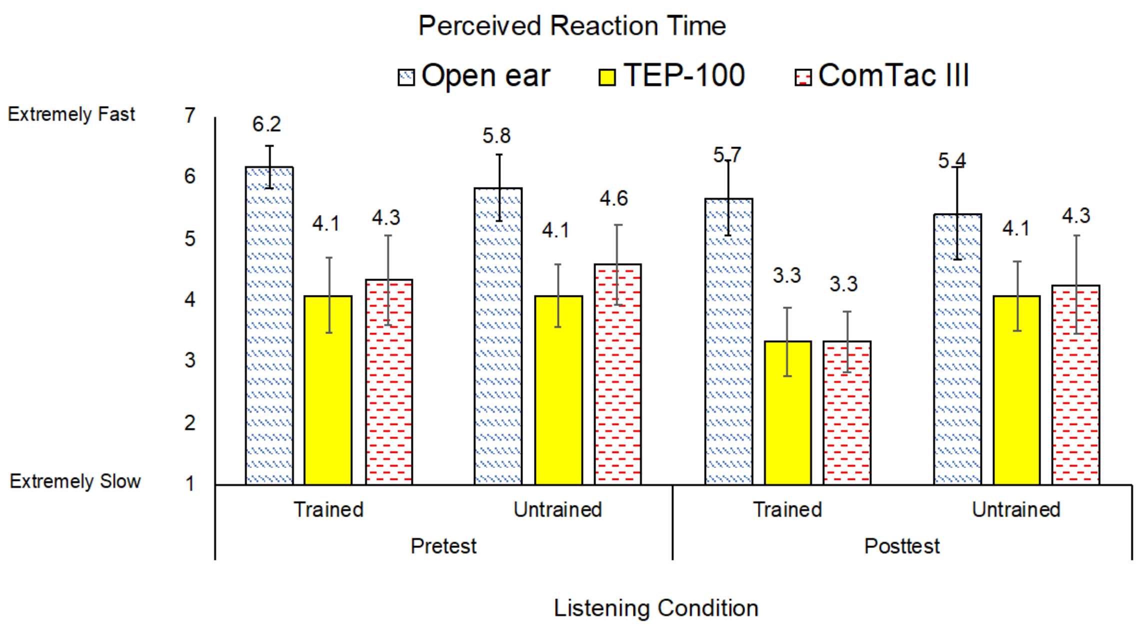


Figure 207. Plotted means and 95% confidence intervals for ratings at pretest and posttest for each group and listening condition for Question 4, Perceived Reaction Time.

Table 160. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 4, Perceived Reaction Time.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	12	13.17	64	0.631
Untrained	12	11.83		

Table 161. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 4, Perceived Reaction Time.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	12	10.08	43	0.078
Untrained	12	14.92		

Mann-Whitney *U* tests were conducted to evaluate group differences in ratings of perceived reaction time at pretest (Table 162) and posttest (Table 163) Results are displayed in Figures 226 (pretest) and 227 (posttest) in Appendix Q. At pretest and posttest, groups showed no significant differences in ratings of perceived reaction times between listening conditions.

Table 162. Results of the Mann-Whitney *U* test evaluating group differences at pretest for each listening condition for Question 4, Perceived Reaction Time.

Training Condition	*n*	Mean Rank	*U*	*p*
Open				
Trained	12	13.67	58	0.373
Untrained	12	11.33		
TEP-100				
Trained	12	12.50	72	1.000
Untrained	12	12.50		
ComTac™ III				
Trained	12	11.63	61.5	0.531
Untrained	12	13.38		

Table 163. Results of the Mann-Whitney *U* test evaluating group differences at posttest for each listening condition for Question 4, Perceived Reaction Time.

Training Condition	N	Mean Rank	*U*	*p*
Open				
Trained	12	13.17	64	0.631
Untrained	12	11.83		
TEP-100				
Trained	12	10.08	43	0.078
Untrained	12	14.92		
ComTac™ III				
Trained	12	9.92	41	0.061
Untrained	12	15.08		

Within-subjects non-parametric analyses were conducted to compare perceived reaction time ratings among devices for each group at pretest and then for each group at posttest. At pretest, Friedman tests showed significant differences in reaction times among the listening conditions for the trained (χ^2[2]=15.14, $p< 0.001$) and untrained (χ^2[2]=15.57, $p<0.001$) groups. Results are provided in Table 164 and Figure 208. Follow-up pairwise comparisons used α=0.016. For the trained group at pretest, Wilcoxon signed-rank tests, Table 165, showed significant differences in ratings of perceived reaction time between the open ear, *M*=6.2, and TEP-100, *M*=4.1, *Z*=2.96, *p*=0.003, and between the open ear and ComTac™ III, *M*=4.3, *Z*=2.83, *p*=0.005. For the untrained group at pretest, Wilcoxon signed-rank tests, Table 166, showed significant differences in ratings of perceived response time between the open ear, *M*=5.8, and TEP-100, *M*=4.1, *Z*=2.84, *p*=0.005 and between the open ear and ComTac™ III, *M*=4.6, Z=2.68, *p*=0.007. Significant differences in confidence ratings were also found at posttest among listening conditions for trained ($\chi^2[2]=14$, $p<0.001$) and untrained ($\chi^2[2]=8.93$, $p<0.012$) groups (Table 167 and Figure 209). For the trained group at posttest, Wilcoxon signed-rank tests, Table 168, showed significant differences in ratings of reaction time between the open ear, *M*=5.7, and TEP-100, Mean=3.3, *Z*=2.83, *p*=0.005, and between the open ear and ComTac™ III, *M*=3.3, *Z*=2.86, *p*=0.004. For the untrained group at posttest, Wilcoxon signed-rank tests, Table 169, showed significant differences in ratings of difficulty between the open ear, *M*=5.4, and TEP-100, *M*=4.1, *Z*=2.56, *p*=0.011.

Table 164. Friedman test results demonstrating significant among listening conditions for the trained and untrained groups at pretest Question 4, Perceived Reaction Time.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**15.14**	**12**	**2**	**<0.001**
Untrained	**15.57**	**12**	**2**	**<0.001**

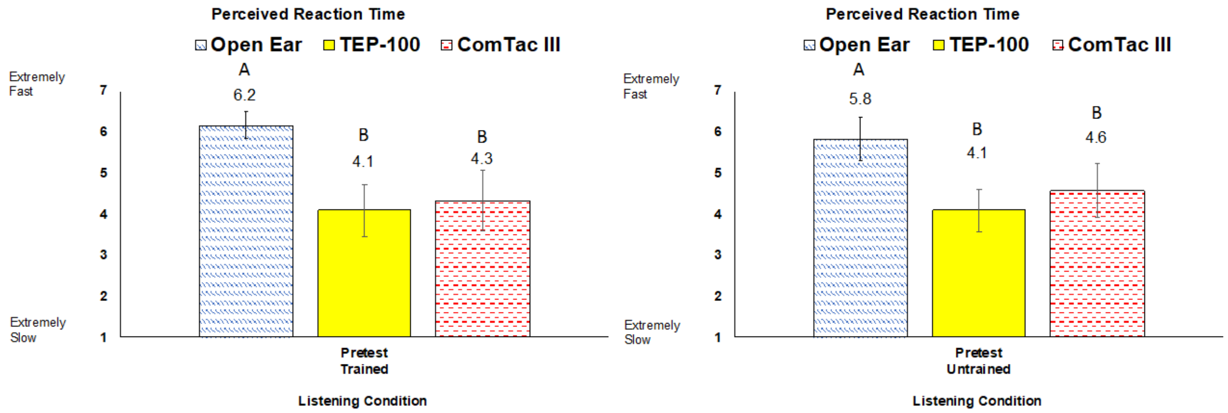


Figure 208. Mean ratings for each group and listening condition at pretest for Question 4, Perceived Reaction Time.

Table 165. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 4 at pretest for the trained group at pretest, Perceived Reaction Time.

Listening Condition	*Z*	*p*
Open - TEP 100	**2.96**	**0.003**
Open - ComTac™ III	**2.83**	**0.005**
TEP-100 – ComTac™ III	0.81	0.417

Table 166. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 4 at pretest for the untrained group, Perceived Reaction Time.

Listening Condition	*Z*	*p*
Open - TEP 100	**2.84**	**0.005**
Open - ComTac™ III	**2.68**	**0.007**
TEP-100 – ComTac™ III	1.73	0.083

Table 167. Friedman test results demonstrating significant differences in ratings among listening conditions for the trained and untrained groups at posttest, Question 4, Perceived Reaction Time.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**14**	**12**	**2**	**0.001**
Untrained	**8.93**	**12**	**2**	**0.012**

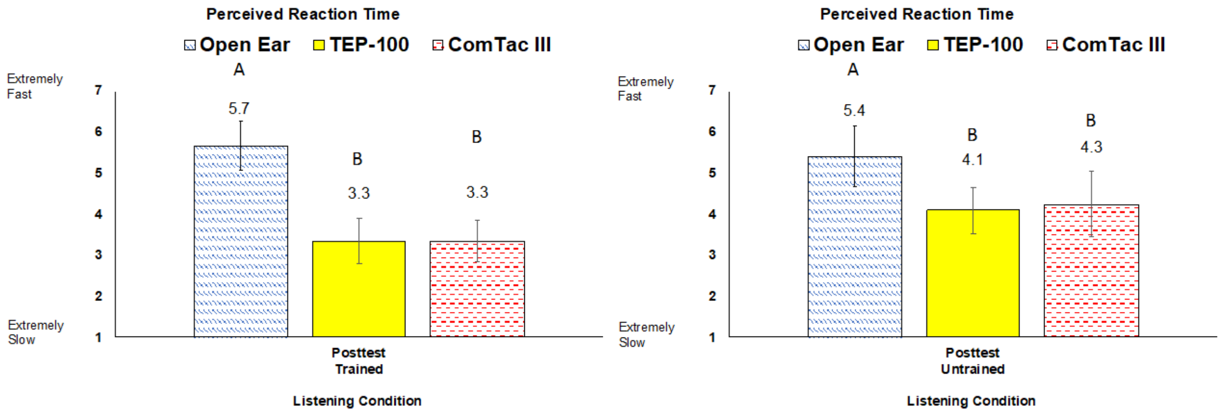


Figure 209. Mean ratings of perceived reaction time for each group and listening condition at posttest for Question 4, Perceived Reaction Time.

Table 168. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 4 for the trained group at posttest, Perceived Reaction Time.

Listening Condition	*Z*	*p*
Open - TEP 100	**2.83**	**0.005**
Open - ComTac™ III	**2.86**	**0.004**
TEP-100 – ComTac™ III	0.06	0.951

Table 169. Table Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for the untrained group at posttest for Question 4, Perceived Reaction Time.

Listening Condition	*Z*	*p*
Open - TEP 100	**2.56**	**0.011**
Open - ComTac™ III	2.26	0.024
TEP-100 – ComTac™ III	0.67	0.501

Question 5. Perceived Comfort

Participants were asked the following: Please rate how comfortable this hearing protection device condition (or open ear) was while wearing it during the experiment from 1 (extremely uncomfortable) to 7 (extremely comfortable). This rating thus was intended to apply to their comfort perception having worn the product during both in the office experiments and in the field tests. Wilcoxon signed-ranks tests showed significant differences in ratings of comfort in the untrained group for the ComTac™ III condition from pretest, M=5.08, to posttest, M=4.42, Z=-2.53, p=0.011 (Table 170 and Figure 210). Results are consistent with the perception of comfort decreasing pretest to posttest in the untrained group for the ComTac™ III. To evaluate differences in ratings of comfort in the experimental versus control groups, Mann-Whitney *U* tests were conducted at pretest (Table 171) and posttest (Table 172) collapsed across all listening conditions. Results were not significant for group differences in perceived comfort at pretest or posttest. Mann-Whitney *U* tests were conducted to evaluate group differences at pretest (Table 180 and Figure 228, Appendix Q) and at posttest (Table 181 and Figure 229, Appendix Q) for each listening condition. Groups showed no significant differences in ratings of comfort according to listening condition at neither pretest nor posttest.

Table 170. Wilcoxon signed-ranks results comparing comfort ratings for pretest versus posttest for each listening condition and group for Question 5, Perceived Comfort.

Listening Condition	Group	*Z*	*p*
Open ear	Trained	-0.38	0.705
Open ear	Untrained	-0.82	0.414
TEP-100	Trained	-0.52	0.607
TEP-100	Untrained	-0.42	0.676
ComTac™ III	Trained	-1.19	0.234
ComTac™ III	**Untrained**	**-2.53**	**0.011**

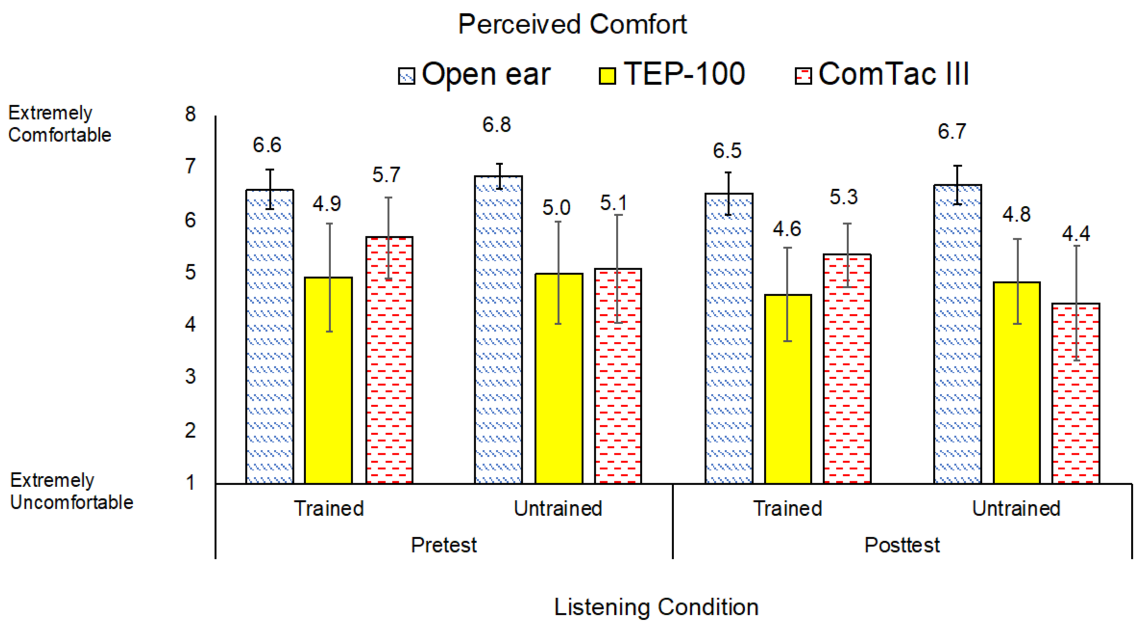


Figure 210. Plotted means and 95% confidence intervals at pretest and posttest for each group and listening condition for Question 5, Perceived Comfort.

Table 171. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 5, Perceived Comfort.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	36	36.42	645	0.972
Untrained	36	36.58		

Table 172. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 5, Perceived Comfort.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	36	36.88	634	0.876
Untrained	36	36.13		

Table 173. Results of the Mann-Whitney *U* test evaluating group differences at pretest for each listening condition for Question 5, Perceived Comfort.

Training Condition	N	Mean Rank	*U*	*p*
Open				
Trained	12	11.42	59	0.320
Untrained	12	13.58		
TEP-100				
Trained		12.58	71	0.952
Untrained		12.42		
ComTac™ III				
Trained		13.58	59	0.438
Untrained		11.42		

Table 174. Results of the Mann-Whitney *U* test evaluating group differences at posttest for each listening condition for Question 5, Perceived Comfort.

Training Condition	*n*	Mean Rank	*U*	*p*
Open				
Trained	12	11.58	61	0.444
Untrained	12	13.42		
TEP-100				
Trained	12	12.42	71	0.953
Untrained	12	12.58		
ComTac™ III				
Trained	12	14.17	52	0.237
Untrained	12	10.83		

Within-subjects non-parametric analyses were conducted to compare perceived comfort ratings among devices for each group at pretest and then for each group at posttest. At pretest, Friedman tests showed significant differences in comfort ratings among the listening conditions for the trained ($\chi^2[2]=10.27$, $p<0.006$) and untrained ($\chi^2[2]=13.26$, $p<0.001$) groups. Results are provided in Table 175 and Figure 211. Follow-up pairwise comparisons used $\alpha=0.016$, for reasons explained previously. For the trained group at pretest, follow-up pairwise comparisons using Wilcoxon tests, Table 176, showed a significant difference between the open ear, $M=6.6$, and TEP-100 condition, $M=4.9$, Z=2.75, $p=0.006$. For the untrained condition at pretest, follow-up pairwise comparisons using Wilcoxon tests, Table 177, showed significant differences at pretest in ratings of comfort between the open ear, *M*=6.8 and TEP-100, *M*=5.0, *Z*=2.68, *p*=0.007, and between the open ear and ComTac™ III, *M*=5.1, *Z*=2.69, *p*=0.007. Significant differences in comfort ratings were also found at posttest among listening conditions for trained ($\chi^2[2]=12.4$, $p<0.002$) and untrained ($\chi^2[2]=14.26$, $p<0.001$) groups (Table 178 and Figure 212). For the trained group at posttest, Wilcoxon signed-rank tests, Table 179, showed significant differences in ratings of comfort between the open ear, *M*=6.5, and TEP-100, *M*=4.6, *Z*=2.69, *p*=0.007. For the untrained group at posttest, Wilcoxon signed-rank tests, Table 180, showed significant differences in ratings of difficulty between the open ear, *M*=6.7, and TEP-100, *M*=4.8, *Z*=2.83, *p*=0.005 and between the open ear and ComTac™ III, *M*=4.4, *Z*=2.67, *p*=0.008.

Table 175. Friedman test results demonstrating significant differences in perceived comfort ratings among listening conditions for the trained and untrained groups at pretest, Question 5, Perceived Comfort.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**10.17**	**12**	**2**	**0.006**
Untrained	**13.26**	**12**	**2**	**0.001**

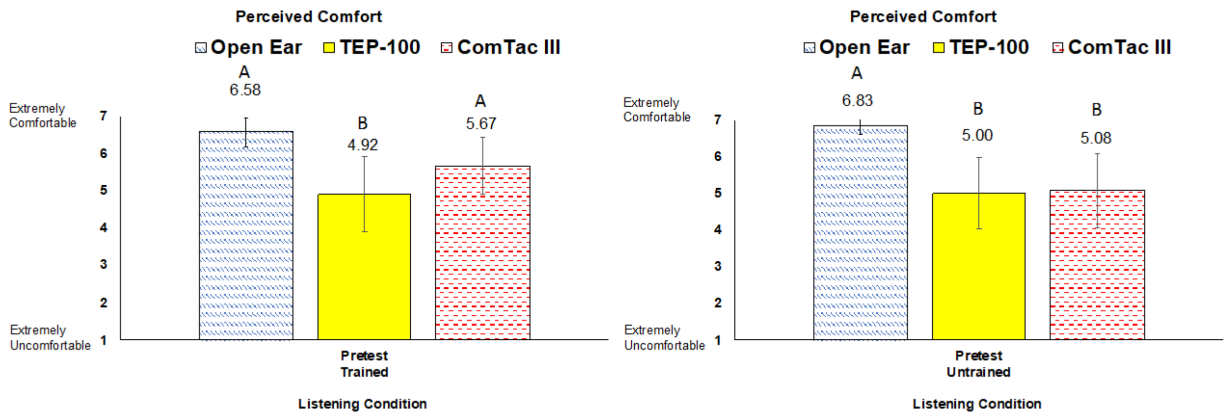


Figure 211. Mean ratings of perceived comfort for each group and listening condition at pretest for Question 5, Perceived Comfort.

Table 176. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 5 at pretest for the trained group (α=0.016) for Question 5, Perceived Comfort.

Listing Condition	*Z*	*p*
Open - TEP 100	**2.75**	**0.006**
Open - ComTac™ III	1.85	0.064
TEP-100 – ComTac™ III	1.38	0.169

Table 177. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 5 at pretest for the untrained group (α=0.016) for Question 5, Perceived Comfort.

Listing Condition	*Z*	*p*
Open - TEP 100	**2.68**	**0.007**
Open - ComTac™ III	**2.69**	**0.007**
TEP-100 – ComTac™ III	0.48	0.629

Table 178. Friedman test results demonstrating significant differences in comfort ratings among listening conditions for the trained and untrained groups at posttest for Question 5, Perceived Comfort.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**12.41**	**12**	**2**	**0.002**
Untrained	**14.26**	**12**	**2**	**0.001**

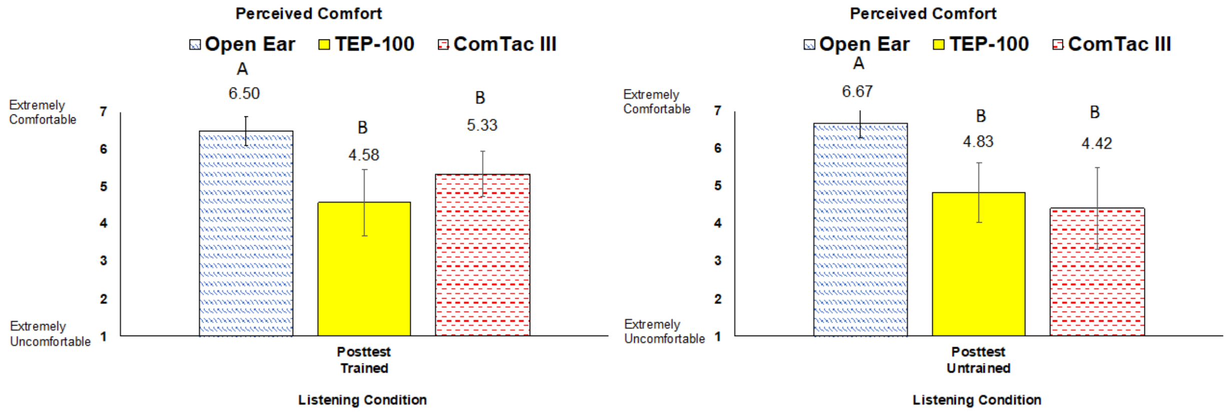


Figure 212. Mean ratings of perceived comfort for each group and listening condition at pretest for Question 5, Perceived Comfort.

Table 179. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III for Question 5 at posttest for the trained group for Question 5, Perceived Comfort.

Listing Condition	*Z*	*p*
Open - TEP 100	**2.69**	**0.007**
Open - ComTac™ III	2.26	0.024
TEP-100 – ComTac™ III	1.54	0.123

Table 180. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at posttest for the untrained group for Question 5, Perceived Comfort.

Listening Condition	Z	*p*
Open - TEP 100	**2.83**	**0.005**
Open - ComTac™ III	**2.67**	**0.008**
TEP-100 – ComTac™ III	0.71	0.478

Question 6. Likelihood of Wearing Device during a Sound Localization Task

Participants were asked the following: How likely would you be to wear this hearing protection device during a task similar to this experiment that required sound localization, if you had access to this hearing protection device [or open ear] from 1 (extremely unlikely) to 7 (extremely likely). Wilcoxon signed-ranks tests showed no significant differences in ratings of likelihood to maintain the current listening condition for the trained and untrained groups for each listening group from pretest to posttest (Table 181 and Figure 213). To evaluate differences in ratings of confidence in the experimental versus control groups, Mann-Whitney *U* tests were conducted at pretest (Table 182) and posttest (Table 183) collapsed across all listening conditions. Results were not significant for group differences in likelihood to wear the device at pretest or posttest collapsed across listening conditions. Mann-Whitney *U* tests were conducted to evaluate group differences at pretest (Table 184 and Figure 230, Appendix Q) and posttest (Table 185 and Figure 231, Appendix Q) for each listening condition. At neither pretest nor posttest, groups showed no significant differences in likelihood to wear the device under similar conditions according to device, or listening condition for Question 6, Likelihood of Wearing Device.

Table 181. Wilcoxon signed-ranks results comparing likelihood of maintaining the current listening condition given a similar task for pretest versus posttest for each listening condition and group for Question 6, Question 6. Likelihood of Wearing Device.

Listening Condition	Group	*Z*	*p*
Open ear	Trained	-0.28	0.783
Open ear	Untrained	-1.00	0.317
TEP-100	Trained	-0.26	0.796
TEP-100	Untrained	-0.18	0.857
ComTac™ III	Trained	-1.42	0.155
ComTac™ III	Untrained	-0.29	0.773

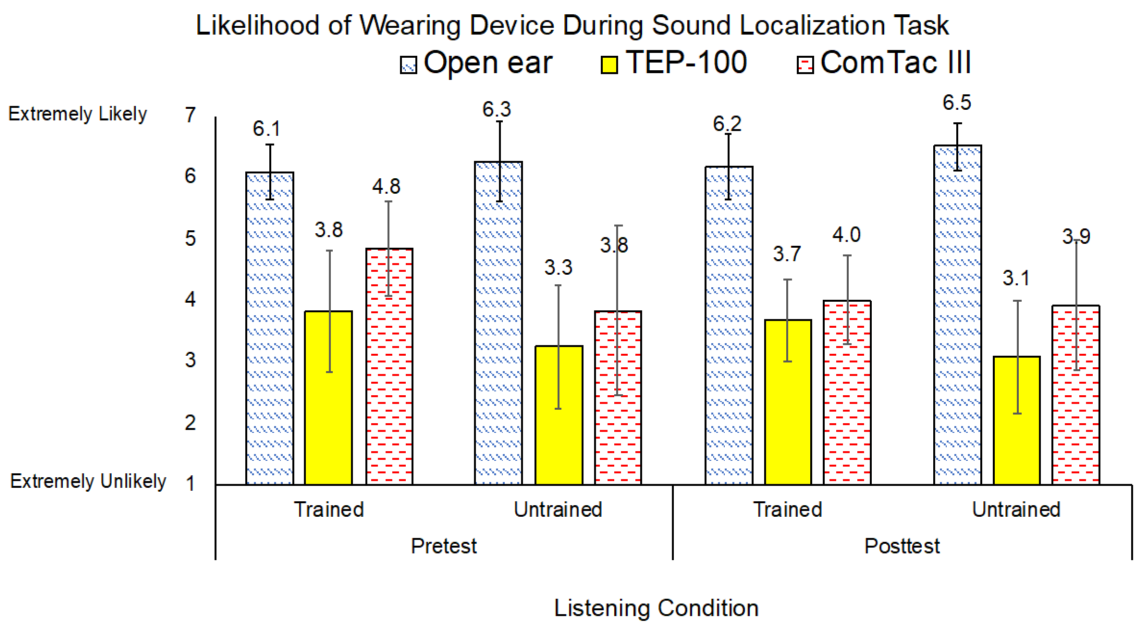


Figure 213. Plotted means and 95% confidence intervals for likelihood to maintain the same listening condition ratings at pretest and posttest for each group and listening condition for Question 6, Likelihood of Wearing Device.

Table 182. Results of the Mann-Whitney *U* test assessing group differences at pretest collapsed across all listening conditions for Question 6, Likelihood of Wearing Device.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	36	38.10	590.5	0.511
Untrained	36	34.9		

Table 183. Results of the Mann-Whitney *U* test assessing group differences at posttest collapsed across all listening conditions for Question 6, Likelihood of Wearing Device.

Training Condition	*n*	Mean Rank	*U*	*p*
Trained	36	36.57	645.5	0.977
Untrained	36	36.43		

Table 184. Results of the Mann-Whitney *U* test evaluating group differences at pretest for each listening condition for Question 6, Likelihood of Wearing Device.

Training Condition	*n*	Mean Rank	*U*	*p*
Open				
Trained	12	11.46	59.5	0.434
Untrained	12	13.54		
TEP-100				
Trained		13.67	58	0.412
Untrained		11.33		
ComTac™ III				
Trained		13.88	55.5	0.333
Untrained		11.13		

Table 185. Results of the Mann-Whitney *U* test evaluating group differences at posttest for each listening condition for Question 6, Likelihood of Wearing Device.

Training Condition	*n*	Mean Rank	*U*	*p*
Open				
Trained	12	11.33	58	0.373
Untrained	12	13.67		
TEP-100				
Trained	12	13.75	57	0.375
Untrained	12	11.25		
ComTac™ III				
Trained	12	12.79	68.5	0.837
Untrained	12	12.21		

Within-subjects non-parametric analyses were conducted to compare participant's likelihood of wearing the device, or keeping an open ear, given similar listening conditions for each group at pretest and then for each group at posttest. At pretest, Friedman tests showed significant differences in ratings among the listening conditions for the trained ($\chi^2[2]=9.95$, $p<$ 0.007) and untrained ($\chi^2[2]=12.76$, $p<0.002$) groups. Results are provided in Table 186 and Figure 214. Follow-up pairwise comparisons used α=0.016 for reasons explained previously. For the trained group at pretest, follow-up pairwise comparisons using Wilcoxon tests, Table 187, showed a significant difference between the open, *M*=6.1, and TEP-100 condition, *M*=3.8, *Z*=2.73, *p*=0.006. For the untrained condition at pretest, follow-up pairwise comparisons using Wilcoxon tests, Table 188, showed significant differences at pretest in ratings between the open ear, *M*=6.3 and TEP-100, *M*=3.3, Z= 3.08, *p*=0.002, and between the open ear and ComTac™ III, *M*=3.8, Z=2.41, *p*=0.016. Significant differences in likelihood ratings were also found at posttest among listening conditions for trained (χ^2[2]=15.45, *p*<0.002) and untrained (χ^2[2]=18.73, *p*<0.000) groups (Table 189 and Figure 215). For the trained group at posttest, Wilcoxon signed-rank tests, Table 190, showed significant differences in ratings between the open ear, *M*=6.2, and TEP-100, *M*=3.7, Z=2.95, *p*=0.003 and between the open ear and the ComTac™ III, *M*=4.0, Z=2.7, *p*=0.007. For the untrained group at posttest, Wilcoxon signed-rank tests, Table 191, showed significant differences in likelihood ratings between the open ear, *M*=6.5, and TEP-100, *M*=3.1, Z=3.09, *p*=0.002 and between the open ear and ComTac™ III, *M*=3.9, Z=2.80, *p*=0.005.

Table 186. Friedman test results demonstrating significant differences among listening conditions for the trained and untrained groups at pretest for Question 6, Likelihood of Wearing Device.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**9.95**	**12**	**2**	**0.007**
Untrained	**12.76**	**12**	**2**	**0.002**

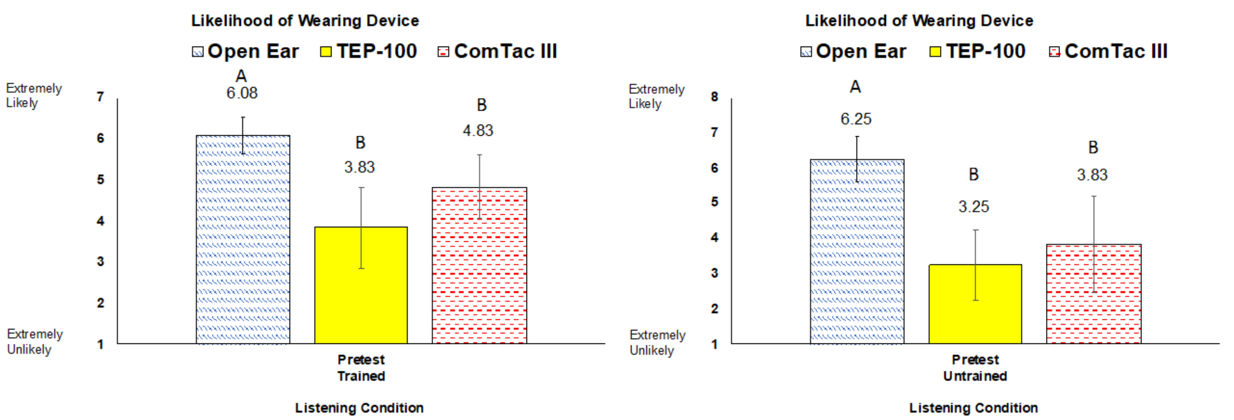


Figure 214. Mean ratings of likelihood of wearing device given similar conditions for each group and listening condition at pretest for Question 6, Likelihood of Wearing Device.

Table 187. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at pretest for the trained group, for Question 6, Likelihood of Wearing Device.

Listing Condition	*Z*	*p*
Open - TEP 100	**2.73**	**0.006**
Open - ComTac™ III	2.20	0.028
TEP-100 – ComTac™ III	1.24	0.217

Table 188. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at pretest for the untrained group, for Question 6, Likelihood of Wearing Device.

Listening Condition	*Z*	*p*
Open - TEP 100	**3.08**	**0.002**
Open - ComTac™ III	**2.41**	**0.016**
TEP-100 – ComTac™ III	0.85	0.397

Table 189. Friedman test results demonstrating significant differences among listening conditions for the trained and untrained groups at posttest, for Question 6, Likelihood of Wearing Device.

	χ^2	*n*	*df*	Asymp. Sig.
Trained	**15.45**	**12**	**2**	**0.002**
Untrained	**18.73**	**12**	**2**	**<0.000**

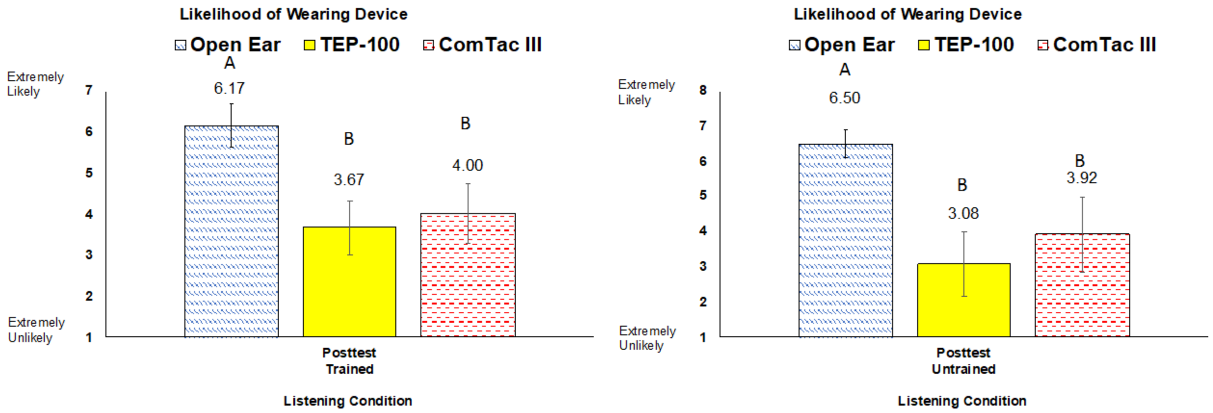


Figure 215. Mean ratings of likelihood of wearing device given similar conditions for each group and listening condition at posttest for Question 6, Likelihood of Wearing Device.

Table 190. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at posttest for the trained group (α=0.016) for Question 6, Likelihood of Wearing Device.

Listening Condition	*Z*	*p*
Open - TEP 100	**2.95**	**0.003**
Open - ComTac™ III	**2.70**	**0.007**
TEP-100 – ComTac™ III	0.50	0.618

Table 191. Wilcoxon signed rank tests, pairwise comparisons open, TEP-100, and ComTac™ III at posttest for the untrained group (α=0.016) for Question 6, Likelihood of Wearing Device.

Listening Condition	*Z*	*p*
Open - TEP 100	**3.09**	**0.002**
Open - ComTac™ III	**2.81**	**0.005**
TEP-100 – ComTac™ III	1.18	0.237

Question 7. Degree of Preparedness as Result of PALAT Training

Only the participants in the trained group were asked the following question at the conclusion of field testing for each listening condition: Rate the degree of preparedness you felt as a result of the training on the localization system (ring of loudspeakers) compared to the task of localizing .22 blank gunshot sounds from 1 (extremely unprepared) to 7 (extremely prepared). In order to compare ratings of the perceived degree of preparedness across listening conditions for the trained group in the field, a Friedman two-way analysis of variance by ranks was performed. Results showed a Chi-square value of 12.88, p=0.002 (Table 192). Follow-up testing for pairwise comparisons was conducted using Wilcoxon signed ranks tests. Testing yielded significant results for degree of preparedness for the open ear, M=5.9, compared to the TEP, M=4.5, Z=2.72, p=0.007, and for the open ear compared to the ComTac™ III, M=4.8, Z=2.27, p=0.010 (Table 193 and Figure 216). Results showed that the participants rated the open ear condition as the most likely they would use for a similar task and significantly more so than the TEP-100 or ComTac™ III condition.

Table 192. Friedman two-way analysis of variance results for listening condition for the trained group by ranks for Question 7, Degree of Preparedness.

χ^2	*n*	*df*	Asymp. Sig.
12.88	**12**	**2**	**0.002**

Table 193. Wilcoxon signed-ranks results comparing perceived degree of preparedness for each listening condition for Question 7, Degree of Preparedness.

Listening Condition	*Z*	*p*
Open ear - TEP-100	**2.72**	**0.007**
Open ear - ComTac™ III	**2.27**	**0.010**
TEP-100 - ComTac™ III	0.63	0.527

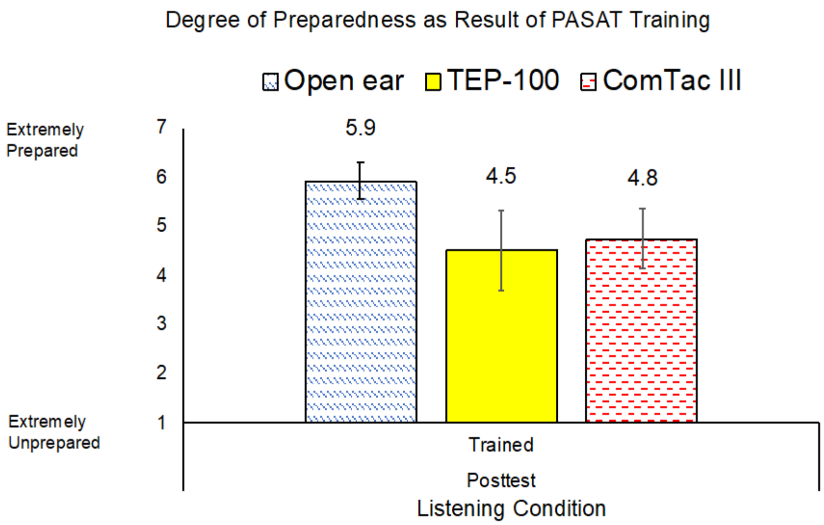


Figure 216. Plotted means and 95% confidence intervals for perceived degree of preparedness for each listening condition ratings at posttest for the trained group for Question 7, Degree of Preparedness.

4.7 Phase III Conclusions: In-Field Investigation of Transfer-of-Training

Overall, the transfer-of-training experiment supported the primary hypothesis that training in-office on the PALAT system transferred to in-field localization performance. The multivariate test only supported a significant main effect on the composite set of dependent variables for the listening condition variable. However, adjusting the alpha level for univariate testing to $\alpha=0.10$ enabled a deeper analysis of the effects of listening condition, stage of training, and group on absolute score. As such, univariate results supported that auditory training in-office resulted in better absolute correct localization of gunshots outdoors. Univariate testing also revealed a significant increase in response time from pretest to posttest suggesting that the in-field localization test was a more difficult task.

In addition to supporting that training significantly improved in-field performance, results supported that in-office training and in-field testing were sensitive to differences in listening conditions. Multivariate and univariate testing both showed significant effects for listening condition. The secondary hypothesis that in-lab performance could predict in-field performance under all three listening conditions was only supported with statistical significance for trained participants in the TEP-100 condition. However, comparing the slopes of the fitted regression lines for trained versus untrained participants for each listening condition showed significant differences in the open ear and TEP-100 conditions.

4.7.1 Listening Condition Conclusions

The open ear resulted in significantly higher performance measures of absolute correct score, ballpark correct score, response time and fewer front-back errors than in the TEP-100 and ComTac™ III conditions. These results were congruent with the in-lab, pretest findings of Casali and Robinette (2014) and Casali and Lee (2106a). The aforementioned studies reported that the open ear outperformed in-the-ear and over-the-ear hearing protection on the measure of absolute correct score. In-field, posttest results were also congruent with Talcott et al.'s (2012) field study where the open ear outperformed in-the-ear and over-the-ear HPDs on all accuracy and response time measures. Use of either TCAPS employed in this study resulted in at least a 50-percent degradation in absolute correct score when collapsed across training stage and training group. Therefore, results aligned with Abel (2008) and Bevis et al. (2014)'s qualitative evidence that reports of greatly reduced situation awareness hearing protection use among servicemembers. Training localization skills while using TCAPS may improve U.S. Military Service Members' confidence in their issued equipment, with strong likelihood of resulting in increased adoption rates.

4.7.2 Training Effect Conclusions

While the results of this and other studies quantified localization loss associated with TCAPS use, evidence supported that certain TCAPS and the open ear are susceptible to training effects. However, collapsed across all listening conditions, mean in-office performance was no different (*M*=46%) than mean absolute performance in the field (*M*=46%). Therefore, only certain listening conditions resulted in training transfer and only in the trained group. For the untrained group, compared to in-office performance, mean absolute performance in the field was, 13.8 % worse for the open ear, 2% worse for the TEP-100 and 1% better for the ComTac™ III. On the other hand, with in-office training, field performance was nearly equivalent (<1% better) for the open ear, 11% better for the TEP-100, and 2% worse for the ComTac™ III. Accuracy results illustrate that performance with the ComTac™ III was not susceptible to the effects of training administered in this study. However, in contrast, training effects, especially the significant results for the trained TEP-100 condition at posttest, suggest that training may improve performance and subsequent trust in device use. Radial plots (Figure 217) illustrate the training effects at each sound source location for each listening condition. A perfect score would result in a large circle along the 100% perimeter ring on the graph. The broader shape seen in the open ear and TEP-100 posttest plots for the trained group display the significant improvement with training compared to the untrained group.

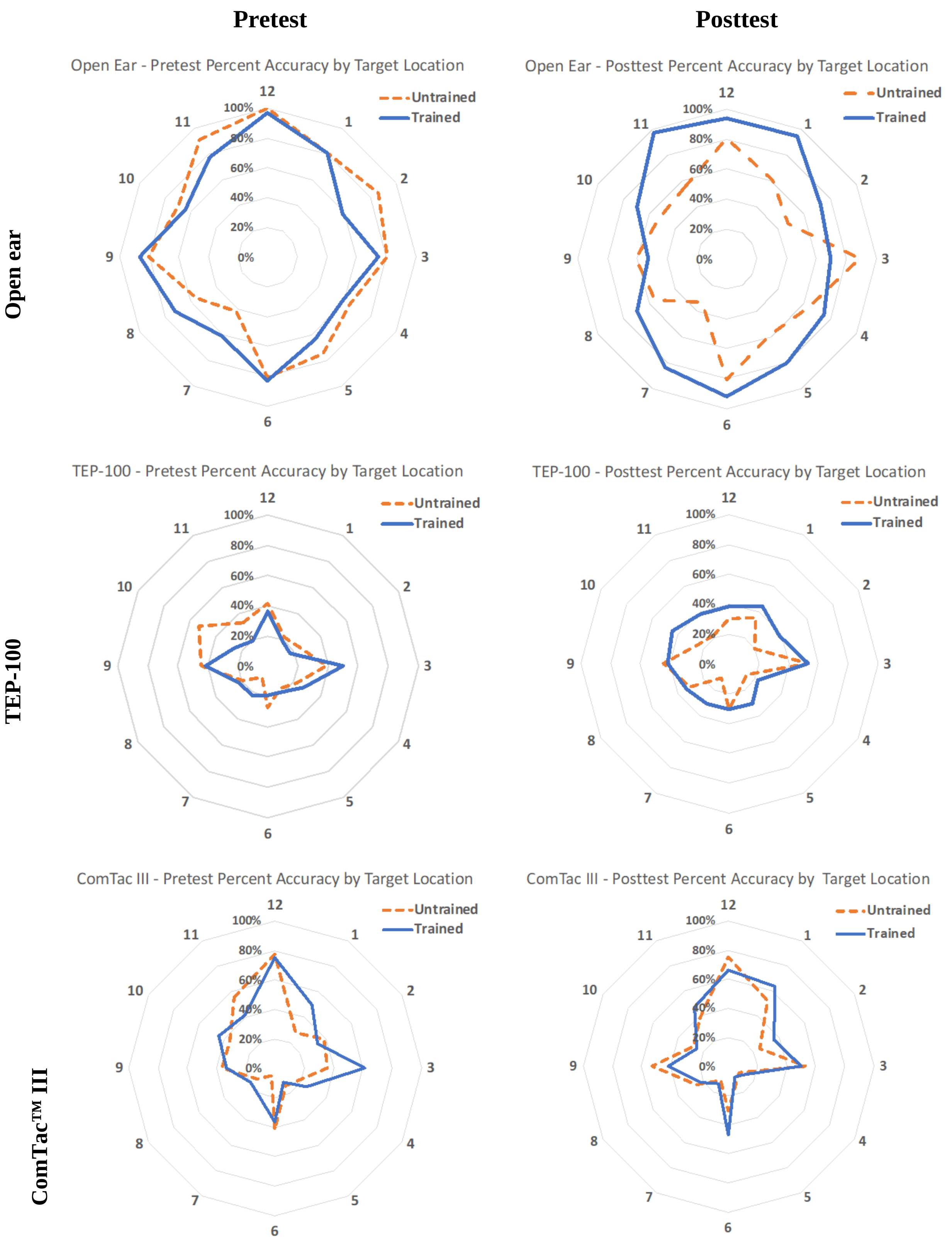


Figure 217. Radial plots of mean absolute correct accuracy percentage for each listening condition during pretest and posttest by group, trained (solid line) and untrained (dashed line).

Localization performance wearing TCAPS

This study was the first of its kind to train auditory localization ability in an office with loudspeakers, followed by testing in a field setting with live, blank gunshots. The TEP-100 listening condition was also unique to this investigation (to Phase II and Phase III of the broader research study). As a result, it is difficult to generalize the results of previous studies to the results of this investigation. However, in previous studies the in-the-ear style HPDs have consistently outperformed over-the-ear style HPDs in localization accuracy. Casali and Lee (2015) compared localization performance for several hearing protectors at unity gain given a 50 dBA dissonant signal amidst 40 dBA of pink noise. The EB15-LE BlastPLG®, an in-the-ear electronic hearing protector, resulted in approximately 46% localization accuracy compared to the ComTac™ III that resulted in approximately 36% accuracy (Casali & Lee, 2015). Performance in the open ear condition resulted in approximately 55% accuracy (Casali & Lee, 2015). The same listening conditions were tested using an 85 dBA signal and 75 dBA of pink noise. The open ear showed approximately 37% accuracy, compared to the 33% for the EB15-LE BlastPLG® and 26% for the ComTac™ III. Similarly, Talcott, Casali, Keady, and Killion (2012) evaluated localization accuracy for a .22 caliber blank stimulus, approximately 100-104 dB pSPL at the participant's ear, amidst 82 dBA of diesel truck noise and 45-50 dBA of rural noise. The investigators used the ballpark criterion to assess mean percent correct localization for the open ear, EB15-LE BlastPLG®, ComTac II over-the-ear electronic earmuff, and the EB1-BlastPLG® in-the-ear electronic hearing protector. Results for the open ear showed mean accuracy results of 88% and 81% for the rural and truck noise, respectively (Talcott et al., 2012). The EB1- BlastPLG® and EB15- BlastPLG® resulted in 61-63% (rural noise) and 59-64% (truck noise) compared to 53% (rural noise) and 43% (truck noise) for the ComTac™ II. Again

EB15- BlastPLG® resulted in 59-64% accuracy compared to 43-53% for the ComTac™ II (Talcott et al., 2012). Consistent with Phase II of this investigation, participants wearing the TEP-100 had, on average, achieved lower absolute accuracy compared to scores obtained while wearing the ComTac™ III. Training on the PALAT system was able to overcome this discrepancy, resulting in a significant difference in localization performance between the trained and untrained participants with the TEP-100 on the posttest. The TEP-100 also resulted in the highest mean absolute localization score for a TCAPS device at posttest. The significant training effect under the TEP-100 listening condition is promising for providing increased situation awareness while protecting Service Members hearing. However, as seen in the radial plots, additional training under the TEP-100 may be needed to achieve localization performance acceptable for most military operational needs.

A visual inspection of the radial plots under the ComTac™ III listening condition shows a concerning trend in poor performance from locations behind the participant, from 4 o'clock to 8 o'clock. Noticeably, the localization accuracy bias towards the frontal plane in the pretest condition using the ComTac™ III did not improve with training. Numerous participants voiced concerns during both the in-lab pretest and in-field posttest that all of stimulus signals seemed to be originating from in front of them while wearing the ComTac™ III. The investigator was unable to identify the primary contributing factor to the poor localization performance behind the listener with the ComTac™ III. However, it is hypothesized that the front-facing microphones of the ComTac™ III may be a contributing factor to its poor rear localization performance.

Front-Back Errors

The implications of the types of errors recorded in the ComTac™ III warrants further discussion. The ComTac™ III yielded significantly more front-back reversal errors than the open ear condition and slightly more front-back reversals than the TEP-100 condition. These results were consistent with previous study results where Talcott et al. (2012) found that earlier generation of the over-the-ear HPD (ComTac™ II) with similar forward-facing microphones resulted in a greater number of front-back reversals compared to the open ear and in-the-ear listening conditions (Etymotic EB1- BlastPLG® and EB15- BlastPLG®). As discussed previously, U.S. Military ground combat personnel are trained to orient in the direction they perceive the enemy threat signal. In a threat-detection scenario, localizing, and thus responding, to a perceived hazard in the opposite (i.e. wrong) plane is projected to have deleterious consequences. To mitigate the possibility of this occurring, a longer training session or training that specifically focuses on the dorsal plane may be needed to yield better localization performance.

Response Time

A significant main effect occurred according to listening condition on response time. Collapsed across training stage, both the TEP-100 and ComTac™ III conditions had significantly higher mean response times than the open ear condition. Mean response times for the TCAPS devices were 0.4 seconds slower during the pretest and 0.5 seconds slower during the posttest. In addition, while not significant, the mean response time in the field environment with gunshot stimuli was higher for every condition. Half a second in military operations could have a detrimental impact on the ability to respond or locate the enemy threat signal. Additional training may be necessary to improve response time while wearing TCAPS devices. Pollastek and Rayner (1998) explain that reaction time can be used to delineate components of mental processing. For example, the authors explained that when items are perceived as similar, reaction time increases, reflecting increased processing time. Therefore, reaction time can reflect time to identify an event and make a decision: key components in situational awareness. (Pollatsek & Rayner, 1998).

While longer response times are consistent with decreased automaticity, or a more difficult task, longer response times also may reflect a difference in task demands. In the field study, participants were instructed to speak his or her response in order to provide a back-up written record. The investigator noted that the single participant who produced outlier data on the measure of response time spoke his response before responding. Given that the task of speaking the response was not part of the in-office experiment, and thus an unfamiliar task, this extra step could have contributed to the response time considerably. Additionally, the experimental apparatus in the field was far less tolerant to deviation, which may have led to participants being far more tentative and deliberate in their responses. Specifically, if the participant triggered the LabView software before he or she was ready, or if the device misfired, the participants knew that no re-starts were allowed. In the lab, participants in the trained group went through many practice sessions after the pretests, whereas the field left no margin of error.

Signal Duration and Head Movement

Adding to the complexity of localizing the in-field gunshot stimulus versus the in-office dissonant signal was the shorter duration of the in-field stimulus. In the lab, the dissonant signal duration was 1000 msec. However, the duration of gunshots from a pistol is less than 50 msec (Maher, 2006). The direct ray of the gunshot, containing broadband information, contained both ILD and ITD information. While not directly measured, gunshots are widely accepted as occurring for less than ½ a second, rendering them an impulse noise hazard. Some reflections of the lower frequency energy may increase the duration of the blank gunshots in the field via reverberance. However, the overall shorter duration of the originating signal creates a more acoustically challenging stimulus to localize. Scharine and Letowski (2005) reported that head movements, which can be particularly helpful in sound localization, are primarily beneficial for sounds of greater than 400 to 500 msec. Therefore, while reflected gunshots sounds should have provided these additional localization cues, but probably not with the stimulus integrity of the directly-emitted one second duration, broadband dissonant sound in the lab. Therefore, the impulse noise of the in-field gunshots contained both ILD and ITD information, but the shorter duration precluded some head-turn benefit available in the in-office environment.

Muller and Bovet (1999) found that the pinna effects and head movement had a synergistic effect on localization accuracy in the horizontal plane. Removing either the pinna or head movement effect resulted in a 10% degradation in localization accuracy (Muller & Bovet, 1999). When just the pinna effects were reduced by filling the troughs of the pinnae with impression material, head movement displacements were larger. However, head movements did not fully compensate for the lost pinna effects (Muller & Bovet, 1999). The aforementioned study aligns with the results of the study. Specifically, at pre-test, the experimental and control groups in the open ear condition resulted in 73-75% absolute accuracy, whereas the ComTac™ III resulted in 33-36% absolute accuracy. While loss of pinna effects cannot entirely account for the degradation in performance recorded in the ComTac™ III pretest condition, the acoustical barrier the circumaural hearing protection causes is irrefutable. Similarly, in the field, the control group in the open condition showed a mean absolute accuracy of 62% whereas the ComTac™ III condition resulted in 34% absolute accuracy. Certainly, the signal processing strategy in the ComTac™ III and differences in stimuli spectra between the office and field environments can contribute to degraded performance. However, the deleterious effects on localization of the loss of pinnae filtering cannot be ruled out given the spectral shaping the pinnae provide.

Additionally, the 13% degradation in the open condition from pretest to the field for the control group aligns with loss of head turn cues. The TEP-100 only partially filled the concha and would have been expected to preserve at least some pinna-filtering cues. However, the effects of the device's signal processing, as previously discussed, are suspected to have interfered with this benefit. Absolute accuracy in the TEP-100 conditions were the same or worse than the ComTac™ III in the office and field environments for both groups. The TEP-100 showed poor absolute accuracy (29% for the control in the pretest and 28% for the experimental). Therefore, when participants had access to both pinna effects in the open ear condition and head movement of the lab, the open-ear in-lab resulted in the highest absolute localization scores (Muller & Bovet, 1999).

Signal Effect on Localization

Another challenge presented by the blank gunshot stimulus versus the dissonant signal was the narrower frequency spectrum in the blanks. The field stimuli generated signals with localizable spectral content from about 800 to 1500 Hz and from 3000-7000 Hz, as shown in Figure 88. The dissonant signal incorporated more low-frequency energy, particularly below 800 Hz, shown in Figure 153. Stevens and Newman (1936) found that frequency had a significant effect on error in horizontal localization, and front-back errors in particular. In their experiment, localization was most accurate below 1000 Hz, but stimuli between 2000 and 4000 Hz rendered the largest localization errors. In this study, the in-office signal incorporated energy at 104 and 295 Hz, spectral content not present in localizable levels in the in-field test. Therefore, the broader frequency signal content present in the in-office study rendered its signal easier to localize than the gunshot in the field, even though the gunshot did provide some ILD and ITD cues. Of course, in the interest of ing the training effect, the additional low frequency energy is appropriate for use in the PALAT system.

Inherent in over-the-ear (muff) hearing protection is the complete loss of pinna effect cues. Surprisingly, the TEP-100 creates less of an obstruction than the ComTac™ III, but the mean absolute performance for the TEP-100 was lower than the ComTac™ III. The investigator believes that the sound processing algorithm, the onset and ramp-up of the compression in the pass-through gain circuit, and/or the fidelity of processing and passing-through localization cues are the sources of lower performance in the TEP-100. The poor localization performance while using the TEP-100 was corroborated via the Phase II localization accuracy training and testing in two test facilities: the DRILCOM hemi-anechoic facility and the in-office PALAT system. The investigator acquired a frequency response curve chart depicting a representative TEP-100 compression curve from 3M™ in order to analyze the effects of compression onset and ramp-up on localization performance. The compression curve previously show in Figure 146 is shown again below in Figure 218 to facilitate results discussion from Phase III. Figure 218 confirms that the TEP-100 begins compressing the auditory signal at 60 dBA as indicated by the knee-point in the minimum volume line (red line) using a broadband pink noise signal (Stergar, Fackler, & Hamer, 2019). At 70 dBA, the sound pressure level used in both the in-office and in-field experiments, the gain circuit using the minimum volume or unity gain setting on the TEP-100 compresses the signal to an output level below the input level. This sound level reduction in a single ear may not have a significant impact on localization in isolation. However, each TEP-100 earpiece acts independently with no communication with the earpiece in the opposite ear resulting in varying degrees of compression for the arriving signal.

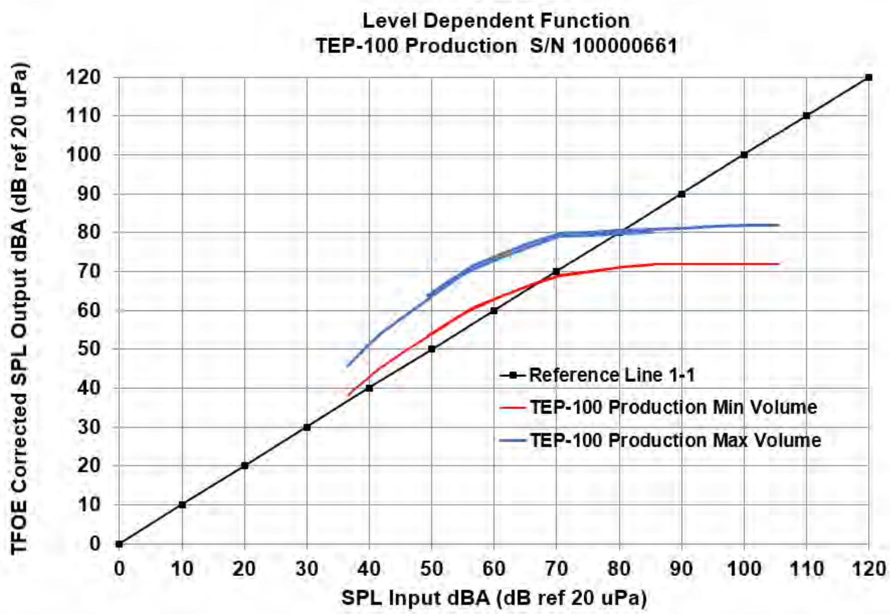


Figure 218. Level dependent function of the TEP-100 measured using a broadband pink noise signal; data are representative, but not considered exact for all samples, per manufacturer (Stergar et al., 2019).

As a result of the independent or unsynchronized compression functions, the head shadow effect which create the interaural level difference localization cues could result in the earpiece on one side of the participant’s head to compress the signal while the opposite earpiece fails to reach the compression threshold. Altering or eliminating interaural level differences would significantly degrade localization performance. To test this scenario, the investigator measured the sound pressure level of the dissonant tonal signal at both 55 dBA and 80 dBA from the 12 loudspeakers in the DRILCOM facility using the KEMAR manikin. Figure 219 displays the sound pressure levels of the dissonant signal at both 55 dBA and 80 dBA at 800 Hz, 5000 Hz, and 8000 Hz 1/3 octave-band frequencies as measured using a Larson Davis® measurement microphone located inside the ear canal of the KEMAR manikin under the open ear and TEP-100 listening conditions. The resulting impacts of the compression settings are illustrated by comparing the radial plots at each 1/3 octave-band frequency between the 55 dBA and 80 dBA signals. At 800, 5000 and 8000 Hz, the difference in sound pressure level recorded with the open ear and the TEP-100 is larger in the near ear (right side) when the signal level reaches the knee point in the compression circuit. This attenuation of the signal in the near ear under the TEP-100 indicates that both the dissonant tone and blank gunshot signals set at 70 dBA were impacted by the TEP-100 compression circuit during the localization testing. In addition, the 80 dBA signal radial plots display the impacts of the independent earpiece circuitry. When the 80 dBA signal originated from the near ear, the compression circuit within the TEP-100 earpiece reduced the signal at all 1/3 octave-band frequencies, indicated by the gap between the open ear SPL, dashed blue line, and TEP-100 SPL, solid orange line, from loudspeakers 12 o'clock to 5 o'clock. Whereas, when the 80 dBA signal originated from the far ear, the compression circuit with the TEP-100 earpiece maintained unity gain presenting a similar SPL as presented with the open ear listening condition, indicated by the overlapping lines of the open ear SPL, dashed blue line, and TEP-100 SPL, solid orange line, from loudspeakers 6 o'clock to 11 o'clock. More testing is needed to confirm the impacts of the independent compression circuitry with each TEP-100 earpiece (of a pair) on localization performance. However, preliminary results indicate the poor localization performance under the TEP-100 listening condition may in part be attributable to the sound processing algorithm, the onset and ramp-up of the compression in the pass-through gain circuit, and/or the fidelity of processing and passing-through localization cues.

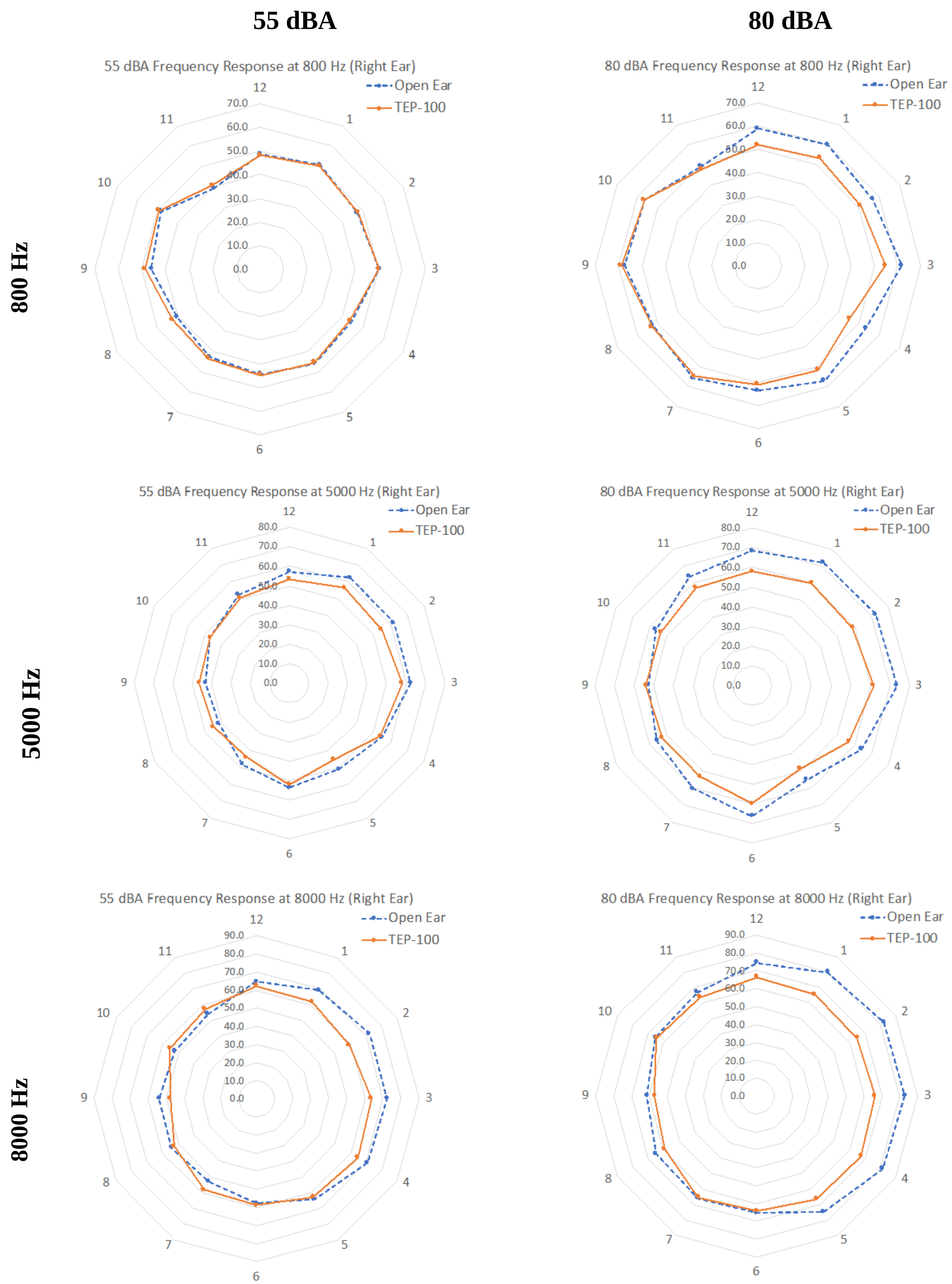


Figure 219. Frequency response radial plots of mean sound pressure level at a single 1/3 octave-band frequency. Measurements recorded using KEMAR manikin with open ear (dashed blue lines) and TEP-100 (solid orange line) in right ear. Dissonant tone set at 55 dBA (left column) and 80 dBA (right column).

Overall, those who received training demonstrated significant differences in performance compared to the untrained group for the open ear and TEP-100 conditions. Holding all other variables constant, the open ear condition resulted in the highest mean performance in the field. However, in the untrained, open ear condition, performance significantly decreased in the field as compared to the office. The trained group in the open ear condition improved, although not significantly. Therefore, results supported that the training overcame the loss in accuracy from the office to the field. The TEP-100 condition was the only condition that exhibited significantly improved performance from pretest in the office to posttest in the field. However, the *overall* mean performance for the TEP-100 condition was the worst compared to all other listening conditions in the office and the field, attesting to certain design issues that seem to exist with this TCAPS-Lite. The ComTac™ III was not susceptible to the transfer-of-training from the lab to the field.

While trends emerge for training effects in the field according to listening condition, the implications of these results remain unknown. Given that the degree of localization accuracy needed for certain duties remains to be defined, attaching meaning to the results of this experiment as to exactly what accuracy is advisable still presents a challenge. In other words, the degree of accuracy a service member requires to perform his or her tasks needs to be defined first, before any criteria for training performance is established. The localization capability may also vary considerably according to the mission or task. A scarcity of literature exists on this topic mainly because simulating poor localization ability is difficult to operationalize. However, Brungart and Sheffield (2016) were able to assess hearing loss and localization effects separately in simulated combat scenarios. A hearing loss simulator was used to either impose hearing loss or disrupt binaural cues. Participants with only a simulated hearing loss adopted compensatory strategies whereas those with degraded localization did not and performed considerably worse than those with hearing loss. Results were suggestive of those with localization loss not realizing the degradation of SA. Building on the work of Brungart and Sheffield (2016), hearing loss and localization loss simulators demonstrate potential in assessing localization requirements. In order to assess the effects of localization loss, the PALAT or DRILCOM apparatus could be used to quantify localization blur and the types of errors associated with the simulator settings. Soldiers could then be assessed while conducting standardized job-related tasks while using the simulator. From performance results, job-related performance could then be predicted given a certain degree of localization loss. The PALAT or DRILCOM system could then quantify localization blur with the use of various assigned HPDs or TCAPS. These products could then be used in the same job-related tasks in order to validate the localization apparatus results against real-world performance. If validated, then hearing protector performance given certain job-related tasks could be predicted and best matched to the service member's duty. While localization requirements are unknown, the results of this study provide data in order to render informed decision-making down to 30° accuracy when that data becomes available. In other words, the current study serves as the initial study of additional field-validated experiments needed to determine specific localization requirements for Service Members that must function in different missions and tasks.

Before training is implemented in any setting, the extent of training extinction should be quantified. The temporal and monetary costs associated with this system is significant as Service Members already have onerous annual training requirements. Even a one-hour annual training requirement can be burdensome to Service Members, especially commanders who are responsible for training compliance. Therefore, use of this training system should include long-term benefits. Kraus et al. (1995) demonstrated that auditory discrimination skills persisted one-month post-experiment via electrophysiological responses. Not only should training extinction be assessed, but electrophysiological measurements should be collected to further quantify the long-term impact, if any, of this training. In other words, evidence of neural pruning given localization training could better quantify the benefits of auditory localization training. It is also recognized that assessment of the extent of training extinction is a large undertaking, well beyond the scope of the current research, and will entail considerable follow-on longitudinal field data collection.

Questionnaire Rating Scales

Regarding *confidence ratings,* training did not result in a significant increase in confidence from pretest to posttest for each listening condition. However, a trend emerged where the open ear was associated with the highest confidence ratings, regardless of stage of training or training group, followed by the ComTac™ III, and the TEP-100. While the differences between pretest and posttest ratings within each group were negligible, differences in confidence ratings among devices were evident. The open ear condition was consistently and significantly associated with the higher confidence ratings than the ComTac™ III and TEP-100. These results also align with the findings of Abel (2008) and Bevis et al. (2014) that perceived loss of situation awareness served as a barrier to wearing hearing protection. Fundamentally, if a user does not have confidence in the key environmental cues transmitted through a device, the likelihood of compliance and HPD usage decreases. Unlike Brungart and Sheffield's (2016) study, participants had other listening conditions by which to compare performance and via feedback given, were aware of the degraded performance imposed by a particular device.

Regarding *perceived accuracy,* trends generally aligned with those of confidence responses. Both trained and untrained participants, regardless of training stage, rated the open ear condition as consistent with the most accuracy, followed by the ComTac™ III and the TEP-100. As in responses regarding confidence, accuracy ratings were not significantly different from pretest to posttest according to listening condition and group. Similar to the confidence ratings, no significant differences emerged comparing the trained versus untrained groups at each training stage. However, significant differences were present for the trained and untrained group at pretest, suggestive of the trained group expressing greater perceived accuracy before any treatment was rendered. This intragroup difference did not persist into the posttest results. Just as in the ratings of confidence question, ratings of perceived accuracy showed significant differences among listening conditions at pretest and posttest. In particular, the open condition resulted in higher ratings of accuracy compared to the TEP-100 and ComTac™ III conditions. Results support that perceived accuracy aligned with actual accuracy on the objective measure of absolute and ballpark accuracy. Of note, participants were informed of their localization accuracy scores prior to populating the questionnaires, and therefore, feedback regarding results may have driven these ratings.

For the question regarding *perceived difficulty,* with a higher rating indicating an easier task, responses followed the same trend as the questions regarding confidence and accuracy. As with the confidence and accuracy questionnaire items, significant differences were present among listening conditions at pretest and posttest for the trained and untrained groups. Listening with the open ear condition was rated as significantly less difficult than the TEP-100 and ComTac™ III. The open ear condition resulted in ratings consistent with the least difficulty, followed by the ComTac™ III, and the TEP-100 was associated with the most difficult localization. One noteworthy trend was the stability of ratings within each condition and group from pretest to posttest. In other words, perceived difficulty did not change even after training, but ratings were consistent with perceived confidence and accuracy. The consistent difficulty ratings from pretest to posttest could have a profound effect on Service Members. The untrained group showed significantly poorer performance on the posttest compared to the pretest for the open ear condition. However, there was no perceived increase in difficulty by the untrained group. This important result potentially suggests that Service Members are not able to perceive the true difficulty associated with localizing various sounds in different environments, obviously posing a dangerous predicament. The PALAT and DRILCOM system could serve as a tool to inform Service Members of their actual localization performance in varying environments, providing them a more accurate situation awareness picture.

On ratings of *perceived reaction time,* an unexpected result emerged in the trained group. Trained group participants rated their posttest reaction time as slower using the TEP-100 and ComTac™ III compared to those in the in the untrained group. Slower response times may have reflected emerging skill acquisition, meaning that participants were more deliberate in their selections given the training. While not significantly different, every listening condition and group rated slower reaction times in the field with the exception of the TEP-100 for the untrained group. Given the lowest accuracy scores associated with this TEP-100 untrained group, the unchanged response time may have reflected a lack of motivation. In other words, given the difficulty of the task, participants may have decided that more time invested into the task did not result in more accurate results. As with previously discussed questionnaire items, reaction time ratings were significantly different for the trained and untrained participants at pretest and posttest. At pretest, both trained and untrained participants showed significantly faster ratings for the open ear than for the TEP-100 or ComTac™ III. At posttest, the trained participants showed significantly faster ratings for the open ear than the TEP-100 or the ComTac™ III. For the untrained condition, only the TEP-100 was rated as significantly slower than the open ear.

Participants' ratings of *comfort* offer another insight into barriers to compliance with hearing protection. Participants consistently reported the open ear as the most comfortable listening condition at each stage of training and for each group. Unexpectedly, wearers of the ComTac™ III in the untrained condition reported significantly less comfort in the posttest compared to the pretest stage. These results would have been more likely in the trained group, in view that these participants wore the devices longer. Otherwise, no significant differences occurred when comparing pretest to posttest comfort ratings. At pretest, the trained and untrained conditions showed significantly less comfort in the TEP-100 condition compared to the open ear. In the untrained condition at pretest, the ComTac™ III was also found to be significantly more uncomfortable than the open ear. The same statistically-significant results for the pretest carried through to the posttest. The data from the comfort question suggest that training with the ComTac™ III reduces perceptions of discomfort. In general, drawing inferences about the comfort of in-the-ear and over-the-ear TCAPS is difficult given that only two devices were used in this study. However, these particular devices demonstrate clear differences between devices. Additionally, an acceptable level of comfort was not addressed to serve as a basis of comparison. In other words, a lower rating of comfort does not necessarily imply that the level of discomfort would lead to non-use of hearing protection.

For the question regarding *likelihood of maintaining the same listening condition given a task that required sound localization,* participants clearly preferred the open ear. This study incorporated immediate feedback due to the training component. Additionally, because of risk to human subjects, hazardous background noise was not included. As such, participants demonstrated a clear preference for the open ear in non-noise hazardous environments and with knowledge about their performance. Had participants not known how good or poor their localization performance was while wearing these devices, they may not have realized their performance decrement as Brungart and Sheffield's (2016) (Brungart & Sheffield, 2016)study suggests. Moreover, the intended use of TCAPS is to protect the user from the physiological and psychoacoustic adverse effects of noise. In other words, this study did not fully incorporate all of the environments where TCAPS devices would be worn and the results should not necessarily be generalized to noise hazardous environments. However, Service Members must operate in quiet environments where unexpected noise hazards can occur, and the results have direct application for that. As such, TCAPS devices should be able to adequately address the need to localize in quiet. While not significantly different, participants rated the open ear higher in the posttest than pretest for both the trained and untrained conditions. As with all other rating scales, significant differences were found at pretest and posttest for the trained and untrained users among listening conditions. For the trained and untrained group at pretest, the open ear condition was rated significantly higher than the TEP-100. In the untrained group, ratings for the open ear were also significantly higher than the ComTac™ III. At posttest, the same significant findings were also recorded for the open ear versus TEP-100 conditions and for the open ear versus ComTac™ III conditions. Results reflect a lesser likelihood of wearing the devices in a quiet field environment for localization, as compared to having an unprotected ear. Results should be interpreted with the understanding that this evaluation only queried the localization aspect of auditory perception, and thus likelihood of use ratings may have differed given different environmental conditions (i.e., in noise).

An additional question was administered only to the trained group in the posttest stage. The question queried the *level of preparedness* the participant felt as a result of the in-lab training. Results showed a significant difference among listening conditions, with the open ear condition resulting in ratings consistent with the highest level of preparedness. Significant differences existed between the open ear versus TEP-100 and the open ear versus ComTac™ III, with the TEP-100 receiving the lowest ratings of preparedness. The mean ratings across listening conditions were all greater than 3.5, the middle rating, suggestive of feeling more prepared than not. In lieu of the training verbiage and substituting with the experience of the pretest, administering this question to untrained personnel as well might have provided a more comprehensive understanding of feeling of preparedness. In other words, administering a similar question to the untrained personnel may have provided an adequate basis of comparison. Otherwise, considerable limitations exist for drawing inferences regarding this question using only the training group.

CHAPTER 5. Implications of the Results

5.1 Limitations of the Research

As previously stated, the study did not evaluate training extinction. While at least a day lapsed between the final office session and field testing, the validity of this scenario of one day between training and field localization actually occurring remains unknown. Furthermore, Service Members generally conduct duty-related training with their issued hearing protection prior to localizing sounds in a field environment or in a deployed setting. The current in-field study served as the first step to converging on a clearer understanding of mechanisms employed in sound localization. Future investigations into the training effect on in-field localization should consider the influence of real-world experience with altered localization cues imposed by hearing protection. Hofman et al., (1998) demonstrated that without feedback regarding performance, listeners with custom-molded impression-filled conchae adapted to the altered localization cues after three to six weeks. Thus, performance should be compared to those receiving formal in-office localization compared to those who only have experiential (i.e., on-the-job in-field) learning with the devices. While experiential learning may improve localization accuracy, it may also influence learning decay. Therefore, training extinction effects should be considered within the context of real world use of these devices.

Another limitation to the study was the validity of using a .22 caliber blank. This stimulus served as one of the few sound stimuli that was technically feasible given that live ammunition could not be used for obvious safety considerations. The quality control of the ammunition was previously discussed. The same brand of ammunition was used throughout the study, but that served as the only investigator-imposed means of quality control; but in any case, all blank shots were significantly suprathreshold and provided the opportunity for localization. In addition to quality control, the field study should be replicated using other types of real-world stimuli. Future investigations may incorporate higher caliber military weapon systems, whistle tube simulators to simulate incoming rocket propelled grenades, or mortar rounds that are relevant to ground-combat Service Members. The portable auditory localization acclimation training (PALAT) system has already been updated to incorporate a variety of military relevant sounds for this purpose.

The terrain of the field environment represents one possible outdoor localization scenario amidst an endless possibility of scenarios that can be encountered in military environments. One environment that varies considerably from the environment used in this study, but which is commonly encountered by U.S. Service Members, is urban operations (UO). Ground combat Service Members are required to conduct UO training. This environment presents many localization challenges due to reflections from buildings, ambient urban noise, and the presence of other non-target stimuli that can tax the listener during the detection/recognition phase of ASA. Clearly, even with its inherently more reflective environment, urban operations do require soldiers to localize sounds of various types, so its importance cannot be overlooked. Additionally, given the presence of buildings and subsequent threats from differing elevations in a UO environment, elevation cues remain paramount to threat localization. However, this experiment only addressed azimuth cues. In order to gain a broader understanding of how training impacts auditory localization, future studies should examine the impact in varied field environments. This study addressed one type of environment (arguably the most common for a field setting) and only employed azimuth cue training with its inherent limitation to the experiment's external validity. However, both the DRILCOM and PALAT systems incorporate elevation training capabilities and customizable training signals, thus they offer the capability for expanded environment investigations.

5.2 Results explained as a possible function of TCAPS design variables

The design of the ComTac™ III's forward-facing microphones most likely accounts for the frontal plane response bias discussed previously. In a forward-facing omnidirectional array, the microphones are most sensitive to sounds in front of the listener with null points at approximately 240° and 300°, or 4 o'clock and 7 o'clock, respectively (Dillon, 2001). The microphones on the ComTac™ III are also more distally mounted from the head than the ITE microphones of the TEP-100. This design conceivably offers greater sensitivity to environmental sounds due to proximity and larger microphone size of the ComTac™ III. Additionally, the passive sound isolation of the earcup may offer an advantage in the signal processing in the ComTac™ III. The ComTac™ III earcups offer greater sound isolation, creating a more robust barrier between the noise source and the entrance to the canal. On the other hand, the TEP-100 receives a higher level of input due to pinnae funneling and resonances. The effect of the higher input at the microphone in the TEP-100 may have been the reason for a more aggressive compression strategy that can distort or disrupt localization cues, particularly those from the sides of the head that are provided through interaural level differences and phase differences. Therefore, the closer proximity of the microphone to the source, larger surface area of the microphone, greater sound isolation of the earcup, and thus, less compression, may account for improved localization performance with the ComTac™ III compared to the TEP-100.

Another possible scenario is that the TEP-100's compression is triggered in the near-ear where the signal is louder (due to its relatively low threshold at compression onset), thus sending the near earplug into compression, rendering a lower level signal to the closer ear. Attenuating the closer signal disrupts the precedence effect where a closer signal is perceived as louder. Another alternative explanation for the poorer localization accuracy is gleaned from the input-output curves obtained in a hearing aid test box (Audioscan, Verifit) shown in Figure 220. In this measurement, the TEP-100 was set to unity gain and a pure tone frequency sweep was delivered at 60 and 90 dB SPL. These measurements were not in-situ measurements, which may better reflect how the outer ear transfer function would affect the frequency response of the output. The response curves show that the output varies considerably according to frequencies above 1600 Hz. For example, input at 1600 Hz is 68 dB for 60 dB input, but only 60 dB at 2000 Hz. Disparities at these frequencies given the same level could disrupt the ratio of frequencies employed in location perception. Thus, the poorer performance from the TEP-100 may be due to the smaller surface area of the microphones, the more aggressive compression strategy, an unequal compression algorithm applied at different frequencies, and an asymmetrical compression strategy with one the on-ear compressing first and applying more attenuation.

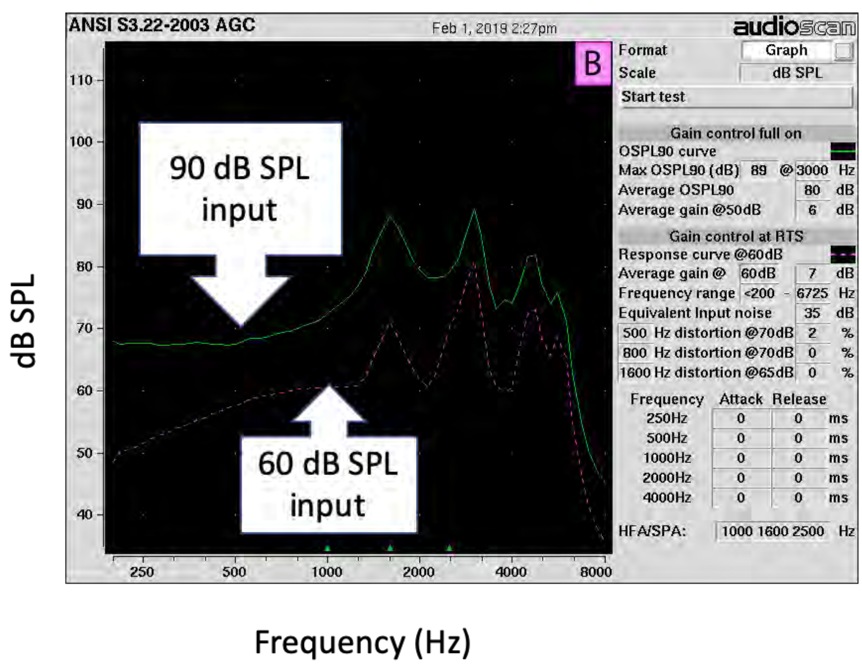


Figure 220. Input-output curve across frequencies using a pure-tone sweep at 60 and 90 dB SPL for the TEP-100 in the unity gain setting. Data were obtained with an Audioscan Verifit system.

5.3 Recommendations for Efficiency and Effectiveness in Training with the Portable Auditory Localization Acclimation Training (PALAT) System

5.3.1 Recommendations for use of the PALAT in various room environments

Portable localization training systems of the type developed and tested in this research also demonstrate promise as a tool to assess the adverse effects of hearing loss. Currently, auditory localization is not incorporated into hearing readiness metrics for U.S. Service Members. Therefore, duty limitations are not aligned with the handicapping effects of impaired localization associated with hearing loss. To assess the impact of differing degrees of hearing loss on auditory localization, the portable PALAT (or the larger DRILCOM system) could quantify localization blur given a certain hearing loss. Assuming field validation of the hearing loss simulator used in Brungart and Sheffield's (2016), one application is that experiments could assess the impact of varying degrees of simulated loss on duty-related tasks. Secondly, advantages can be derived from the fact that the PALAT is designed to fit within a standard examination room inside an audiology clinic for possible inclusion in hearing readiness metrics. Service Members with hearing loss could thus be assessed with PALAT in order to predict the functional impact of their localization accuracy as part of a standard test battery. The goal of these evaluations would be for clinicians and leaders to better align duties given a service member's hearing loss. Given that a standard clinical encounter in audiological setting lasts approximately one hour, localization evaluations could reasonably be assessed using the PALAT in the open ear and with a few TCAPS within this timeframe.

The PALAT system and training protocol may be implemented as means of training to counter auditory localization loss associated with loss of auditory accuracy. Undoubtedly, future research should address the effects of localization training in those with localization loss, either temporarily-imposed by hearing protection use or permanently from hearing loss. Localization remediation in hearing-impaired listeners through auditory training has yet to be thoroughly investigated. Furthermore, the possibility of restoring localization cues and the effect on localization accuracy through TCAPS has also yet to be evaluated. Therefore, the effects of localization loss remediation could be evaluated, and possibly treated, using the PALAT system and training protocol. The results of this study also support that training delivery of five learning units (LUs) can be administered within an hour, a timeframe that is currently the standard audiological evaluation time in most clinics. Also, training does not necessitate the presence of a clinician as the PALAT is intended to be executed independently by the listener, or trainee in this case.

5.3.2 Recommendations for use of the PALAT System as setup and deployed by a trainee

The PALAT system was specifically designed to be user-operated by a trainee with no prior experience with localization training. The current portable PALAT system has a diameter measurement of seven feet. It is recommended that trainees select a room that allows for one meter of space between the perimeter of the system and any reflective surface. This will reduce the likelihood of reflected sound rays interfering with the training and testing. If possible, the trainee should try to place the PALAT system in the center of the room and remove or spread out any furniture in the room along the walls as far as possible from the portable system. Use of curtains, closing window blinds, and carpeting can mitigate reflections. However, the investigator understands that space is often limited and that the PALAT system may need to be operated in less than ideal environments. As such, in this research, the PALAT system was deployed in a semi-reverberant room with multiple reflective surfaces, during both Phase II and Phase III. This room likely mimicked the types of rooms in which the PALAT system would be deployed in military base facilities, so the results are considered to be generalizable to those applications. In addition, the PALAT system was intentionally placed off-center within the room with several speakers located only two feet from the wall to mimic what was considered as "non-optimal" placement that would likely be established in actual military practice. As seen in the results of Phase II and Phase III, in spite of these training room non-optimalities, the PALAT system was still able to impart improvements in localization training from pretest to the last learning unit in all three listening conditions.

It is also recommended that the PALAT system be employed in a quiet room setting that is free from any continuous or frequent loud noise sources. Ideally, the ambient environmental noise should remain below the 55 dBA masking noise produced by the PALAT system. The trainee should test the ambient noise by turning on the pink noise or alternative masking noise source and listening to hear if any external noise sources can be heard. Training and testing sessions should be paused temporarily in the presence of any occasional loud noises. An additional feature that is being incorporated in the next software generation of the PALAT system will continuously measure the ambient noise and adjust the masking noise source and training signal to maintain the desired signal-to-noise ratio. This "automatic gain control" feature will need to be tested to identify the impacts on localization training.

5.3.3 Recommendations for use of the PALAT System given time likely available and time likely required

As previously discussed, this investigation did not measure the impacts of various training durations on localization performance. In addition, the military, or any other industry, has yet to define a standard for the minimum or desired level of localization performance. With that in mind, the training protocol developed in Phase I of the overarching investigation aimed to reduce the training time to a reasonable duration based on current U.S. service member training availability. The investigator equated localization training to rifle marksmanship training, based on their personal U.S. Army experiences. Typically, U.S. Service Members conduct basic rifle marksmanship at least twice a year and are provided with a few hours of training via simulator and practice ranges before being tested. The current training protocol employed during Phase II and Phase III was limited to three, one-hour training sessions. Participants were allowed to take a break at any point during their training. Based on results from Phase II and Phase III, one hour of training (5 LUs) for the open ear listening condition resulted in an asymptotic level of localization accuracy performance, with diminishing and negligible benefits of administering further LUs. Both the TEP-100 and ComTac™ III mean localization accuracy performance failed to reach the open ear performance levels with the same amount of training. As such, additional training may be needed if the goal is to achieve localization performance similar to the open ear while wearing certain TCAPS devices. Furthermore, more testing is needed to identify the frequency and duration of refresher training needed to maintain a desired level of localization performance.

Finally, the PALAT system can be used in a testing sense to determine when a TCAPS device places too high a training burden on users, and perhaps should be eliminated from consideration as a result. Toward this end, further research needs to be performed to determine at what point in the training process a TCAPS device should be eliminated if it requires an inordinate amount of training to bring the trainee up to a criterion level (or not at all). Casali and Lee (2016a), found, for instance, that one particular in-the-ear prototype TCAPS never did asymptote in its learning curve even after 12 LUs, and never reached 86% absolute correct performance of the open ear performance level -- thus, it was recommended that its development be discontinued.

5.4 Implications for military implementation

5.4.1 Relevance to TCAPS design, selection, and procurement

As with any hearing protector, the best protector is the one most appropriate for the duty and environment. Acoustically and historically speaking, the major determining factor in hearing protection selection is how much attenuation is needed to provide adequate defense of the ears in noise exposures encountered. Required attenuation and noise reduction ratings (NRRs) are generally known and routinely evaluated, and are required by the U.S. Environmental Protection Agency to be published on packaging for all HPDs sold in the U.S. (U.S. Environmental Protection Agency (EPA), 2002). The most important other determining factor in selection is the extent to which auditory components of situation awareness should be preserved to effectively perform job-related duties. Currently, how performance differs using a hearing protector compared to the open ear is mostly known, and that can be useful in communicating risk. For example, Table 6 demonstrated that given the effective range of a weapon, increasing localization blur, as is likely certain hearing protectors, can considerably increase the visual search area. However, the implications of the risks of degraded situation awareness cues can only be partially predicted, since they are different for every situation. Nonetheless, the lack of auditory situation awareness has been clearly evidenced as a causal factor in many accidents (Casali, 2019).

In the absence of a known need, the open ear's capability should serve as the gold standard in regards to assessing acceptable risk with TCAPS use from a products liability standpoint. However, this is most likely not necessary in all noise-hazardous situations, but it does ensure that the least amount of acceptable risk is adopted. Exceptions to where the open ear should not hold true as the standard is when Service Members operate in an enclosed environment. For example, those monitoring unmanned aerial vehicles may need to communciate with fellow operators, but do not need to localize particular threats in their environment, due to being enclosed in safe surroundings. Other examples of where hearing protection and communication are required, but not localization, are inside a tracked vehicle (e.g., military tank) or in an aviation setting. Weapons instructors at an outdoor range would not necessarily need open ear-equivalent localization capabilities, but would need to clearly hear incoming and outgoing communication. Therefore, from a liability perspective, quantifying localization loss, or any other degraded aspect of ASA, could better inform users of implied risk in using the device. Although not ideal, research can quantify the full scope of the effects of TCAPS use on ASA, and inform the user. Describing the risk and letting the stakeholder decide what device is most appropriate offers the best way to balance acceptable risk with safety requirements.

In the U.S. Military, Service Members do not typically have the opportunity to procure their own TCAPS. Commanders, acquisitions personnel, and occasionally clinicians render decisions regarding procurement of these devices. Employing a standardized testing system, especially a portable one, could better inform TCAPS stakeholders of the associated risks and benefits for these devices. Generally, one TCAPS is not well-suited for all scenarios where hearing protection would be needed. While the psychoacoustic needs may remain unknown, generally missions and duties have associated ASA requirements. Accordingly, a list of requirements generated by the end users and commanders thereof could be matched with capabilities requisite in an HPD or TCAPS to meet those requirements. As of December, 2019, a program of record exists for the Army for TCAPS, but the requirements are not disseminated.

5.4.2 Relevance to ground combat service member duties and mission

As policy, the U.S. Army does not release hearing loss metrics specific to units for security reasons. However, in 2016, hearing readiness metrics from an Army special operations unit showed signficantly lower incidence of hearing loss compared to conventional infantry units on the same installation (Klingseis, 2017). The difference in hearing thresholds between the two groups of units could not be accounted for due to differences to age or rank. One main difference between the two types of units was that in the special operations unit, ComTac™ III hearing protection was mounted on the helmet via rail attachment and integrated into the communication system. In other words, use of hearing protection was part of the standard personal protective equipment ensemble, and not a separate item. Conventional forces can use a variety of HPDs or TCAPS, but do not have such an ensemble requirement. Therefore, while a causal relationship between the use of the TCAPS integrated into standard equipment has not been established, audiometric data in this example and others supports a strong positive relationship between TCAPS use and hearing loss prevention.

Another implication of the finding of Klingseis (2017) is the importance of TCAPS not just restoring ASA cues and preventing hearing loss, but also providing U.S. Service Members with an operational, even tactical advantage. In other words, the goal of TCAPS should eventually be to achieve performance beyond that of the open ear and improving overall warfighter performance. For example, noise reduction algorithims integrated into hearing protection could enable improved speech transmission and understanding, more accurate localization, and improved threat detection. In addition to the psychoacoustic advantage afforded by improved designs, less noise exposure could also lead to positive implications of decreasing workload and fatigue. Gaining a strategic advantage through device use would certainly improve compliance, thus reducing hearing loss. While the ComTac™ III as evaluated herein may present challenges to certain aspects of ASA, compliance with wear of these devices is irrefutable, as demonstrated by Klingseis (2017). However, a current shortfall exists in communicating the associated risk of using a certain TCAPS devices to the user. Specifically, Service Members are not necessarily aware of the adverse, often lethal effects of not being able to localize well, especially when detection is improved with devices which provide amplification, such as the ComTac™ III. In other words, the service member perceives that they can hear better (the detection benefit) and thus views the TCAPS as a performance enhancer, not realizing that other aspects of hearing, such as localization, are compromised. While TCAPS offer a means to improve performance, in their current technology state they have associated risks that should be adequately communicated to the end users. Armed with this knowledge, stakeholders in military TCAPS programs could more accurately define the requirements for manufacturers.

5.4.3 Implications for NIHL reduction

Given the ever-present risk to U.S. Service Members of noise exposure due to training-related and unexpected exposures from hostile actions, compliance with hearing protection usage policies presents a unique set of challenges. The heightened risk of noise exposure is illustrated in the 30% greater likelihood of severe hearing loss in Service Members compared to non-veteran counterparts (Groenwold, Tak, & Matterson, 2011). While hearing protection is widely available to Service Members, compliance obviously lags the identified risk. All U.S. military personnel are required to undergo annual training that explains the risk of hearing loss and how to mitigate noise exposure. Despite these efforts, Service Members often choose not to wear hearing protection, but sometimes with good reason (Abel, 2008; Bevis et al. 2014). As the studies of Casali and Robinette (2014), Casali and Lee (2016a), Casali and Lee (2016b), Brown et al., (2015), Giguère et al, (2013) among many others, hearing protection use can degrade aspects of situation awareness. However, not all aspects of ASA are critical to each duty. Quantifying the risk inherent in each TCAPS application can assist Service Members in selecting devices that are best aligned with their operational needs, and will also help focus the pre-deployment training that may be advisable with a given product.

Not only does employing standardized testing of TCAPS devices improve device compatibility, but standardized training can better ensure confidence in associated TCAPS use. To overcome the tradeoff of choosing between sufficient protection or situation awareness, Service Members must have evidence-based confidence that TCAPS use will not compromise survivability or lethality. Training on aspects of ASA, including localization, serves as one method of instilling confidence. Establishing confidence in issued equipment is a common practice in the military. For instance, Service Members are required to test their gas masks in gas chambers to experience how the masks protects their breathing in the presence of CS (orto-chlorobenzylidene-malononitrile) gas. Likewise, Service Members test safety harnesses, parachutes, weapon systems, etc. to instill a sense of confidence in their equipment. However, TCAPS devices are typically stored in company supply rooms for accountability, and issued prior to training exercises or deployments without testing or training of the user. Training Service Members on the PALAT system while wearing their issued TCAPS device demonstrates strong potential as a means to improve localization performance and increase confidence in the fidelity of situation-awareness related cues. Conceivably, increased confidence would manifest as increased adoption rates of TCAPS devices and compliance, especially in noise-hazardous environments where detection and localization are critical. As the metrics illustrate, higher compliance with TCAPS use is associated with lower rates of hearing loss (Klingseis, 2017). Additionally, quanitfying the degradation to ASA with TCAPS use can assist manufacturers in better understanding the requirements that generate their designs, and improve future generations of their products. Standardized testing and training of TCAPS devices can improve device compatibility selection, confidence, and manufactuer design.

5.4.4 Cost-Benefit of Implementing the PALAT System for Training

The cost of implementing the PALAT system is relatively low compared to the potential benefits of auditory localization training. The full extent that the role of auditory localization performance imparts on military (or other) mission success has not been established. However, numerous studies, including the findings from Phases II and III of this investigation, have proven that TCAPS impede natural performance of auditory localization, increase the number of front-back reversal errors, and slow response times. The perceptibly-degraded auditory localization performance also causes some service members to forgo the use of TCAPS in hazardous noise environments in order to maintain auditory situation awareness, as a consequence of lost confidence. The PALAT system demonstrated the ability improve auditory localization after a fairly short training regimen. Participants also indicated, via survey responses, that they were able to perceive the benefits of training under both TCAPS conditions. The combination of increased performance and perception of improved auditory situation awareness could lead to increased adoption rates of TCAPS among service members. Increased adoption could help reduce the high prevalence rates of Noise Induced Hearing Loss (NIHL) among U.S. service members, as covered in detail in Chapter 1 of this dissertation.

The monetary costs associated with fielding the PALAT system to military units are relatively low given the ability of a user to operate the system without the need for an experimenter or a laboratory facility. As such, costs are limited to the production and fielding of the PALAT system and the opportunity cost associated with the time it takes to train auditory localization. A single PALAT system is estimated to cost $16,000 circa 2020. The investigator recommends semi-annual training for each service member under both the open ear and with assigned TCAPS device. While sufficient training is recommended to achieve a standard performance threshold as a criterion for “stopping,” the Phase II and III studies showed significant training benefits with only 1.5 hours of training per listening condition. As a result, the investigator concludes that the benefits of improved auditory localization performance and increased TCAPS adoption rates would far outweigh the cost of one day of auditory training every six months per service member and $16,000 per PALAT system.

5.5 Summary of Applications and Recommendations for Implementation of the PALAT System

The PALAT system was designed to fill an operational gap in the military and other industries where personnel are frequently exposed to hazardous noise sources but maintain auditory situation awareness. The three convergent studies in this investigation successfully (Phase I) developed an improved auditory localization training protocol, (Phase II) designed and validated a portable auditory localization training system capable of imparting similar training benefits as a full-scale, laboratory grade system, and (Phase III) demonstrated the transfer-of-training effect from a dissonant tonal complex stimulus trained in a semi-reverberant office to a real-world environment using military relevant gun shot stimulus. Based on the findings from the convergent studies, the following table summarizes the applications and recommendations for implementing the PALAT system (Table 194).

Table 194. PALAT System: Applications and Recommendations for Implementation.

1. Applications for PALAT System
   - A. Training with open ear and with HPDs or TCAPS, pre-deployment
     - Military operations require Service Members detect and locate hazardous threats accurately and in a timely manner. The PALAT system provides a practical and efficient training system to improve localization.
     - Service Members often received new equipment, including TCAPS, within a few months or weeks of deployments. The PALAT system provides the ability for user-driven training while wearing TCAPS in a variety of settings that are conducive to pre-deployment training, including training after the duty day in the barracks.
   - B. Determination of auditory fitness for duty, including localization acuity and blur
     - No current standard exists for testing or training localization within the U.S. Military.
     - The PALAT system provides a cost-effective apparatus that can be widely fielded to military units to collect normative data to help establish standards for both baseline proficiency and optimal performance.
     - Equipped with 24 azimuthal loudspeakers, localization acuity and blur can be measured before and after training to help develop standards.
     - The portability of the PALAT system provides a validated testing apparatus that can be operated in a deployed environment to screen for localization performance degradation following exposure to hazardous noise.
   - C. Re-training after hearing loss occurs
     - The PALAT system provides a means to establish a baseline localization performance score for each Service Member with open ear and their assigned TCAPS. The personalized performance scores can be used to screen reduced auditory situation awareness as a result of Noise-Induced Hearing Loss.
     - In the event of Noise-Induced Hearing Loss, temporary or permanent, the PALAT system provides a method to test and train localization performance to achieve baseline standards.
   - D. Confidence-building with HPDs or TCAPS
     - The PALAT system provides a medium to instill confidence within Service Members that their TCAPS device provides the requisite protection while maintaining auditory situation awareness.
     - The military requires testing of standard issued personal protective equipment including gas masks, hazmat suits, and weapons but has not established program to establish confidence in hearing protection devices.
   - E. Compliance testing and other aspects of vetting HPDs or TCAPS
     - The PALAT system demonstrated it is capable of detecting differences in auditory localization performance, both accuracy and response time, among TCAPS devices. This study also demonstrated the variability between users on localization performance under open ear and two TCAPS devices. The PALAT system provides an efficient manner to measure individual localization proficiency with assigned TCAPS devices.
2. Room Environment Considerations for PALAT System
   - A. Size and shape of training site
     - Ideally the room should offer one meter between the perimeter of the PALAT system (the back of the loudspeakers) and any reflective surface.

- Recommend using a room that offers minimum number of corners (square shape symmetrical room preferred).
- Place the PALAT system in the center of the room when possible.

B. Reverberance and reflective surfaces
- The PALAT system should be operated in a quiet room. Ideally the ambient noise level should be below 55 dBA and free of continuous noise or frequent loud noises.
- Close blinds, curtains, doors and drawers within the training room to minimize acoustic reflections.

C. Presence of furniture and fixtures
- Remove or spread out furniture along the perimeter walls of the room.
- If possible, cover large reflective surfaces with blankets or sound-absorbent material.
- Minimize the size of furniture used inside the loudspeaker array. Use a comfortable chair with a low back height to prevent blocking auditory signals presented behind the user. If a small desk or stand is used to hold the tablet, ensure the height of the tablet and stand are below the neck of the seated user.
- Do not block the loudspeakers. Ensure the user can clearly see all 24 loudspeakers when turning their head and neck.

D. Ambient noise considerations
- Place a sign outside of the training room notifying others that auditory training is being conducted.
- Temporarily pause the training or testing in the presence of occasional loud noises. The user-controlled PALAT system allows the user to initiate the auditory signal when the environment is conducive for training or testing.
- Ensure the masking noise is playing during all training and testing.

3. Training Protocol Considerations for PALAT System

A. Time available and time required for training
- Recommend training to achieve a criterion level established for each military occupational specialty. In the absence of such a standard, use a 1:2 ratio rule to train twice as long with a TCAPS device than with the open ear.
- One training session consisting of five learning units is recommended for the open ear.
- As demonstrated in Phase II of this study, training under TCAPS conditions may require two training sessions for a total of three hours of training and testing.
- Take short breaks between learning units and longer breaks between training sessions.

B. Training considerations for relevant operational tasks
- Train like you fight. Use your assigned TCAPS device during training. The Phase II study found variations between devices of the same TCAPS when measuring the unity gain sound pressure levels indicating that each device may have a different impact on localization performance.
- Train and test using a variety of military relevant stimuli to keep the training interesting and improve localization performance using different spectral content.

C. Frequency of required training
- Recommend semi-annual training and testing consistent with weapon qualification and gas mask training requirements.
- Conduct training and testing when assigned a new TCAPS devices or following any injuries to the auditory system.

The first stage of PALAT implementation requires a system usability study and exploratory examination of use in actual practice. The usability study will serve to identify how service members interact with the PALAT system and ways to improve the apparatus. The Virginia Tech Auditory Systems Laboratory personnel have requested that the exploratory examination of the PALAT system occur in four locations capable of allowing military service members to use the system under observation of researchers or perhaps audiology practitioners who are familiar with auditory experimentation and who are briefed on the PALAT system. The Phase II and III investigation involved participants with no prior experience with advanced hearing protection devices. The exploratory examination will also serve as a test to confirm that similar training benefits can be imparted on service members who have experience using TCAPS (or other augmented HPDs) and have operated in military environments where auditory situation awareness is critical for survivability. Testing the PALAT system with the intended end users will help to identify ways to improve the design of the PALAT system as well as confirm the training benefits that should occur in military service members.

Final Conclusions

The negative impacts on auditory situation awareness introduced by Hearing Protection Devices (HPDs) or Tactical Communications and Protective Systems (TCAPS) have been well documented through previous research and focus group interviews with military service members. As a result, the U.S. military identified the need for a portable system capable of imparting auditory localization acquisition skills at a similar level as the proven full-scale laboratory grade DRILCOM system which requires a large, hemi-anechoic room. A series of studies conducted at the Virginia Tech – Auditory Systems Laboratory demonstrated that a Portable Auditory Localization Acclimation Training (PALAT) system equipped with an improved training protocol was capable of replicating the auditory localization training benefits of DRILCOM in a semi-reverberant office environment. In addition, the study evidenced that training benefits from using the PALAT system with a dissonant tonal complex training signal could be transferred to an in-field environment using a military relevant signal of actual gunshots. Finally, the PALAT system was demonstrated to be capable of detecting differences in auditory localization performance between the open ear and with TCAPS devices and, more importantly, between two TCAPS devices.

The PALAT system was developed using mostly commercial off-the-shelf products to provide a portable system capable of being operated by a trainee in an office or barracks environment. Extensive research into design elements and audio components was conducted to select optimized components and build a system capable of reproducing auditory signals used in the DRILCOM test battery at low cost. A thorough Subject Matter Expert analysis using a human factors design selection algorithm unanimously recommended the Cambridge Audio Minx Min 12 loudspeaker, one of the most critical components of the PALAT system.

Participant absolute correct response scores were slightly lower on the PALAT system for all listening conditions during the pretest possibly indicating the portable system may be more challenging to initially localize sounds. However, training rates were consistent with (open ear and TEP-100) or better than (ComTac^{TM} III) the DRILCOM system for all listening conditions after five learning units totaling 1.5 hours for each listening condition. Multiple factors including the semi-reverberant room environment, visibility of the PALAT loudspeakers, increased number of loudspeakers (24 instead of 12), or potential near field effects of low frequencies may account for the slightly more challenging localization task while using the PALAT system. A slightly more difficult localization task could be an advantage for the PALAT in a training sense, because it may be a closer representation of localization tasks in real-world scenarios as demonstrated in Phase III (discussed in the following chapter). A more challenging task could also reduce the possibility of a ceiling effect. More testing is needed to identify which independent variables have the greatest effects on auditory localization. However, subjective ratings of both systems showed that participants were not able to perceive a difference in auditory localization performance between the two systems and preferred the more efficient PALAT system tablet user interface.

Over the course of two studies, the PALAT system demonstrated the ability to distinguish differences in auditory localization performance between listening conditions. The open ear condition significantly outperformed both TCAPS devices in absolute correct response and front-back reversal errors on both the PALAT and DRILCOM systems. The PALAT system also proved sensitive to detecting differences between TCAPS devices. Surprisingly, the in-the-ear TCAPS (TEP-100) was outperformed by the circumaural over-the-ear TCAPS (ComTac^{TM} III). Further testing of the in-the-ear TCAPS identified that the independent, or unsynchronized, compression processing algorithm in each ear piece reduced the sound pressure levels and spectral content of the signal on one side of the listeners head altering monaural and binaural localization cues, which may have contributed to its poorer localization accuracy results. Based on the totality of these findings, the PALAT system clearly offers a portable, less expensive option to test and screen future TCAPS devices based on how they impact auditory situation awareness.

The final in-field study demonstrated a transfer-of-training benefit from training on the PALAT system in an office environment with a broadband stimulus (dissonant tonal complex) to a real world in-field environment using actual blank gunshots. Open ear performance remained consistent between the pretest in-office and posttest in-field in the trained group but significantly declined for the untrained group. The trained group significantly outperformed the untrained group in the field after only five learning units of training with the in-the-ear TCAPS. This result evidenced the benefits of the PALAT system in instilling localization skills that transferred to the field environment. The over-the-ear TCAPS demonstrated no transfer of training effect to the field environment. Participants struggled most with signals originating from 4, 5, 7, and 8 o'clock positions indicating possible design impacts of the forward-facing microphones on the ComTac™ III. Additional studies are needed to test if training for longer periods could result in improved performance, better training transfer results, and/or how long the training benefits are retained in the actual operational environments. In the interim, these series of studies support that the Portable Auditory Localization Acclimation Training (PALAT) system, equipped with an improved training protocol, offers a feasible, beneficial, and low-cost system to efficiently train auditory localization under various listening conditions in a non-laboratory environment.

REFERENCES

3M. (2016a). *TEP 100 Tactical Earplug Brochure.* Retrieved from www.3M.com: https://multimedia.3m.com/mws/media/1001819O/tep-100-tactical-earplug-brochure-single-pgs.pdf

3M. (2016b). *3M PELTOR Tactical Comm and Hearing Protection Brchure 2016.* Retrieved from www.3M.com: https://multimedia.3m.com/mws/media/1417830O/3m-peltor-tactical-comm-and-hearing-protection-brchure-2016.pdf

Abel, S. (2008). Barriers to Hearing Conservation Programs in Combat Arms Occupations. *Aviation, Space, and Environmental Medicine, 79*(6), 591-598.

Abel, S. M., & Paik, J. S. (2004). The benefit of practice for sound localization without sight. *Applied Acoustics, 65*, 229-241.

Abel, S. M., Boyne, S., & Roesler-Mulroney, H. (2009). Sound localization with an army helmet worn in combination with an in-ear advanced communications system. *Noise & Health, 11*(45), 199-205.

Abel, S., Tsang, S., & Boyne, S. (2007). Sound localization with communications headsets: Comparison of passive and active systems. *Noise and Health, 9*(37), 101.

Abouchacra, K. S., & Letowski, T. (2012). *Localization of a Speech Target in Nondirectional and Directional Noise as a Function of Sensation Level.* Aberdeen Proving Ground, MD: Army Research Laboratory.

Abouchacra, K., & Letowski, T. (2001). Localization accuracy of speech in the presence of non-directional and directional noise maskers. *Paper presented at the 17th International Congress on Acoustics* (pp. 1-3). Rome, Italy: Sessions.

Ades, H. W., & Engstrom, H. (1974). Anatomy of the Inner Ear. In H. W. Ades, A. Axelsson, I. L. Baird, G. v. Bekesy, R. L. Boord, C. B. Campbell, . . . E. G. Wever, W. D. Keidel, & W. D. Neff (Eds.), *Auditory System: Anatomy Physiology (Ear)* (Vol. 5, p. 737). New York: Springer-Verlag Berlin Heidelberg.

AES 2-2012. (2012). *AES2-2012 - Standards for acoustics - Methods of measuring and specifying the performance of loudspeakers for professional applications.* New York: Audio Engineering Society, Inc.

Alali, K. (2011). *Azimuthal localization and detection of vehicular backup alarms under electronic and non-electronic hearing protection devices in noisy and quiet environments.* Blacksburg, VA: Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University.

Alali, K. A., & Casali, J. G. (2011). The challenge of localizing vehicle backup alarms: Effects of passive and electronic hearing protectors, ambient noise level, and backup alarm spectral content. *Noise & Health, 13*(51), 99-112.

Alali, K., & Casali, J. G. (2012). Auditory backup alarms: distance-at-firstdetection via in-situ experimentation on alarm design and hearing protection effects. *Work: A Journal of Prevention, Assessment, and Rehabilitation, 41*, 3599-3607.

Altshuler, M. W., & Comalli, P. E. (1976). Effect of stimulus intensity and frequency on median horizontal plane sound localization. *The Journal of Auditory Research, 15*, 262-265.

American National Standards Institute (ANSI). (1974). Hearing Protectors And Physical Attenuation Of Earmuffs, Method For Measurement Of Real-ear Protection Of,. *ANSI S3.19-1974*. New York: Acoustical Society of America.

American National Standards Institute (ANSI). (2010). Methods For The Measurement Of Insertion Loss Of Hearing Protection Devices In Continuous Or Impulsive Noise Using

Microphone-In-Real-Ear Or Acoustic Test Fixture Procedures. *ANSI S12.42-2010*. New York: Acoustical Society of America.

American National Standards Institute (ANSI). (2019). Methods for Measuring the Effect of Head-worn Devices on Directional Sound Localization in the Horizontal Plane. *ANSI/ASA S3.71-2019*. New York: Acoustical Society of America.

Arazi, S. C. (2017). *Antenna Fundamentals.* Retrieved October 2017, from MTI Wireless Edge LTD.: http://www.mtiwe.com/?CategoryID=353&ArticleID=163

Bauer, R. W., Matuzsa, J. L., Blackmer, R. F., & Glucksberg, S. (1966). Noise Localization after Unilateral Attenuation. *The Journal of the Acoustical Society of America, 40*(2), 441-444.

Behringer. (2012). Behritone C50A Studio Monitor. British Virgin Islands.

Berger, E. H. (1981). Re-examination of the low-frequency (50-1000 Hz) normal threshold of hearing in free and diffuse sound fields. *The Journal of the Acoustical Society of America, 70*(6), 1635-1645.

Bevis, Z., Semeraro, H., van Besouw, R., Rowan, D., Lineton, B., & Allsopp, A. (2014, March-April). Fit for the frontline? A focus group exploration of auditory tasks carried out by infantry and combat support personnel. *Noise and Health, 16*(69), 127.

Blauert, J. (1969/1970). Sound localization in the median plane. *Acoustica, 22*, 205-213.

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

Borwick, J. (2001). *Loudspeaker and headphone handbook* (3rd ed.). Boston: Focal Press.

Bose. (2017). *FreeSpace 3 Technical Data Sheet.* Retrieved January 10, 2018, from pro.Bose.com: https://pro.bose.com/en_us/products/loudspeakers/background_foreground/freeSpace-3-surface-mount-satellites.html#v=fs_3_system_black

Boston Acoustics. (2017). *SoundWare XS Satelite Technical Data Sheet.* Retrieved from bostonacoustics.com: http://www.bostonacoustics.com/DocumentMaster/US/lit_SoundWare_XS51.pdf

Brungart, D. S. (2014, March 10). *20Q: Auditory Fitness for Duty.* Retrieved July 26, 2017, from AudiologyOnline: http://www.audiologyonline.com/articles/20q-auditory-fitness-for-duty-12528

Brungart, D., & Sheffield, B. (2016). The operational impacts of hearing impairment: The important roles that psychology and context play in determining the performance of military members with degraded hearing. *Paper presented at the meeting of the National Hearing Conservation Association.* San Diego.

Butler, R. A. (1986). The bandwidth effect on monaural and binaural localization. *Hearing Research, 21*, 67-73.

Butler, R. A., & Musicant, A. D. (1993). Binaural localization: Influence of stimulus frequency and linkage to covert peak areas. *Hearing Research, 67*, 220-229.

Cambridge Audio. (2017). *Minx Min 12 Technical Data Sheet.* Retrieved from cambridgeaudio.com: https://www.cambridgeaudio.com/usa/en/products/minx

Casali, J. G. (2010a). Powered Electronic Augmentations in Hearing Protection Technology Cira 2010 including Active Noise Reduction, Electronically-Modulated Sound Transmission, and Tactical Communications Devices: Review of Design, Testing, and Research. *International Journal of Acoustics and Vibration, 15*(4), 45.

Casali, J. G. (2010b, December). Passive Augmentations in Hearing Protection Technology Circa 2010 including Flat-Attenuation, Passive Level-Dependent, Passive Wave Resonance, Passive Adjustable Attenuation, and Adjustable-Fit Devices: Review of

Design, Testing, and Research. *International Journal of Acoustics and Vibration, 15*(4), 187-195.

Casali, J. G. (2012a). Hearing protection devices: Regulation, current trends, and emerging technologies. In C. G. Refereed book chapter in LaPrell, D. Henderson, R. R. Fay, & A. N. Popper, *Noise-Induced Hearing Loss: Scientific Advances, (Handbook of Auditory Research Series)* (pp. 257-284). New York: Springer, Chapter 12.

Casali, J. G. (2012b). Sound and Noise: Measurement and Design Guidance. In G. Salvendy, *Handbook of Human Factors and Ergonomics* (Forth Edition ed., pp. 612-642). Hoboken, New Jersey: John Wiley & Sons, Inc.

Casali, J. G., & Gerges, S. N. (2006). Protection and enhancement of hearing in noise. In R. C. Williges, *Reviews of human factors and ergonomics* (Vol. 2, pp. 195-240). Santa Monica, CA: Human Factors and Ergonomics Society.

Casali, J. G., & Lee, K. (2015). *Objective metric-based assessments for efficient evaluation of auditory situation awareness characteristics of tactical communication and protective systems (TCAPS) and augmented hearing protective devices (HPDs) (Report No. W81XWH-13-C-0193).* Virginia Polytechnic Institute and State University, Industrial Systems Engineering. Blacksburg: Auditory Systems Laboratory.

Casali, J. G., & Lee, K. (2016a). *Objective Metric-Based Assessments for Efficient Evaluation of Auditory Situation Awareness Characteristics of Tactical Communications and Protective Systems (TCAPS) and Augmented Hearing Protective Devices (HPDs), Final Report.* Contract #W81XWH-13-C-0193, Department of Defense, Hearing Center of Excellence, Audio Lab 11/7/16-2-HP (final accepted date: 1/14/16, approved for public release by DTIC on 11/7/16).

Casali, J. G., & Lee, K. (2016b). An Objective, Efficient Auditory Situation Awareness Test Battery for Advanced Hearing Protectors and Tactical Communications and Protective Systems: DRILCOM (Detection-Recognition/Identification-Localization-Communication). *172nd Meeting of the Acoustical Society of America* (pp. 1-60). Honolulu, HI: Acoustical Society of America.

Casali, J. G., & Lee, K. (2019). Learning to localize a broadband tonal complex signal with advanced hearing protectors and TCAPS: the effectiveness of training on open-ear vs. device-occluded performance. *International Journal of Audiology, 58*(S1), S65-S73.

Casali, J. G., & Lee, K. (2020). What you don't hear can kill you – the Conundrum of Balancing Hearing Protection and Auditory Situation Awareness: Guidance for the Hearing Conservationist. *Presented at the National Hearing Conservation Association Conference.* Destin, FL: NHCA.

Casali, J. G., & Robinette, M. B. (2014, December). Effects of user training with electronically-modulated sound transmission hearing protectors and the open ear on horizontal localization ability. *International Journal of Audiology,* 1-9.

Casali, J. G., & Tufts, J. (in press). Auditory situation awareness and speech communications in noise. In D. e. Meinke, *The Noise Manual* (6th ed., p. Chapter 14). Fairfax, VA: American Industrial Hygiene Association.

Casali, J. G., Ahroon, W. A., & Lancaster, J. A. (2009). A field investigation of hearing protection and hearing enhancement in one device: For soldiers whose ears and lives depend upon it. *Noise & Health, 11*(42), 69-90.

Casali, J. G., Mauney, D. W., & Burks, J. A. (1995). Physical vs. psychophysical measurement of hearing protector attenuation - a.k.a. MIRE vs. REAT. *Sound and Vibration, 29*(7), 20-27.

Cave, K. M. (2019). *Evaluation of an Auditory Localization Training System for Use in Portable Configurations: Variables, Metrics and Protocol.* Blacksburg, VA: Virginia Polytechnic Institute and State University.

Cave, K. M., Thompson, B. S., Lee, K., & Casali, J. G. (2019). Optimisation of an auditory azimuth localisation training protocol for military service members. *International Journal of Audiology*, 1-9.

Chandler, D. W., Grantham, D. W., & Leek, M. R. (1993). Auditory spatial resolution in the horizontal plane as a function of reference angle: Microstructure of the azimuth function. *The Journal of the Acoustical Society of America, 93*, 2350-2351.

Clasing, J. E., & Casali, J. G. (2014). Warfighter auditory situation awareness: Effects of augmented hearing protection/ enhancement devices and TCAPS for military ground combat applications. *International Journal of Audiology, 53*, S43-S52.

Clifton, R. K., Clarkson, M. G., Gwiazda, J., Bauer, J. A., & Held, R. M. (1988). Growth in Head Size During Infancy: Implications for Sound Localization. *Developmental Psychology*, *24*(4), 477-483.

Coleman, P. D. (1963). An analysis of cues to auditory depth perception in free space. *Psychological Bulletin, 60*(3), 302-315.

Davis, R. J., & Stephens, S. D. (1974). The effects of intensity on the localization of different acoustical stimuli in the vertical plane. *Journal of Sound and Vibration, 35*(2), 223-229.

Defense Manpower Data Center (DMDC). (2017, July). *DoD Personnel, Workforce Reports & Publications.* Retrieved October 14, 2017, from DMDC: https://www.dmdc.osd.mil/appj/dwp/glossary.jsp

Defense Manpower Data Center (DMDC). (2018, July). *DoD Personnel, Workforce Reports & Publications.* Retrieved October 14, 2017, from DMDC: https://www.dmdc.osd.mil/appj/dwp/glossary.jsp

Department of Defense. (2015). *Department of Defense Design Criteria Standard Noise Limits MIL-STD-1474E.* Washington, DC.

Department of the Air Force. (2013). *Air Force Instruction 48-123.* Washington, DC.

Department of the Army. (2015). *DA Pamphlet 40-501: Army Hearing Program.* Washington, DC.

Department of the Army. (2017). *Ranger Handbook.* Alexandria, VA: Army Publishing Directorate.

DePass, D. (2017, May 18). *3M wins military hearing protection contract.* Retrieved from Star Tribune: www.startribune.com

Dillon, H. (2001). Advanced signal processing schemes for hearing aids. In H. Dillon, *Hearing aids.* Turramurra, New South Wales, Australia: Boomerang Press.

Dixon, W. J. (1951). Ratios Involvling Extreme Values. *The Annals of Mathematical Statistics*, 22(1), 68-78.

DOEHRS-DR. (2016). *Defense Occupational Health Readiness System Data Repository (DOEHRS-DR)*. Retrieved from U.S. Army Public Health Center: https://doehrswww.apgea.army.mil/doehrsdr

Donahue, A. M., & Ohlin, D. W. (1993). Noise and the Impairment of Hearing. In D. P. Deeter, & J. C. Gaydos, *Occupational Health: The Soldier and the Industrial Base* (Vol. Vol. 2, pp. 207-252). Washington DC: Office of the Surgeon General at TMM Publications Borden Institute.

Driscoll, D. P., & Royster, L. H. (2003). Noise Control Engineering. In E. H. Berger, L. H. Royster, J. D. Royster, D. P. Driscoll, & M. Layne, *The Noise Manual* (5th ed., pp. 279-378). Fairfax, VA: American Industrial Hygiene Association.

Duda, R. O., & Martens, W. L. (1998). Range dependence of the response of a spherical head model. *Journal of the Acoustical Society of America, 104*(5), 3048-3058.

Dufour, J., Ratelle, A., Leroux, T., & Gendron, M. (2005). Auditory localization training model: Teamwork between audiologist and O&M specialist—pre-test with a visually impaired person using bilateral cochlear implants. *International Congress Series, 1282*, 109-112.

Dunn, O. J., & Clark, V. A. (1987). *Applied Statistics: Analysis of Variance and Regression.* New York: John Wiley & Sons.

Eargle, J. (2003). *Loudspeaker Handbook* (2nd ed.). Boston: Kluwer Academic Publishers.

Emanuel, D. C., Maroonroge, S., & Letowski, T. R. (2009). Auditory function. In C. E. Rash, M. B. Russo, T. R. Letowski, & E. T. Schmeisser, *Helmet-mounted displays: Sensation, perception and cognition issues* (pp. 279-306). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Emanuel, D. C., Maroonroge, S., & Letowski, T. R. (2009). Auditory Function. In C. E. Rash, M. B. Russo, T. R. Letowski, & &. E. Schmeisser, *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues* (pp. 307-332). Fort Rucker, Alabama: U.S. Army Aeromedical Research Laboratory.

Ericson, M. A. (2000). Velocity Judgments of Moving Sounds in Virtual Acoustic Displays. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 44*, pp. 710-713 . San Diego: Sage Journals.

Fletcher, H., & Munson, W. A. (1933). Loudness, its definition, measurement and calculation. *Journal of the Acoustical Society of America, 5*, 82-108.

Gelfand, S. A. (2010). *Hearing: An Introduction to Psychological and Physiological Acoustics* (5th Edition ed.). London: Informa Healthcare.

Getzmann, S. (2003). A Comparison of the Contrast Effect in Sound Localization in the Horizontal and Vertical Planes. *Experimental Psychology, 50*(2), 131-141.

Getzmann, S., & Lewald, J. (2007). Localization of moving sound. *Perceptions & Psychophysics, 69*(6), 1022-1034.

Getzmann, S., Lewald, J., & Guski, R. (2004). Representational momentum in spatial hearing. *Perception, 33*, 591-599.

Giguere, C., Laroche, C., & Vaillancourt, V. (2013). Advanced hearing protection and auditory awareness in individuals with hearing loss. *Proceedings of Meetings on Acoustics. 19*, pp. 1-6. Montreal, Canada: Acoustical Society of America.

Grantham, D. W. (1986). Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *The Journal of the Acoustical Society of America, 79*(6), 1939-1949.

Groenwold, M. R., Tak, S., & Matterson, E. (2011). Severe hearing impairment among military veterans- United States, 2010. *Journal of the American Medical Association, 306*(11), 1192-1194.

Hajicek, J. J., Myrent, N., Li, Q., Barker, D., & Coyne, K. M. (2010). Protocols for Improved Understanding of Situational Awareness Effects of Head-Borne PPE. *2010 IEEE International Conference on Technologies for Homeland Security (HST)* (pp. 127-131). Waltham, Mass.: IEEE.

Hansen, C. H. (2001). Fundamentals of Acoustics. In B. Goelzer, C. H. Hansen, & G. A. Sehrndt, *Occupational exposure to noise: evaluation, prevention and control* (pp. 23-52). Dortmund, Germany: World Health Organization. Retrieved from University of Adelaide: http://www.portal.pmnch.org/occupational_health/publications/noise1.pdf

Hartmann, W. M. (1983). Localization of sound in rooms. *The Journal of the Acoustical Society of America, 74*(5), 1380-1391.

Heald, S. L., & Nusbaum, H. (2017). Understanding Sound: Auditory Skill Acquisition. *Psychology of Learning and Motivation, 67*, 1-42.

Hebrank, J., & Wright, D. (1975). The effect of stimulus intensity upon the localization of sound sources on the median plane. *Journal of Sound and Vibration, 38*(4), 498-500.

Held, R. (1955). Shifts in binaural localization after prolonged exposure to atypical combinations of stimuli. *The American Journal of Psychology, 68*(4), 526-548.

Hofman, P. M., & Van Opstal, A. J. (2003). Binaural weighting of pinna cues in human sound localization. *Experimental Brain Research, 148*(4), 458-470.

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

Humes, L. E., Joellenbeck, L. M., & Durch, J. S. (2005). *Noise and Military Service: Implications for Hearing Loss and Tinnitus.* Washington, D.C.: National Academies Press.

IEEE Std 219-1975. (1975). *IEEE Recommended Practice for Loudspeaker Measurements.* New York: The Institute of Electrical and Electronics Engineers, Inc.

Jokel, C., Yankaskas, K., & Robinette, M. B. (2019). Noise of military weapons, ground vehicles, planes and ships. *The Journal of the Acoustical Society of America, 146*((5)), 3832–3838.

Kapralos, B., Jenkin, M. R., & Milios, E. (2008). Virtual Audio Systems. *Presence, 17*(6), 527-549.

Keppel, G., & Wickens, T. D. (2004). *Design and Analysis A Researcher's Handbook.* Upper Saddle River, New Jersey: Pearson Education Inc.

Killion, M. C. (1978). Revised estimate of minimum audible pressure: Where is the ’’missing 6 dB’’? *The Journal of the Acoustical Society of America, 63*(5), 1501-1508.

Klingseis, K. H. (2017). Hearing loss in infantry soldiers at Fort Benning. *Army Public Health Course.* Fort Dix: Army Public Health Command.

Lee, K., & Casali, J. G. (2016). Effects of low speed wind on the recognition/ identification and pass-through communication tasks of auditory situation awareness afforded by military hearing protection/ enhancement devices and tactical communication and protective systems. *International Journal of Audiology, 55,* 1-9.

Lee, K., & Casali, J. G. (2017). Development of an audiorty situation awareness test battery for advanced hearing protectors and TCAPS: Detection subtest of DRILCOM (Detection-Recognition/Identification-Localization-Communication). *International Journal of Audiology, 56*, 22-33.

Letowski, T. R., & Letowski, S. T. (2012). *Auditory Spatial Perception: Auditory Localization ARL-TR-6016.* Aberdeen Proving Ground, MD: Army Research Laboratory.

Letowski, T. R., Scharine, A. A., Gaston, J. R., Amrein, B. E., & Ericson, M. A. (2012b). *The U.S. Army Research Laboratroy's Auditory Research for Dismounted Soldier: Present (2009-2011) and Future ARL-SR-239.* Aberdeen Proving Ground, MD: Army Research Laboratory.

Maher, R. C. (2006, April 4). *Summary of gunshot acoustics.* Retrieved from Montana State University: R Maher Publications: http://www.montana.edu/rmaher/publications/maher_aac_0406.pdf

Makous, J. C., & Middlebrooks, J. C. (1990). Two-dimensional sound localization by human listeners. *The Journal of the Acoustical Society of America, 87*(5), 2188-2200.

Maroonroge, S., Emanuel, D. C., & Letowski, T. R. (2009). Basic Anatomy of the Hearing System. In C. E. Rash, M. B. Russo, T. R. Letowski, & E. T. Schmeisser, *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues* (pp. 279-306). Fort Rucker, Alabama: U.S. Army Aeromedical Research Laboratory.

McBeath, M. K., & Neuhoff, J. G. (2002). The Doppler effect is not what you think it is: Dramatic pitch change due to dynamic intensity change. *Psychonomic Bulletin & Review, 9*(2), 306-313.

McIlwain, S. D., Gates, K., & Ciliax, D. (2008, December). Heritage of Army Audiology and the Road Ahead: The Army Hearing Program. *American Journal of Public Health, 98*(12), 2167-2172.

McMullen, K. A., & Wakefield, G. H. (2017). The effects of training on real-time localization of headphone-rendered, spatially processed sounds. *Proceedings of the Human Factors and Ergonomics Society 2017 Annual Meeting* (pp. 1557-1561). Austin, TX: Human Factors and Ergonomics Society.

Meister, D. (1985). *Behavioral Analysis and Measurement Methods.* New York: John Wiley & Sons.

Melzer, J., Scharine, A., & Amrein, B. E. (2012). Soldier auditory situation awareness: The effects of hearing protection, communications headsets, and headgear. In P. Savage-Knepshield, J. Lockett, & J. Martin, *Designing Soldier Systems: Issues in Human Factors, Chapter 9* (pp. 173-196). Ashgate.

Mershon, D. H., & King, E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics, 18*(6), 409-415.

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology, 42*(1), 135-159.

Mikaelian, H. M. (1969). Adaptation to rearranged ear-hand coordination. *Perceptual and Motor Skills, 28*, 147-150.

Milley, M. A. (2015, August 14). *CSA's speech at change of responsibility.* Retrieved July 30, 2017, from Army.mil: https://www.army.mil/article/154050

Mills, A. W. (1958). On the minimum audible angle. *The Journal of the Acoustical Society of America, 30*(4), 237-246.

Moller, H., & Pedersen, C. S. (2004, April-June). Hearing at low and infrasonic frequencies. *Noise Health, 6*(23), 37-57.

Moore, B. C. (1997). Space perception. In B. Moore, *An Introduction to the Psychology of Hearing.* San Diego, CA: Academic Press.

Moore, B. C. (2004). *An introduction to the psychology of hearing.* San Diego, CA: Elsevier Academic Press.

Muller, B. S., & Bovet, P. (1999). Role of pinnae and head movements in localizing pure tones. *Swiss Journal of Psychology, 58*(3), 170-179.

Mystic Marvels LLC. (2018, June). *PNG-400*. Retrieved from Mystic Marvels: www.mysticmarvels.com/png400.html

Navy and Marine Corps Public Health Center. (2008). *Technical Manual 6260.51.99-2.* Washington, DC: Department of the Navy.

Neuhoff, J. G. (2001). An Adaptive Bias in the Perception of Looming Auditory Motion. *Ecological Psychology, 13*(2), 87-110.

Newell, P., & Holland, K. (2007). *Loudspeakers: for music recording and reproduction.* Amsterdam;Boston: Elsevier/Focal.

Newton, V. E., & Hickson, F. S. (1981). Sound localization Part II: A clinical procedure. *The Journal of Laryngology and Otology, 95*, 41-48.

Noble, W., & Byrne, D. (1990a). A comparison of different hearing aid systems for sound localization in the horizontal and vertical planes. *British Journal of Audiology, 24*(5), 335-346.

Noble, W., & Byrne, D. (1991). Auditory localization under conditions of unilateral fitting of different hearing aid systems. *British Journal of Audiology, 25*(4), 237-250.

Noble, W., Byrne, D., & Lepage, B. (1994). Effects on sound localization of configuration and type of hearing impairment. *The Journal of the Acoustical Society of America, 95*(2), 992-1005.

Noble, W., Murray, N., & Waugh, R. (1990b). The effects of various hearing protectors on sound localization in the horizontal and vertical planes. *American Industrial Hygiene Association Journal, 51*(7), 370-377.

Oldfield, S. R., & Parker, S. P. (1984). Acuity of sound localization: a topography of auditory space. I. Normal hearing conditions. *Perception, 13*, 581-600.

Palca, J. (2016, June 3). *Army's smart earplug damps explosive noise, but can enhance whispers*. Retrieved July 21, 2017, from NPR: http://www.npr.org/sections/health-shots/2016/06/03/480173016/armys-smart-earplug-damps-explosive-noise-but-can-enhance-whispers

PEO Soldier. (2017). *PEO Soldier Portfolio FY17*. Retrieved July 26, 2017, from Program Executive Office Soldier: http://www.peosoldier.army.mil/portfolio/#187

Perrett, S., & Noble, W. (1995). Available response choices affect localization of sound. *Perception & Psychophysics, 57*(2), 150-158.

Perrott, D. R., & Musicant, A. D. (1977). Minimum auditory movement angle: Binaural localization of moving sound sources. *The Journal of the Acoustical Society of America, 62*(6), 1463-1466.

Perrott, D. R., & Saberi, K. (1990). Minimum audible angle thresholds for sources varying in both elevation and azimuth. *The Journal of the Acoustical Society of America, 87*(4), 1728-1731.

Peters, L. J., & Garinther, G. R. (1990). *The effects of speech intelligibility on crew performance in an M1A1 tank simulator.* Aberdeen Proving Grounds, Maryland: U.S. Army Human Engineering Laboratory.

Pickles, J. O. (1988). *An Introduction to the Physiology of Hearing.* San Diego, CA: Academic Press.

Pituch, K. A., & Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS.* New York, NY: Routledge.

Pollack, I., & Rose, M. (1954). Intensity discrimination thresholds under several psychophysical procedures. *Journal of the Acoustical Society of America, 26*, 1056-1059.

Pollatsek, A., & Rayner, K. (1998). Behavioral experimentation. In W. Betchel, & G. Graham, *A companion to cognitive science* (pp. 352-370). Malden: Blackwell.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research (4th ed.).* Upper Saddle River, NJ: Pearson Prentic Hall.

Price, G., Kalb, J. T., & Garinther, G. R. (1989). Toward a measure of auditory handicap in the Army. *Annals of Otology, Rhinology & Laryngology, 98*(5), 42-52.

Rakerd, B., & Hartmann, W. M. (1985). Localization of sound in rooms, II: The effects of a single reflecting surface. *The Journal of the American Society of America, 78*(2), 524-533.

Riesz, R. R. (1928). Differential intensity sensitivity of the ear for pure tones. *Physical Review, 31*, 867-875.

Rosenblum, L. D., Carello, C., & Pastore, R. E. (1987). Relative effectiveness of three stimulus variables for locating a moving sound source. *Perception, 16*, 175-186.

Rowland, R. C., & Tobias, J. V. (1967). *Interaural intensity difference limen AM 67-10.* Springfield, VA: Federal Aviation Administration.

Russell, G. (1977). Limits to behavioral compensation for auditory localization in earmuff listening conditions. *The Journal of the Acoustical Society of America, 61*(1), 219-220.

Sabin, A. T., Macpherson, E. A., & Middlebrooks, J. C. (2005). Human sound localization at near-threshold levels. *Hearing Research, 199*, 124-134.

Samsung. (2016, December 8). *The Auditory System - Overview*. Retrieved July 15, 2017, from Samsung Developers Conference: https://www.samsungdevcon.com/the-auditory-system-overview/

Scharine, A. A., & Letowski, T. R. (2005). *Factors Affecting Auditory Localization and Situational Awareness in the Urban Battlefield ARL-TR-3474.* Aberdeen Proving Ground, MD: Army Research Laboratory.

Scharine, A. A., Cave, K. D., & Letowski, T. R. (2009). Auditory Perception and Congnitive Performance. In C. E. Rash, M. B. Russo, T. R. Letowski, & E. T. Schmeisser, *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues* (pp. 391-490). Fort Rucker, Alabama: U.S. Army Aeromedical Research Laboratory.

Scharine, A., Mermagen, T., MacDonald, J., & Binseel, M. (2007). Effects of ear coverage and reflected sound on the localization of sound. *The Journal of the Acoustical Society of America, 121*(5), 3094-3094.

Scheaffer, R. L., & McClave, J. T. (1990). *Probability and Statistics for Engineers.* Boston: PWS-Kent Publishing Company.

Schechter, M. A., Fausti, S. A., Rappaport, B. Z., & Frey, R. H. (1986). Age categorization of high-frequency auditory threshold data. *The Journal of the Acoustical Society of America, 79*(3), 767-771.

Shaw, E. A. (1974a). The External Ear. In H. W. Ades, A. Axelsson, I. L. Baird, G. v. Bekesy, R. L. Boord, C. B. Campbell, . . . S. Rauch, E. G. Wever, W. D. Keidel, & W. D. Neff (Eds.), *Auditory System: Anatomy Physiology (Ear)* (Vol. 5, p. 737). New York: Springer-Verlag Berlin Heidelberg.

Shaw, E. A. (1974b). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *The Journal of the Acoustical Society of America, 56*, 1848-1974.

Sivian, L. J., & White, S. D. (1933). On minimum audible fields. *The Journal of the Acoustical Society of America*, 288-321.

Skvarenina, T. L. (2002). *The Power Electronics Handbook.* Boca Raton: CRC Press.

Steinberg, J. C., & Snow, W. B. (1934). Physical Factors. *Electrical Engineering, 13*(2), 245-258.

Stergar, M., Fackler, C., & Hamer, J. (2019). Correspondance with 3M Engineer. Indianapolis, IN.

Stevens, S. S., & Newman, E. B. (1936). The Localization of Actual Sources of Sound. *The American Journal of Psychology, 48*(2), 297-306.

Stewart Audio. (2018, June). *AV30MX-2*. Retrieved from Stewart Audio: www.stewartaudio.com

Talcott, K., Casali, J. G., Keady, J. P., & Killion, M. C. (2012). Azimuthal auditory localization of gunshots in a realistic field environment: Effects of open-ear versus hearing protection-enhancement devices (HPEDs), military vehicle noise, and hearing impairment. *International Journal of Audiology, 124*(1), S20-S30.

Thurlow, W. R., & Mergener, J. R. (1970). Effect of Stimulus Duration on Localization of Direction of Noise Stimuli. *Journal of Speech and Hearing Research, 13*(4), 826-838.

Tran, P. K., Amrein, B. E., & Letowski, T. R. (2009). Audio Helmet-Mounted Displays. In C. E. Rash, M. B. Russo, T. R. Letowski, & &. E. Schmeisser, *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues* (pp. 175-234). Fort Rucker, Alabama: U.S. Army Aeromedical Research Laboratory.

U.S. Army Public Health Center. (2017). *Hearing Readiness.* Retrieved July 30, 2017, from U.S. Army Medical Department: https://phc.amedd.army.mil/topics/workplacehealth/hrc/Pages/Hearing%20Readiness.aspx

U.S. Department of Defense. (2018). *2017 Demographics: Profile of the Military Community*. Alexandria, VA: Department of Defense (DoD), Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy.

U.S. Environmental Protection Agency (EPA). (1979). Noise Labeling Requirements for Hearing Protectors. *40CFR211, 44, 190*, 56130-56147. Federal Register.

U.S. Environmental Protection Agency (EPA). (2002). *Product noise labeling. Code of Federal Regulations, 40 CFR Part 211 [originally as 44 FR 56139, Sept. 28, 1979].*

United States Department of Veterans Affairs. (2006). *Annual Benefits Report Fiscal Year 2006.* Washington, D.C.: Veterans Benefit Administration.

United States Department of Veterans Affairs. (2016). *Annual Benefits Report Fiscal Year 2016*. Retrieved July 22, 2017, from http://www.benefits.va.gov/REPORTS/ abr/ABR-Compensation-FY16-0613017.pdf

USMC. (2017). *Enemy Threat Weapons.* United States Marine Corps: The Basic School. Camp Barrett, VA: Marine Corps Training Command.

Van Wanrooij, M. W., & Van Opstal, J. (2005). Relearning sound localization with a new ear. *The Journal of Neuroscience, 25*(22), 5413-5424.

Vause, N. L., & Grantham, W. (1999). Effects of Earplugs and Protective Headgear on Auditory Localization Ability in the Horizontal Plane. *Human Factors: The Journal of the Human Factors and Ergonomics, 41*(2), 282-294.

Viehweg, M. A., & Campbell, R. A. (1960). Localization difficulty in moaurally impaired listeners. *Transactions of the American Otological Society, 48*, 226-240.

Vliegen, J., & Van Opstal, A. J. (2004). The influence of duration and level on human sound localization. *The Journal of the Acoustical Society of America, 115*(4), 1705-1713.

Ward, W. D., Royster, L. H., & Royster, J. D. (2003). Anatomy and Physiology of the Ear: Normal and Damaged Hearing. In L. H. Royster, J. D. Royster, D. P. Driscoll, & M. Layne, *The Noise Manual* (5th Edition ed., pp. 101-121). Fairfax, Virginia: American Industrial Hygiene Association.

Westerberg, J. A., Balhorn, A. R., Tyshynsky, R. S., Olson, T. J., Brichetto, D. E., Gaston, M. L., & Loebach, J. L. (2016). *Novel Audiovisual Continuous Performance Task for Testing Working Memory and Decision-Making.* Retrieved August 4, 2017, from St. Olaf College: http://pages.stolaf.edu/cis-jwesterberg/files/2016/02/MUPC-AVCPT.pdf

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2013). *Engineering Psychology and Human Performance* (4th ed.). Boston: Pearson.

Wickens, C. D., Lee, J., Liu, Y., & Becker, S. G. (2004). *An Introduction to Human Factors Engineering.* Upper Saddle River, New Jersey: Pearson Education Inc.

Wolfle, D. (1946). Military training and the useful parts of learning theory. *Journal of consulting psychology, 10*(2), 73-75.

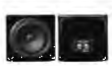
Wright, B. A., & Zhang, Y. (2006). A review of learning with normal and altered sound-localization cues in human adults. *International Journal of Audiology, 45*(1), S92-S98.

Zahorik, P., Brungart, D. S., & Bronkhorst, A. W. (2005). Auditory distance perception in humans: A summary of past and present research. *Acta Acustica united with Acustica, 91*, 409-420.

Appendix A. Review of Auditory Localization Apparatus Designs

Study	Setting	Shape	Dimensions	# of Speakers Type (if indicated)	Horizontal Displacement Angle (degrees)	Noise Generator	Amplifier
Abel & Paik, 2004 Abel et al., 2007 Abel et al., 2009	double-walled semi-reverberant sound proof	circular horizontal	~6 feet diameter	8 speakers (Minimus 3.5; Radio Shack Corp,	15, 75, 105, 165, 195, 255, 285, 345 degrees	Type 1405; Bruel & Kjaer Instruments,	Realistic SA-150; Radio Shack Corp, Fort Worth, Tx 2x Sony STR DE-135
Alali & Casali, 2011	hemi-anechoic room	circular horizontal	12 feet	4 speakers (Infinity SM-155)	45 degrees	Atlas Soundolier, Model GPN-1200A	2x Phoenix Gold VSS2 speaker selectors 2x AudioControl C-
Casali & Lee, 2016a Casali & Lee, 2016b Casali & Robinette, 2014 Clasing & Casali, 2014 Lee & Casali, 2016 Lee & Casali, 2017	hemi-anechoic room	circular horizontal	12 feet	12 speakers (Behritone C50A)	30 degrees (azimuthal)	QSC CX1102	
Makous & Middlebrooks, 1990	sound-attenuated room	circular horizontal	~6 feet diameter	36 piezoelectric speakers (Motorola model KSN 1072A)	10 degrees	IBM PC/AT computer	
Nawaz 2014	anechoic chamber	circular horizontal		20 speakers	18 degrees		
Dufour et al., 2005	sound booth	semi-circular horizontal		11 speakers	18 degrees		
Butler, 1986	sound-attenuated room	semi-circular horizontal	4.5 feet from listener	21 speakers (Braun, Output, Compact/C)	7.5 degrees (from 195 to 345 degrees on the left side)		
Noble et al., 1990 Noble et al., 1994	anechoic chamber	semi-circular horizontal and vertical	~6 feet diameter	20 speakers (11 horizontal)	18 degrees (from 0 to 180 degrees azimuthal)		
Vause & Grantham, 1999	anechoic chamber	semi-circular horizontal	~6 feet diameter	20 speakers	8 degrees		
Westerberg et al., 2016	classroom	semi-circular horizontal	~8 feet diameter	37 speakers	5 degrees (-90 to 90 degrees)		
Abouchacra & Letowski, 2001 Abouchacra & Letowski, 2012	anechoic chamber	sound boom		6 speakers (Bose 108515K)	45 degrees (although boom could present any angle)		
Oldfield & Parker, 1984	anechoic chamber	sound boom	~6 feet diameter	2 speakers (5 cm T52 KEF)	any angle	Bruel & Kjaer 1405	Harman/Kardon HK 505
Butler, 1993	sound-attenuated room	hemispheric surface	4.5 feet from listener	100 tweeters (Realistic, Model 40-1289)	15 degrees (from 0 to 180 degrees azimuthal)		
Van Wanrooij & Van Opstal, 2005	dark, sound-attenuated room	hemispheric surface	100 cm away from listener	58 small broadband speakers (MSP-30; Monacor International, Bremen, Germany)	20 (from 0 to +/- 60 degrees)		
Vliegen & Van Opstal, 2004	dark, sound-attenuated room	hemispheric surface	100 cm away from listener	58 small broadband speakers (MSP-30; Monacor International, Bremen, Germany)	20 (from 0 to +/- 60 degrees)		
Rakerd & Hartmann,	anechoic	arc	10 feet	8 speakers	3 degrees		
Perrett & Noble, 1995	semi-anechoic chamber	quadrant	~6 feet diameter	6 speakers horizontal (Realistic Midrange Tweeters)	15 degrees (horizontal quadrant)		
Getzmann, 2003	sound-attenuated room	crosswise (horizontal and vertical wall in front of listener)	~3 feet in front	9 speakers (Soundcraft CX-320, 7 cm diameter) (5 horizontal)	-20, -7.5, 0, 7.5, 20 degrees horizontal	CoolEdit96 at a 16-bit resolution	Uher stereo-amplifier UMA-2000
Sabin et al., 2005	anechoic chamber	motorized arc	~6 feet diameter	2 speakers (1x horizontal and 1x	any angle -90 to 90 degrees		Adcom GFA-535II
Thurlow & Mergener, 1970	anechoic chamber	front and side walls	6 feet diameter	14 speakers (7 low frequency speakers, University C&W) (7 high frequency speakers,		Grason-Stadler (model 455)	

Appendix B. Loudspeaker Alternative Screening

Loudspeaker Alternatives	Picture	Driver diameter	Size	Weight	Driver type	Power source	Terminal type	Sensitivity SPL	Frequency range	Price
Bosch LP1-UC10E-1 Sound projector uni-directional		4 in	7.3 x 11.8 in	6.6 lb	single cone	15 W	3 pole screw	86 dB	75 - 20000	$105
Bosch LBC3941/11 Sound projector		4 in	6.5 x 7.87 in	3.3 lb	single cone	9 W	2m 5 wire cable	96 dB	130 - 18000	$65
MM-4XPD Directional Miniture Self-Powered Loudspeaker		4 in	4 x 4 in	5.2 lb	single cone	self powered, DC power	spring clip	113 dB	120 - 18000	$1,800
Bose Acoustimass 6 series V Virtually Invisible series II speakers		2.5 in	3.3 x 3.2 x 3.7 in	1.2 lbs	single cone	12 W	Bare wire, spring clip	84 dB	170 - 20000	$ 600 for 6 and subwoofer
Bose Virtually Invisible 300 wireless surround speakers		2.5 in	3.3 x 3.2 x 3.7 in	1.2 lbs	single cone	12 W	Bare wire, spring clip	84 dB	170 - 20000	$ 150 (2 for $300)
Bose FreeSpace 3 Satelite		2.5 in	3.0 x 3.0 x 4.0 in	1.8 lbs	single cone	12 W	Bare wire, spring clip	84 dB	170 - 20000	$ 150 (2 for $300)
REI outdoor tech buckshot 2.0 speaker		1 in	3.6 x 1.87 x 1.53 in	0.5 lbs	single cone	lithium ion battery	Bluetooth			
Theater solutions TS4ODB Indoor or Ourdoor speakers		4.25 & 1 in	7.75 x 5.75 in	7 lbs	single cone	10 - 150 W	Bare wire, spring clip	90 dB	80 - 20000	$275 for 4 speakers
Theater solutions TS38W		3 in	6.5 X 4 in	>4 lbs	single cone	10 - 150 W	Bare wire, spring clip	90 dB	120 - 20000	$100 for 2
Acoustic Audio AA051B		3 in	4.6 x 4.3	>4 lbs	single cone	20 - 480 W	Bare wire, spring clip	89 dB	80 - 20000	$100 for 2
JBL Control 1 Pro (Pair)		5.25 in	9 x 6 in	4 lbs	two way, ported	150 W	Bare wire, spring clip	87 dB 114 pe	80 - 20000	$165 for 2
MixCube Passive		5 in	6.5 x 6.5 x 6.5 in	7 lbs	single cone	50 - 200 W	2-wire, screw	98 dB	90 - 17000	$160 for 2
MixCube Active		5 in	6.5 x 6.5 x 6.5 in	7 lbs	single cone	60 W	XLR & TRS combo jack	104 dB	90 - 17000	$250
A23 Bookshelf speaker		3.5 in & 1 ir	8.5 x 5.5 x 5 in	4.2 lbs	2 way	10 - 150 W	2-wire, screw	86 dB	80 - 25000	$150
Boston Acoustics SoundWare		4.5 in & .75	6.5 x 6.5 x 6.5 in	4.9 lbs	2 way	15 - 100 W	2-wire, screw	87 dB	90 - 20000	$100

Loudspeaker Alternatives	Picture	Driver diameter	Size	Weight	Driver type	Power source	Terminal type	Sensitivity SPL	Frequency range	Price
Boston Acoustics SoundWare XS Satelite		2.5 in & .5 in	4.25 x 4.5 x 3.75 in	1.0 lbs	2 way, coaxial	10 - 100 W	2-wire, screw	85 dB	150 - 20000	$76
G One Active Speaker		3 in & .75 in	7.7 x 4.75 x 4.5 in	3.7 lbs	2 way	25 W	1x RCA analog	96 dB	67 - 25000	$400
Cambridge Audio Minx Min 12		2.25 in	3.1 x 3.1 x 3.3 in	0.95 lbs	Flat panel BMR single	25 - 200 W	5 way binding post	86 dB	150 - 20000	$75
Kipsch G-12		3.5 in & 1 in	12 x 6 x 2.4	3.75 lbs	2 way	50 - 200 W	2-wire, screw	91 dB	69 - 24000	$200
Behringer 1CBK		5.5 in & .5 in	13.8 x 7.6 x 10.7 in	8.8 lbs	2 way	25 - 100 W	2-wire, screw	86 dB	60 - 23000	$80 for 2

Appendix C. Subject Matter Expert Questionnaire

Subject Matter Expert PALAT System Loudspeaker Analysis

You are being asked based on your expertise to evaluate loudspeaker criteria and alternatives for the Portable Auditory Localization Acclimation Training (PALAT) system, intended for training military service members to localize sounds in the horizontal plane (i.e., azimuth) by ear in their environment.

System Requirements

Loudspeakers are the most critical component of the PALAT system. The PALAT system will consist of a circular array of 12 directional loudspeakers separated by 30° increments horizontally. An overhead schematic of the basic design of the layout appears in Figure 1. The loudspeakers selected must be capable of replicating auditory training results obtained with the full-scale, laboratory grade DRILCOM system while meeting the systems requirements to include being no larger than 7 feet in diameter, storing in a small shipping trunk that can be hand-carried, being user-controlled by a laptop PC, and enabling military service members to assemble, conduct training, and disassemble in various training environments. In addition, the PALAT system loudspeakers must be able to accurately reproduce auditory signals across a major portion of the audible frequency spectrum, i.e., at least from about 250 to 10000 Hz, with a flat frequency response at an overall sound pressure level (SPL) of at least 85 dBA. As a result, it is important that the loudspeaker design parameters and engineering specifications be carefully evaluated when choosing the PALAT loudspeakers to maximize acoustical performance and meet system requirements.

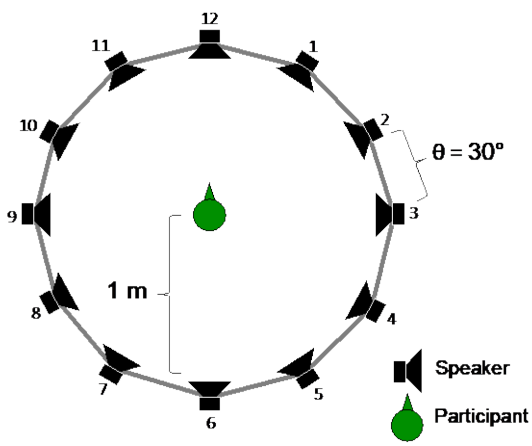


Figure 1. Schematic depicting overhead view of the loudspeaker layout.

System Functional Analysis

The Meister's plan will be used as a system functional analysis tool to evaluate and prioritize loudspeaker system design criteria and to select the most effective alternative. Table 1 defines the design criteria chosen to evaluate the loudspeaker alternatives in order to meet PALAT system requirements.

Table 1. Loudspeaker design criteria.

Criteria	Definition & requirements
Performance requirements	
Frequency response	The frequency response is the range of frequencies over which a loudspeaker produces a sound pressure level that remains within a specific ±dB tolerance level of its nominal sensitivity level. Typically, the tolerance level is set at ±3 dB for mid to high frequencies and ±6 dB for low frequencies depending on the size and quality of the loudspeakers (Emanuel, Maroonroge, & Letowski, 2009). Frequency response is an output measure based on a constant level input of pure tone frequencies (Borwick, 2001). The measurement is given by a stated frequency range in Hz within a dB SPL tolerance range, e.g. 200 Hz – 16000 Hz (±3 dB). Frequency response is often measured on-axis, i.e., directly in front of the loudspeaker, at a distance of 1 meter (m) with 1 Watt (W) of power. A flat frequency response over a broad frequency spectrum means the loudspeaker is capable of reproducing the input sound accurately. The PALAT system will need to reproduce military relevant sounds that span the frequency spectrum including the low frequency sounds of gunshots and explosions, and high frequency sounds of the whistle of incoming rocket propelled grenades or mortars, and the clicks emitted by charging of a rifle (Clasing & Casali, 2014). In addition to presenting military relevant sound across the audible spectrum, to train localization, the PALAT system *must be able* to accurately produce sounds that provide interaural time difference cues below 1500 Hz and interaural level difference cues above 3000 Hz (Casali & Tufts, in press). *Based on the requirements above, the PALAT system must have a flat frequency response of 250 Hz – 10000 Hz within ±3 dB.*
Total harmonic distortion	Total harmonic distortion (THD) is the amount of amplitude distortion present in a signal as a result of mechanical or magnetic nonlinearities in the loudspeaker or impurities in the voltages and currents in the power system (Eargle, 2003; Skvarenina, 2002). The distortions occur at integral multiples (i.e., harmonics) of the fundamental frequency of the signal. Calculated relative levels of harmonics compared to the fundamental can be expressed in percentages or decibels (dB) as in the table below (Newell & Holland, 2007): 0 dB — 100 % -10 dB — 30 % -20 dB — 10 % -30 dB — 3 % -40 dB — 1 % In the PALAT system, the performance effect on localization based on THD will vary depending on the frequency spectrum and the sound

	pressure level of the signal presented. *<u>As a general rule, THD should be minimized.</u>*
Ability to reproduce DRILCOM signal	The PALAT system shall reproduce the localization training results of a proven full-scale, laboratory grade DRILCOM system consisting of 12 integrally-powered, 5.25-inch loudspeakers. Previous studies using the DRILCOM system demonstrated the ability to improve the open ear's absolute correct performance by over 25% and demonstrated that participants using certain TCAPS can learn and perform at similar ballpark levels to the open ear with relatively little training (Casali & Lee, 2016a; Casali & Robinette, 2014). *<u>The PALAT system loudspeakers must be able to reproduce the localization training signals with similar fidelity as produced by the DRILCOM loudspeakers in order to achieve comparable localization training effects.</u>*
Sensitivity	Loudspeaker sensitivity is the sound pressure level in dB SPL measured at 1 m distance from the loudspeaker in response to a 1 W signal (Tran, Amrein, & Letowski, 2009). The test signal is usually pink noise limited to the frequency range output of the loudspeaker (Borwick, 2001). The sensitivity measurement combined with the power rating helps to determine if the loudspeaker can produce the desired dB SPL for localization training in background noise. Sensitivity is measured at the approximate distance between the listener and the front of the loudspeaker in the PALAT system. *<u>To ensure that a small amplifier will be sufficient to power the system, the minimum sensitivity for the PALAT system loudspeaker must be approximately 75 dBA with an ample power rating to achieve 85 dBA at the listener's ear.</u>*
Power requirements	
Recommended power rating	The power rating, also referenced as power capacity or maximum input power, is the highest continuous power that a loudspeaker can receive without being damaged or without producing sound distortion beyond a specified level. Power rating is measured in Watts (W) and is usually characterized by a specific percent of nonlinear distortions that the loudspeaker cannot exceed under normal conditions (Emanuel, Maroonroge, & Letowski, 2009). The power rating is usually listed in terms of continuous and peak power (Eargle, 2003). Power rating plays a role in acoustical performance in terms of being able to produce the desired SPL at the listener's ear. A doubling of power is required to increase the sound pressure level by 3 dB, e.g. to increase from 84 dB at 1 meter using 1 watt of power to 87 dB would require 2 watts (Eargle, 2003).

	Higher power ratings are preferred in order to ensure that the loudspeaker can produce 85 dBA at the listener's ear without distorting the signal.
Impedance	Impedance in loudspeakers is the amount of electrical resistance to power present within the electrical and mechanical components. A higher impedance, measured in Ohms, requires more voltage in order to produce the same Wattage of power. The PALAT system will require a small central amplifier to distribute power to the loudspeakers. *Lower impedance loudspeakers are preferred if the sensitivity ratings and power ratings provide enough headroom to produce signals at 85 dBA.*
Portability requirements	
Weight	The PALAT system will incorporate 12-loudspeakers and must be capable of being assembled, operated, and disassembled by two personnel. *As a result, loudspeakers of minimal weight are preferred to increase portability of the system.*
Physical Dimensions	*Loudspeakers of minimal size are preferred to increase portability.*
Durability & usability requirements	
Driver type	Dynamic drivers are the most prevalent loudspeakers used in auditory localization research and have proven to demonstrate the ability to not only produce signals that are localizable with HPDs and TCAPS but also to perform well in the few studies that have tested localization training (e.g. Abel & Paik, 2004; Wright & Zhang, 2006; Casali & Robinette, 2014; Casali & Lee, 2016a). The PALAT system will be limited to small, full-range drivers due to portability requirements. However, in addition to traditional full-range drivers, there are several loudspeaker driver design options for the PALAT system including two-way coaxial loudspeakers and flat panel loudspeakers. It is important to evaluate the driver type on their durability and their effect on auditory localization. *The PALAT system loudspeaker driver type must consistently produce a signal that is localizable at a distance of 1 meter under open ear listening conditions and while wearing hearing protection devices.*
Wire terminal type	In order to increase usability, the PALAT system loudspeaker wire terminals must establish and maintain a secure electrical connection during repeated assembly, operation, and disassembly. Some users may be inexperienced with audio equipment. As a result, the loudspeaker wire terminals should allow for easy identification of the positive and negative terminals.

	Based on the requirements above, the PALAT system loudspeaker wire terminals must provide a secure electrical connection and easily identifiable terminal polarity.

Criterion Comparison
The first step in the Meister plan is to conduct a pairwise comparison for every design criterion in order to determine the value or weight of each criterion (Meister, 1985). In this section you are asked to compare every pair of criteria and select the criterion that is most important for the system. The selected criterion is assigned a value of 1 and a value of 0 is assigned to the criterion that is less important (Meister, 1985).

For each pair, please **select** and **circle** the criterion that you feel is **most important** to the system. (You must select a preference; no ties are allowed.)

Criterion A	vs	Criterion B
Frequency response		Total harmonic distortion
Frequency response		Ability to reproduce DRILCOM signal
Frequency response		Sensitivity
Frequency response		Recommended Power Rating
Frequency response		Impedance
Frequency response		Weight
Frequency response		Physical dimensions
Frequency response		Driver type
Frequency response		Wire terminal type
Total harmonic distortion		Ability to reproduce DRILCOM signal

Total harmonic distortion	Sensitivity
Total harmonic distortion	Recommended Power Rating
Total harmonic distortion	Impedance
Total harmonic distortion	Weight
Total harmonic distortion	Physical dimensions
Total harmonic distortion	Driver type
Total harmonic distortion	Wire terminal type
Ability to reproduce DRILCOM signal	Sensitivity
Ability to reproduce DRILCOM signal	Recommended Power Rating
Ability to reproduce DRILCOM signal	Impedance
Ability to reproduce DRILCOM signal	Weight
Ability to reproduce DRILCOM signal	Physical dimensions
Ability to reproduce DRILCOM signal	Driver type
Ability to reproduce DRILCOM signal	Wire terminal type
Sensitivity	Recommended Power Rating

Sensitivity	Impedance
Sensitivity	Weight
Sensitivity	Physical dimensions
Sensitivity	Driver type
Sensitivity	Wire terminal type
Recommended Power Rating	Impedance
Recommended Power Rating	Weight
Recommended Power Rating	Physical dimensions
Recommended Power Rating	Driver type
Recommended Power Rating	Wire terminal type
Impedance	Weight
Impedance	Physical dimensions
Impedance	Driver type
Impedance	Wire terminal type
Weight	Physical dimensions

Weight	Driver type
Weight	Wire terminal type
Physical dimensions	Driver type
Physical dimensions	Wire terminal type
Driver type	Wire terminal type

Alternative Comparison
The next step in the Meister plan analysis is to complete a pairwise comparison of each alternative for every design criterion based on objective data and subject matter expertise subjective evaluation. Data and information are presented below for three loudspeaker alternatives which meet all of the minimal feasible design criteria. Data results for the three alternatives, referred to as A, B, and C, are color coded:

Alternative A (blue) **Alternative B (green)** **Alternative C (red)**

After evaluating the data presented, conduct a pairwise comparison for each pair of alternatives selecting your preferred alternative for each design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		B
B		C
C		A

Frequency response
In principle, the PALAT system must have a flat frequency response of 250 Hz – 10000 Hz within ±3 dB.

Table 2 below displays the manufacturer-reported frequency response range within ±3 dB at 1 watt measured at 1 meter. Although 1 watt at 1 meter and a ±3 dB tolerance is the standard, Alternative B manufacturer did not specify ±3 dB and Alternative C did not specify 1 watt at 1 meter (Alt A, 2017; Alt B, 2017; Alt C, 2017).

Table 2. Manufacturer-reported frequency response.

	Alternative A	Alternative B	Alternative C
Frequency response	210 Hz – 16 kHz	120 Hz – 20 kHz	150 Hz – 20 kHz

An on-axis frequency response test was conducted by the author at the Virginia Tech – Auditory Systems Laboratory using a manual-stepped pure tone sinusoidal signal from 100 Hz to 20000 Hz. The test was conducted in an anechoic chamber with a 1-inch Larson-Davis measurement microphone placed 1 meter from the cone of the loudspeaker. Measurements were recorded using a Larson-Davis 2800 Model Spectrum Analyzer. The audio signal was generated using Audacity® 2.2.0 and presented via a MacBook Pro laptop with a Kemo® Electronic 12W audio amplifier. The output voltage was manually measured and set to produce 1 W. Of note, Alternative B and Alternative C are 8 Ohm loudspeakers and the output voltage was set to ~2.83 Vrms[1] at 1000 Hz. Alternative A is a 6 Ohm loudspeaker and the output voltage was set to ~2.45 Vrms at 1000 Hz. The volume and output voltage were not adjusted during the testing in order to try and maintain a constant voltage as specified by industry standards (AES 2-2012, 2012). Some deviations in frequency response may be attributable to frequency response limitations of the computer soundcard or amplifier. However, the tests were consistent across all Alternatives. Figure 2 displays the measured sound pressure level (dB SPL) in every 1/3-octave frequency band from 100 Hz – 20000 Hz.

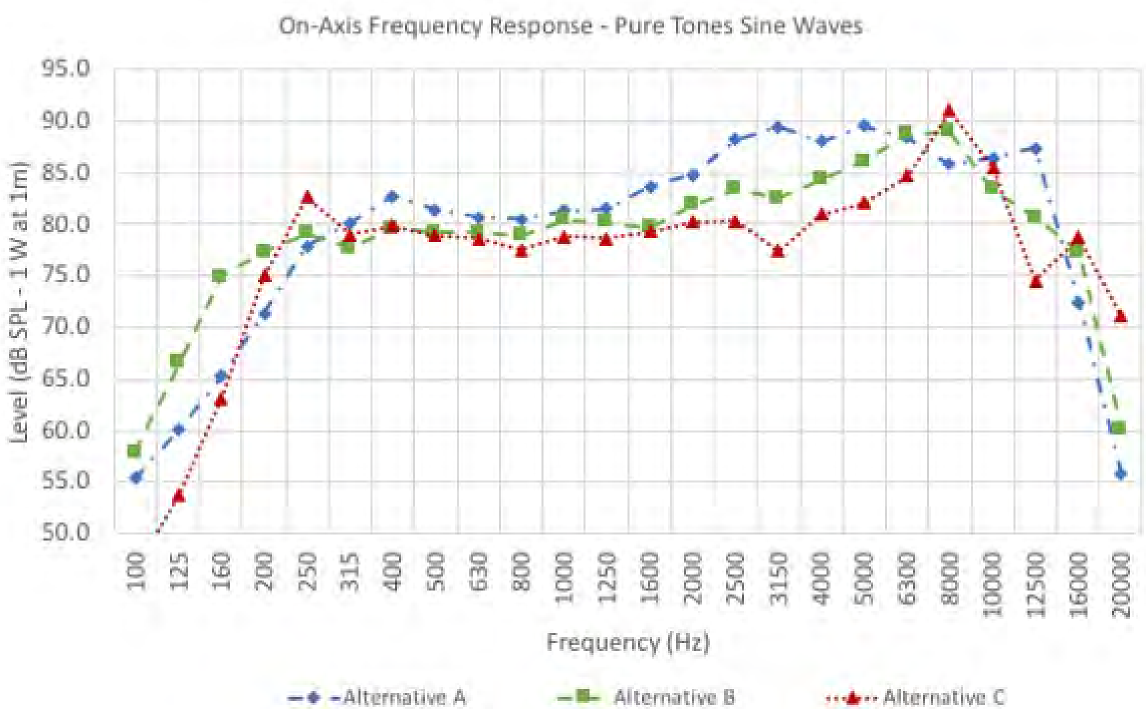


Figure 2. On-Axis frequency response – stepped sine (dB SPL 1 watt at 1 meter).

An additional frequency response test was conducted under similar conditions above using the Room EQ Wizard® computer software to generate a sinusoidal sweep and record the frequency response at each frequency (rather than in 1/3-octave bands as previously discussed). A MiniDSP UMIK-1 USB measurement microphone was used as the input source. Figure 3 displays the results of the computer-generated frequency response test.

[1] Vrms - Volts root mean square

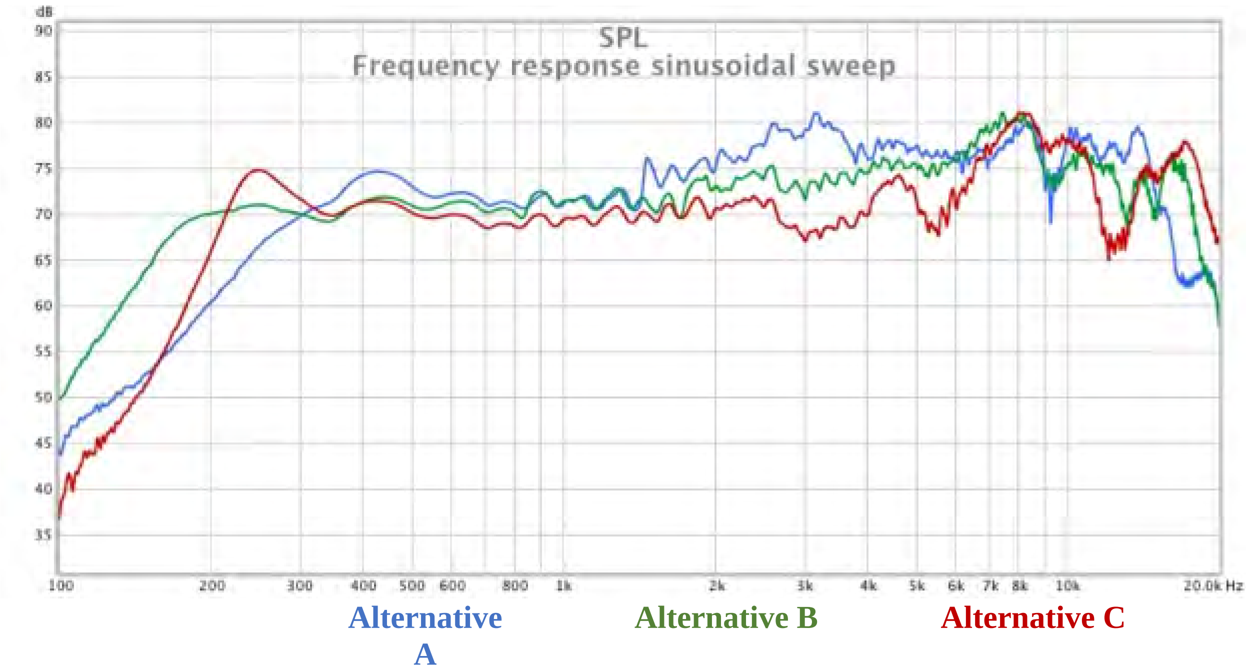


Figure 3. On-Axis frequency response – sine sweep (dB SPL 1 watt at 1 meter).

For each pair, please **select** and **circle** your preferred alternative for the frequency response design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		B
B		C
C		A

Total Harmonic Distortion

As a general rule, Total Harmonic Distortion (THD) should be minimized.

Harmonic distortion was measured with both a stepped sine and sine sweep measuring the second- and third-harmonic components compared with the fundamental frequency.

The first harmonic distortion test was conducted by the author at the Virginia Tech – Auditory Systems Laboratory using a manual-stepped pure tone sinusoidal signal from 100 Hz to 20000 Hz. This allowed for measurements of the 6th-harmonic up to 3150 Hz and 3rd-harmonic at 6300 Hz. The test was conducted in an anechoic chamber using the same measurement set-up as used above in the frequency response test (1 W at 1 m). Again, some of the distortion may be attributable to the computer soundcard or amplifier, but measurements were consistent across all alternatives. Figure 4 displays the total harmonic distortion in dB below the fundamental frequency. Figure 5 displays the total harmonic distortion as a percentage referenced to the fundamental frequency.

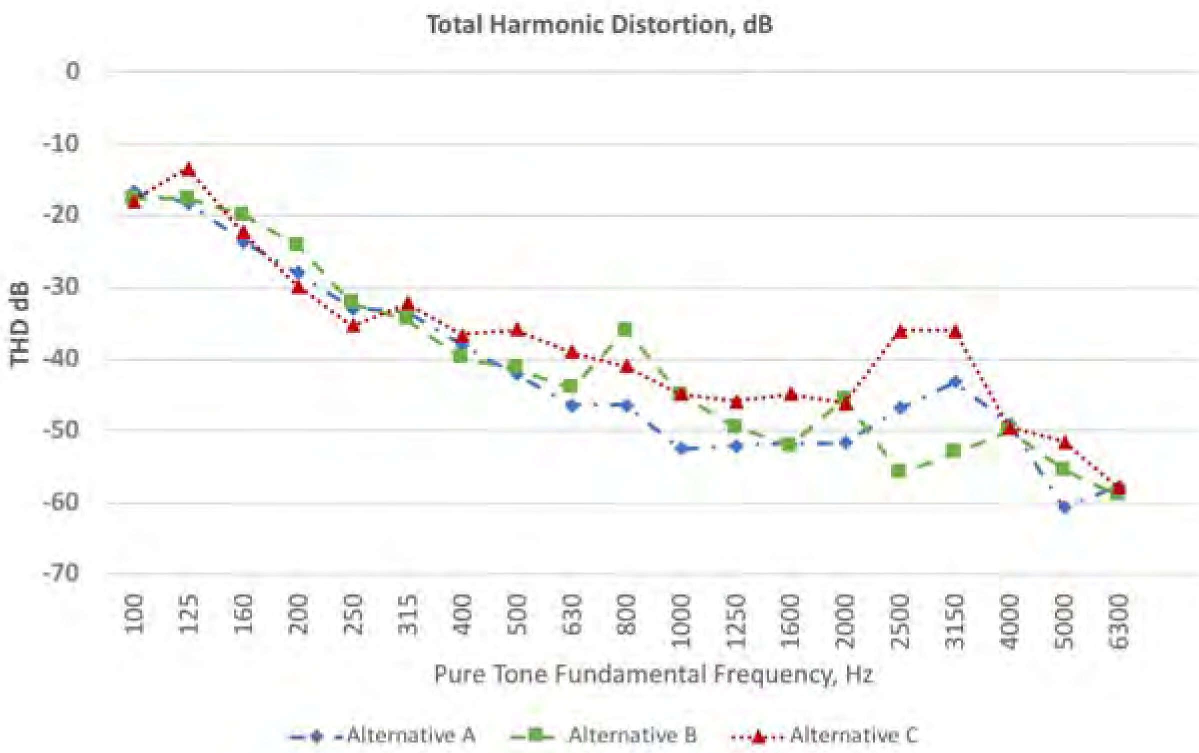


Figure 4. Total harmonic distortion reference to the fundamental frequency (dB).

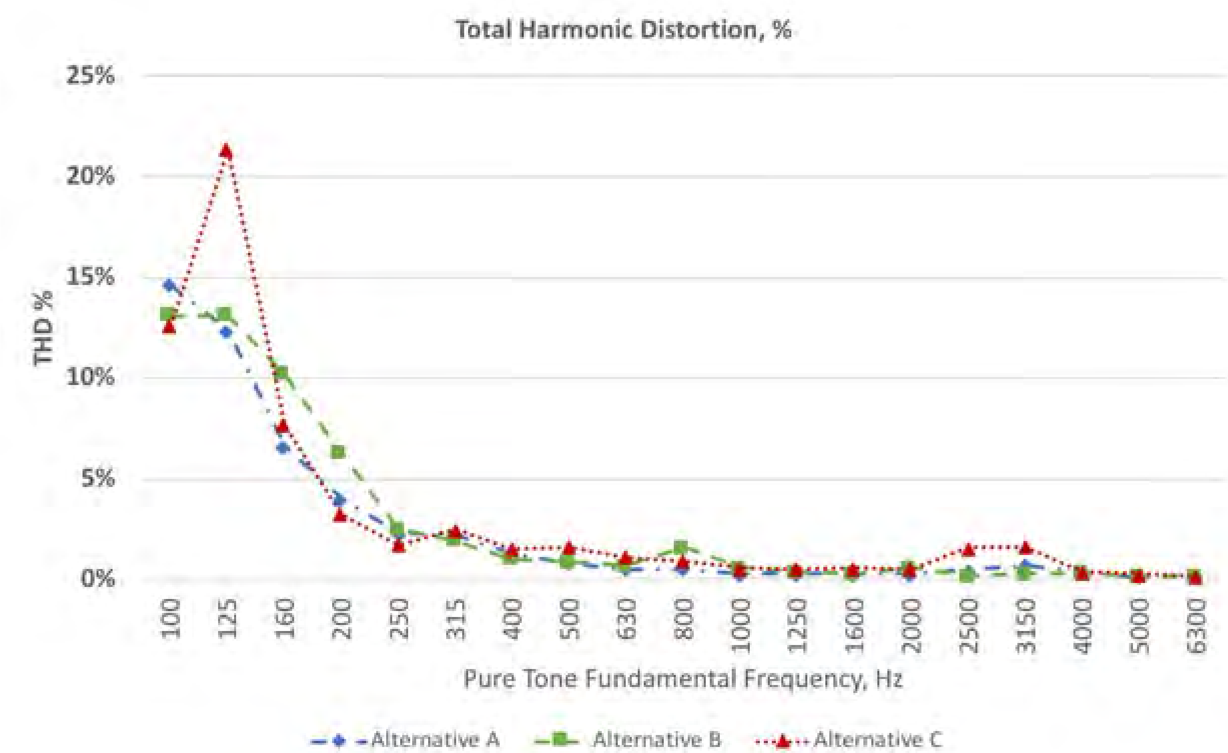


Figure 5. Total harmonic distortion reference to the fundamental frequency (%).

An additional harmonic distortion test was conducted under similar conditions above using the Room EQ Wizard® computer software to generate a sinusoidal sweep and record the distortion. Figures 6-8 displays the levels of distortion in dB, total harmonic distortion, 2nd-harmonic, and 3rd-harmonic, referenced to the level of the fundamental (AES 2-2012, 2012).

Alternative A

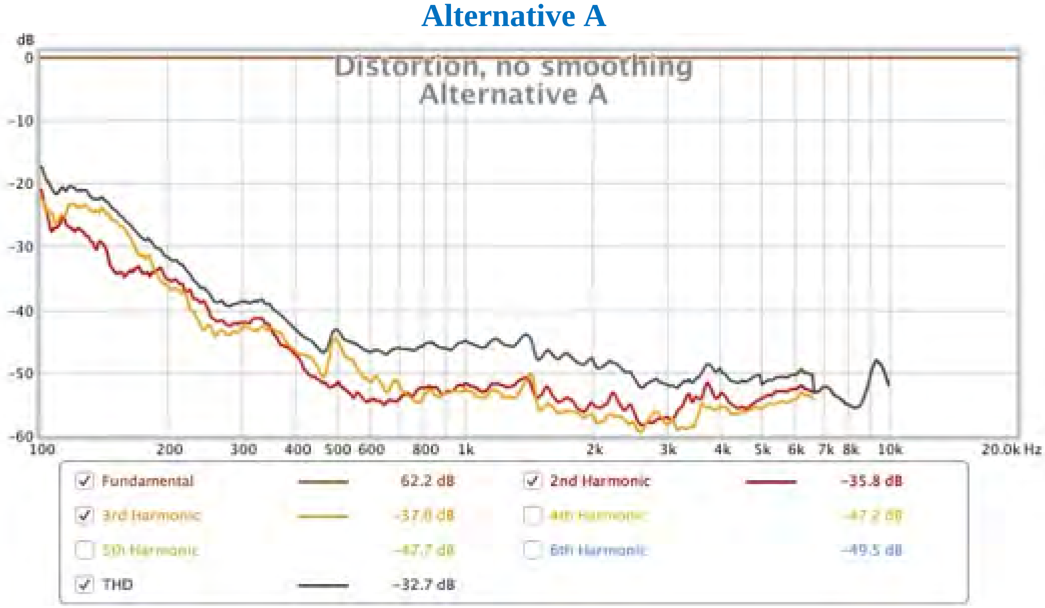


Figure 6. Alternative A total harmonic distortion reference to the fundamental frequency (dB).

Alternative B

Figure 7. Alternative B total harmonic distortion reference to the fundamental frequency (dB).

Alternative C

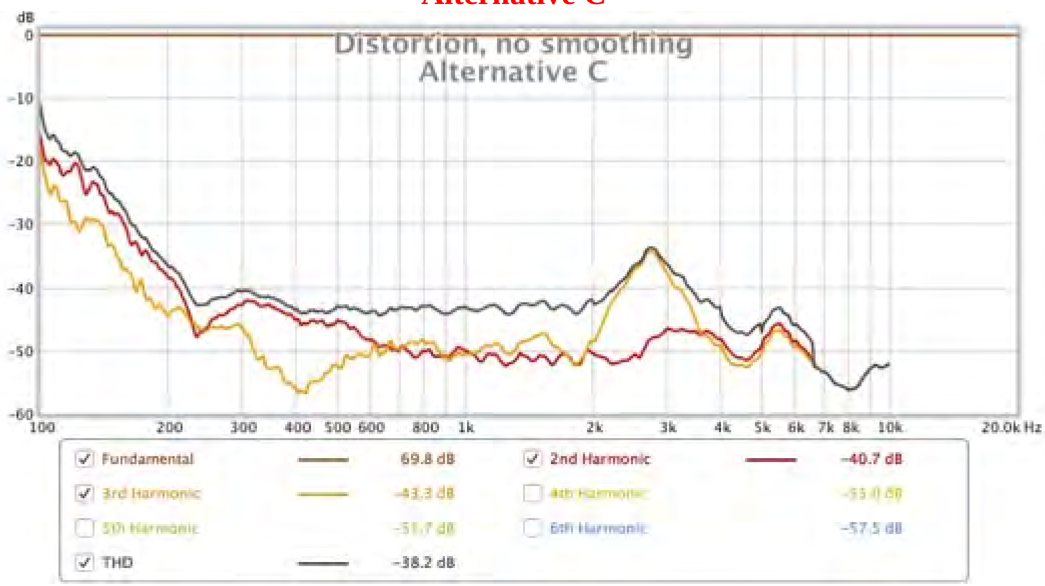


Figure 8. Alternative C total harmonic distortion reference to the fundamental frequency (dB).

For each pair, please **select** and **circle** your preferred alternative for the total harmonic distortion design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		**B**
B		**C**
C		**A**

Ability to reproduce DRILCOM signal

The PALAT system loudspeakers must be able to reproduce the localization training signals with similar fidelity as produced by the DRILCOM loudspeakers in order to achieve comparable localization training effects.

A test was conducted in the hemi-anechoic DRILCOM lab room to measure the sound pressure level across the frequency spectrum from 100 Hz to 10000 Hz for a dissonant tone signal and four military relevant signals, a whistle of an incoming artillery round (Whistle), the rotor sounds of an approaching Apache helicopter (Apache), spoken foreign language (Arabic), and an AK-47 three round burst (AK-47). Each loudspeaker was calibrated at 55 dBA and 80 dBA for the dissonant tone signal. The DRILCOM loudspeaker, Behringer Behritone C50A, was measured at a distance of 1.5 meters from the measurement microphone and the three PALAT alternative loudspeakers were measured at a distance of 1 meter from the measurement microphone. The graphs below display the measured sound pressure level (dB SPL) for each 1/3-octave band frequency and the absolute deviation from the DRILCOM reference loudspeaker. The deviation graph on the far right plots the total absolute deviation, logarithmic sum across all frequencies, in order to show the total absolute delta.

Dissonant tone signal at 55 dBA

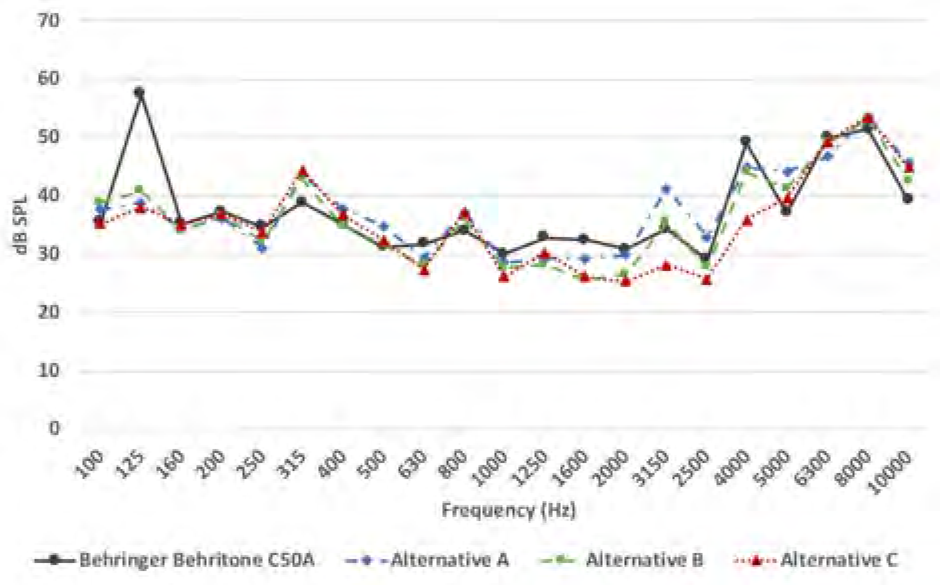


Figure 9. Sound pressure level of dissonant tone signal tone at 55 dBA.

Dissonant tone signal at 55 dBA

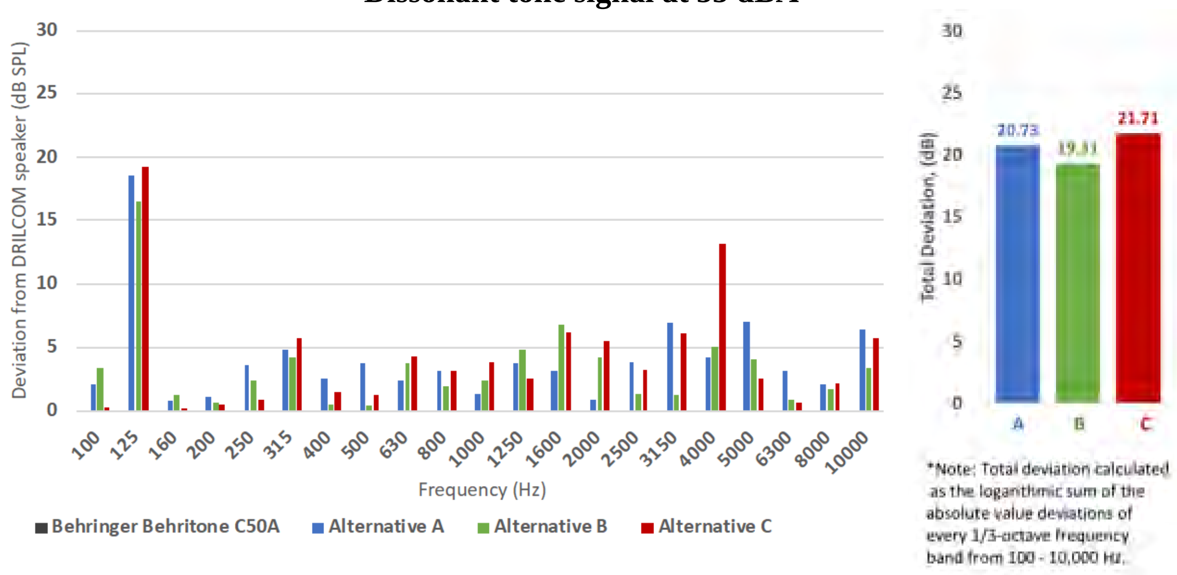


Figure 10. Sound pressure level deviations from DRILCOM loudspeaker for dissonant tone signal tone at 55 dBA.

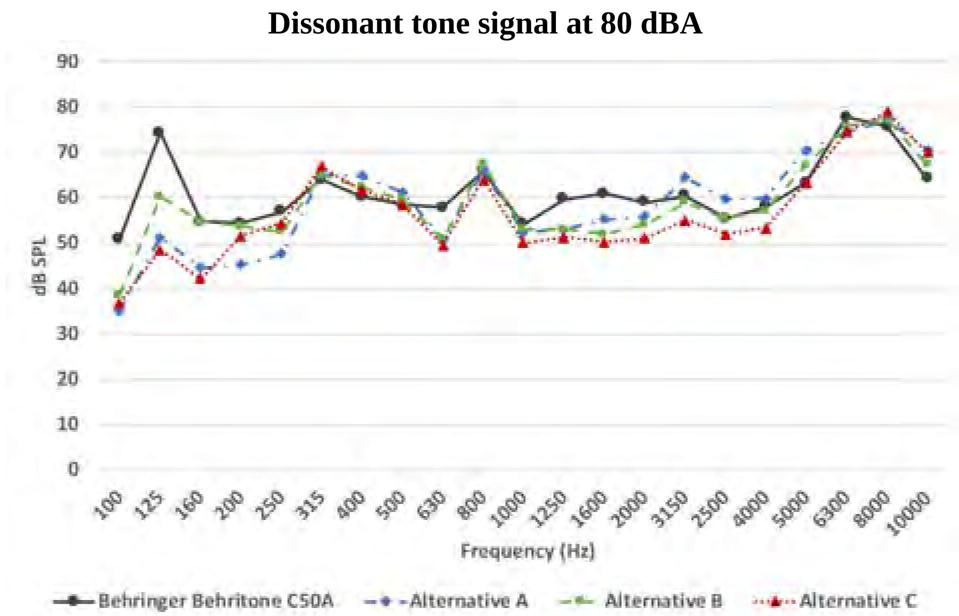


Figure 11. Sound pressure level of dissonant tone signal tone at 80 dBA.

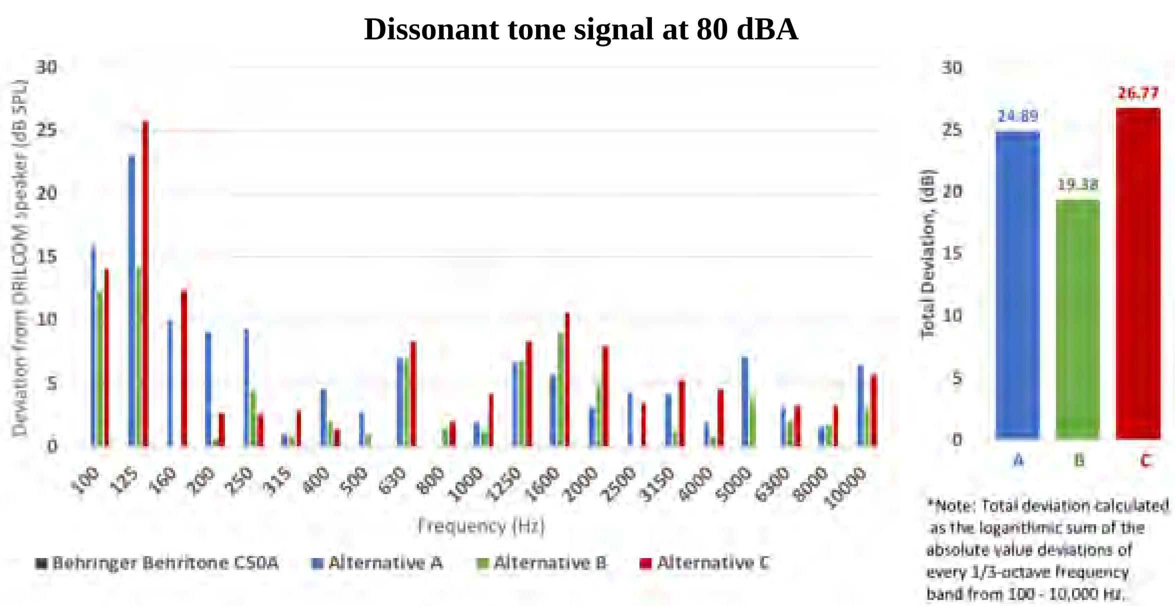


Figure 12. Sound pressure level deviations from DRILCOM loudspeaker for dissonant tone signal tone at 80 dBA.

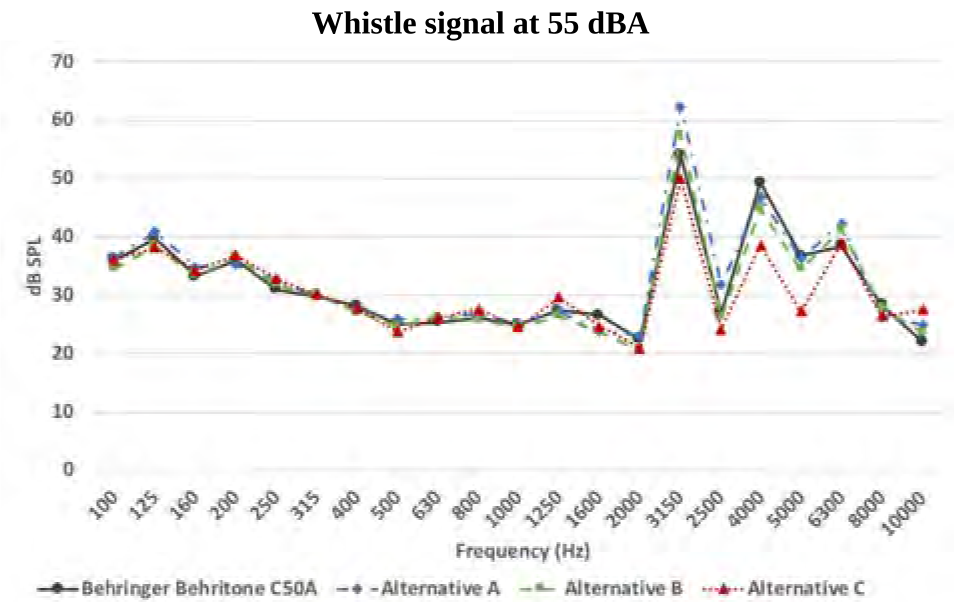


Figure 13. Sound pressure level of Whistle signal at 55 dBA.

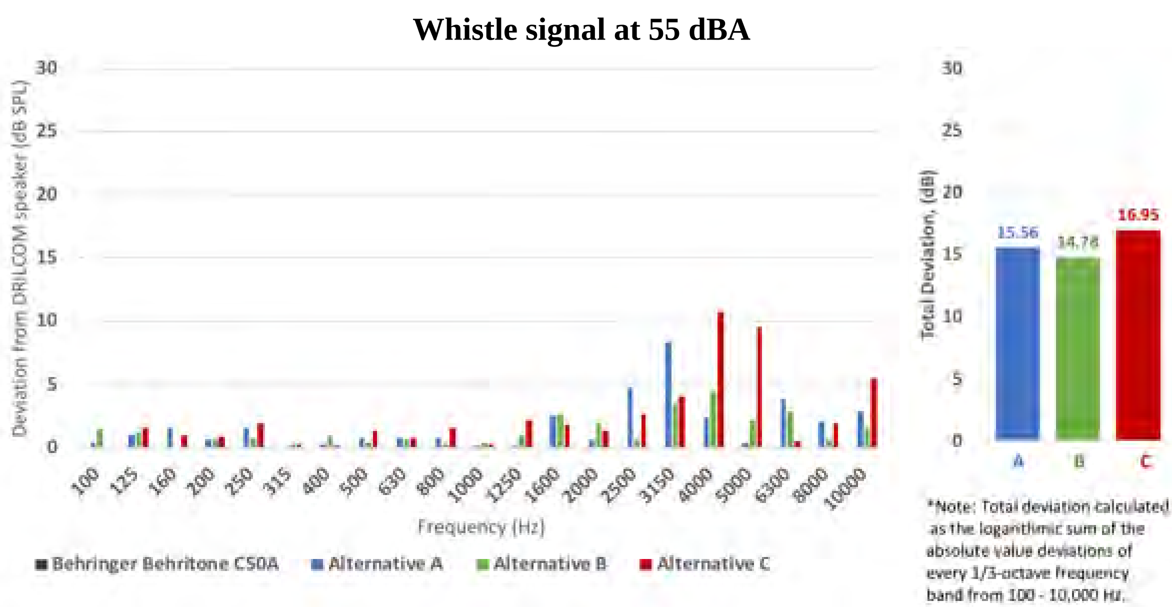


Figure 14. Sound pressure level deviations from DRILCOM loudspeaker for Whistle signal at 55 dBA.

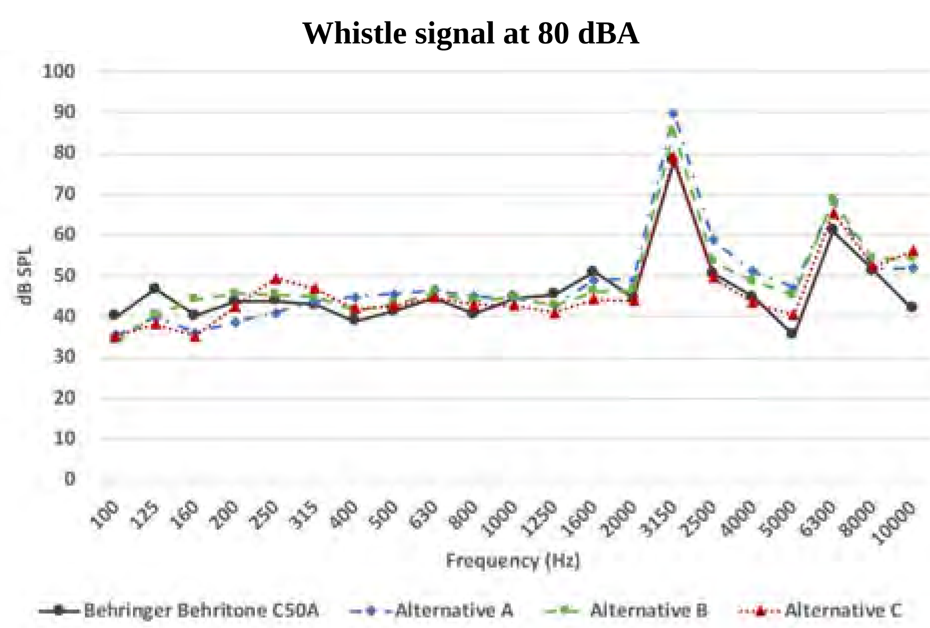


Figure 15. Sound pressure level of Whistle signal at 80 dBA.

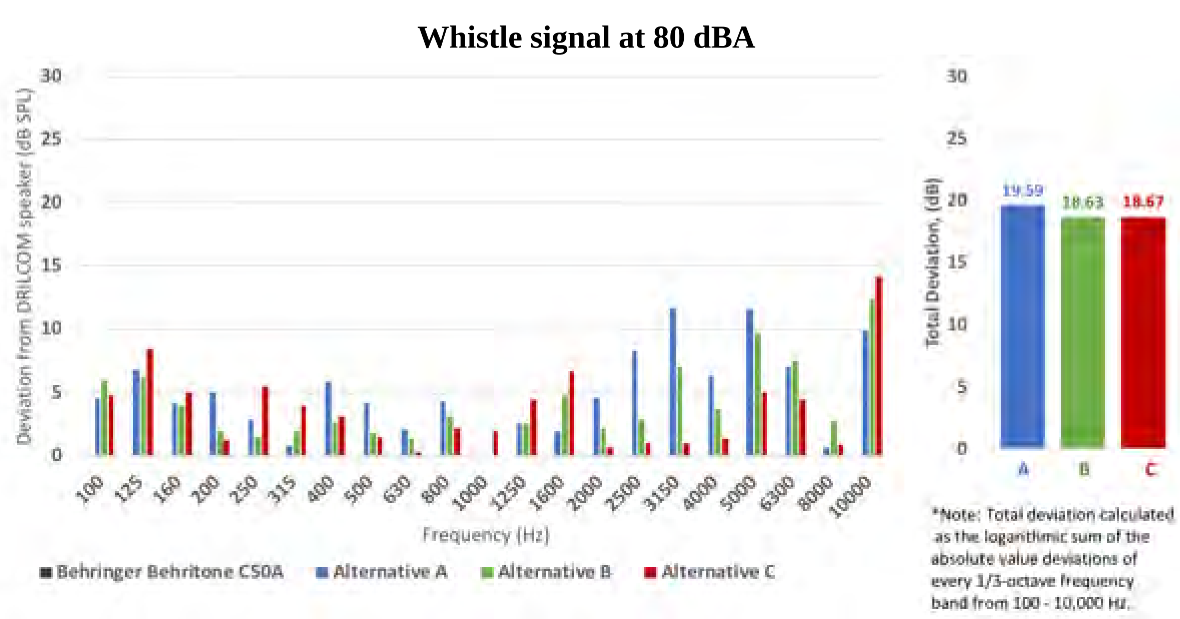


Figure 16. Sound pressure level deviations from DRILCOM loudspeaker for Whistle signal at 80 dBA.

Apache signal at 55 dBA

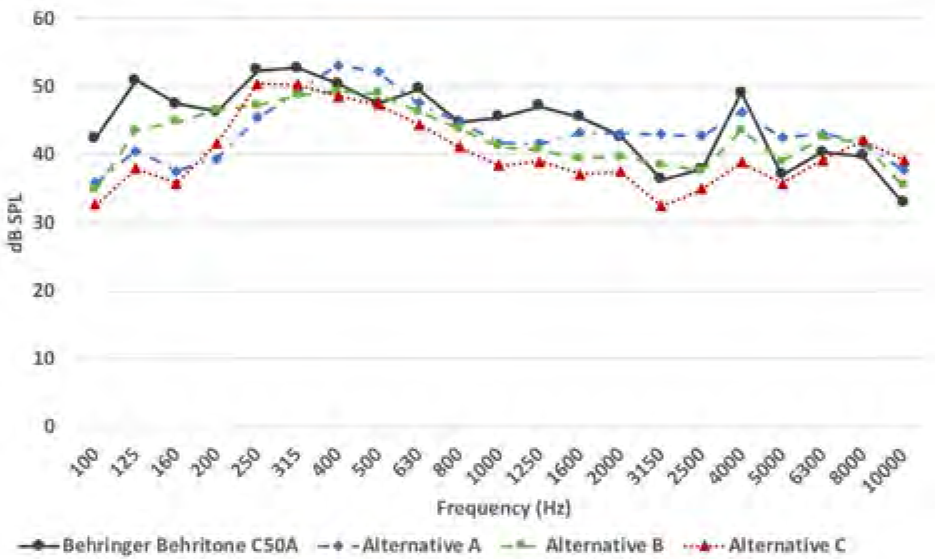


Figure 17. Sound pressure level of Apache signal at 55 dBA.

Apache signal at 55 dBA

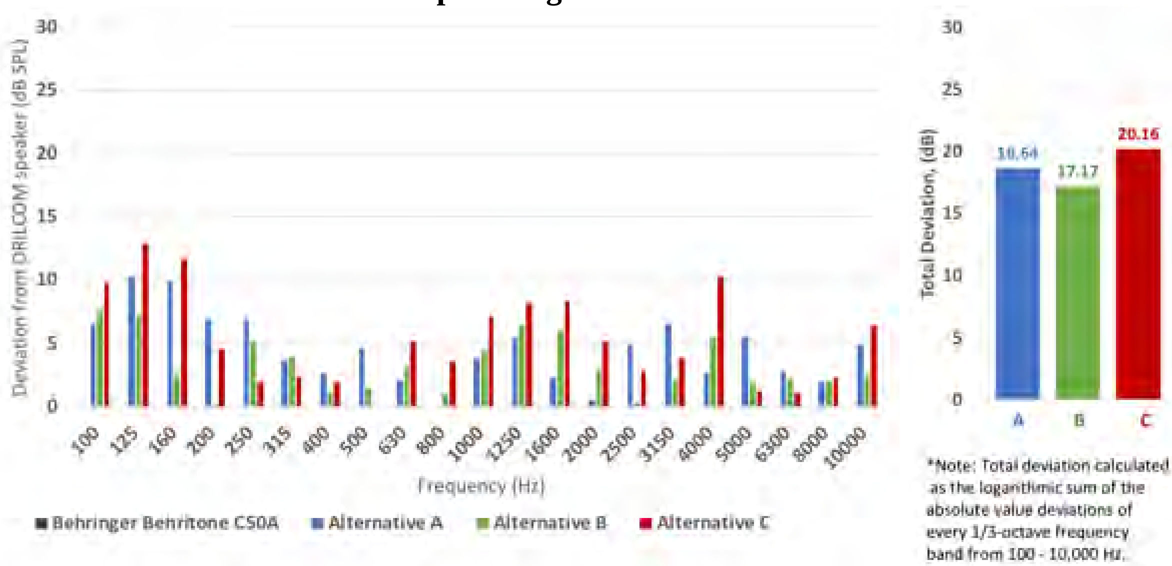


Figure 18. Sound pressure level deviations from DRILCOM loudspeaker for Apache signal at 55 dBA.

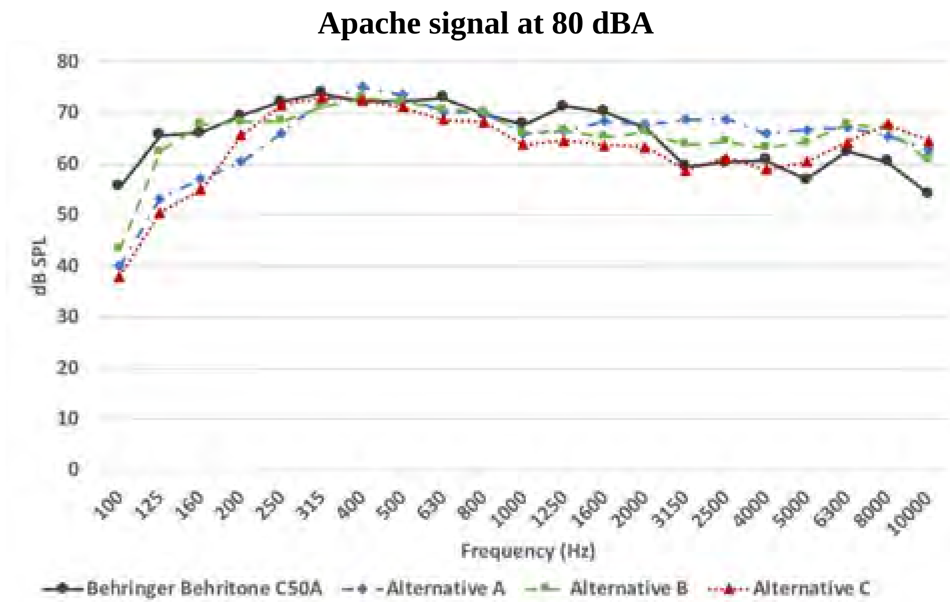


Figure 19. Sound pressure level of Apache signal at 80 dBA.

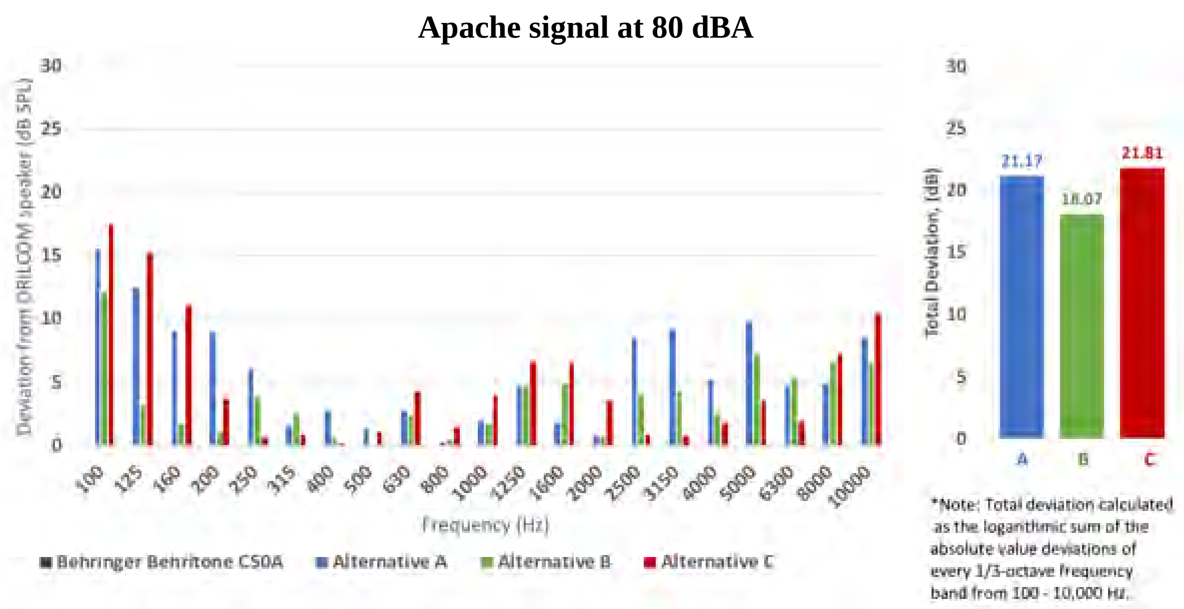


Figure 20. Sound pressure level deviations from DRILCOM loudspeaker for Apache signal at 80 dBA.

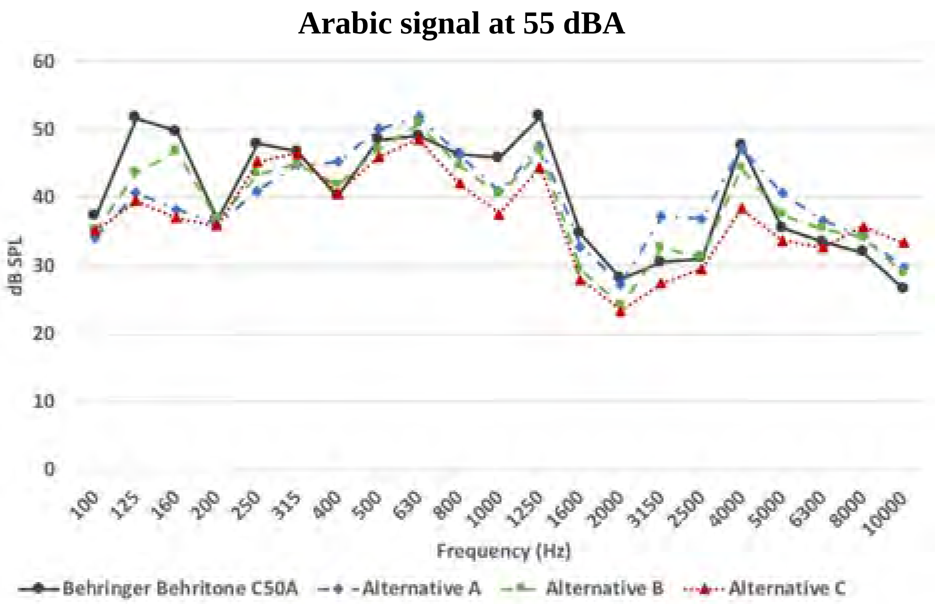


Figure 21. Sound pressure level of Arabic signal at 55 dBA.

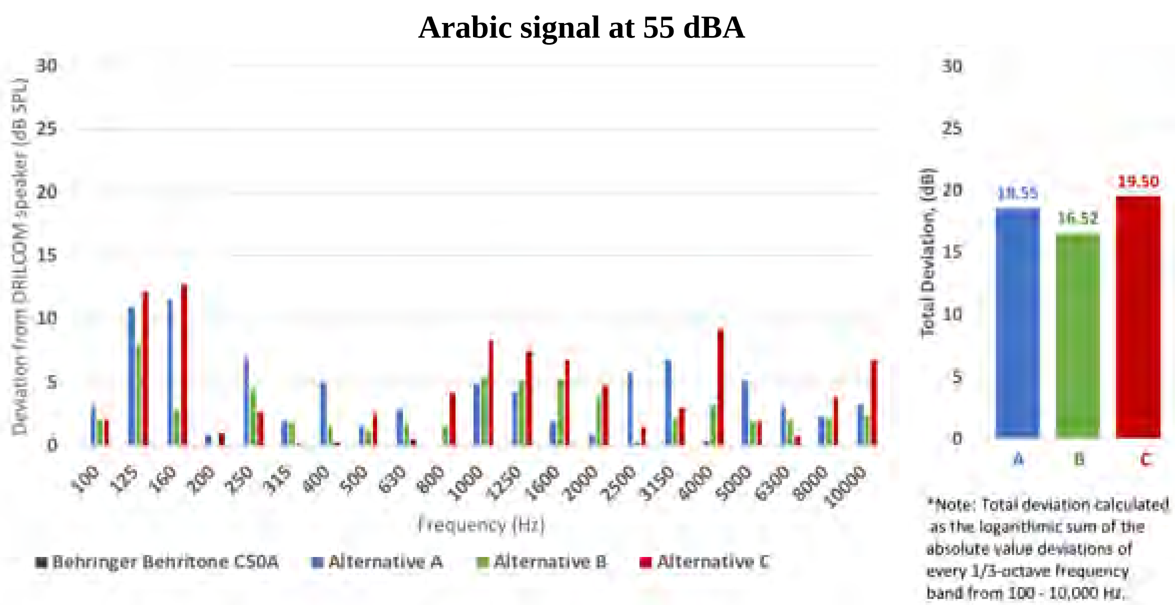


Figure 22. Sound pressure level deviations from DRILCOM loudspeaker for Arabic signal at 55 dBA.

Arabic signal at 80 dBA

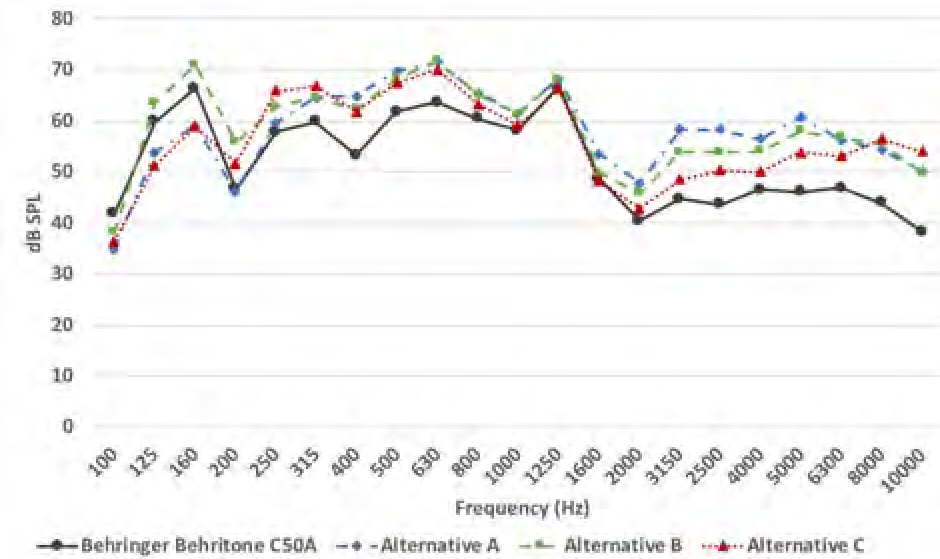


Figure 23. Sound pressure level of Arabic signal at 80 dBA.

Arabic signal at 80 dBA

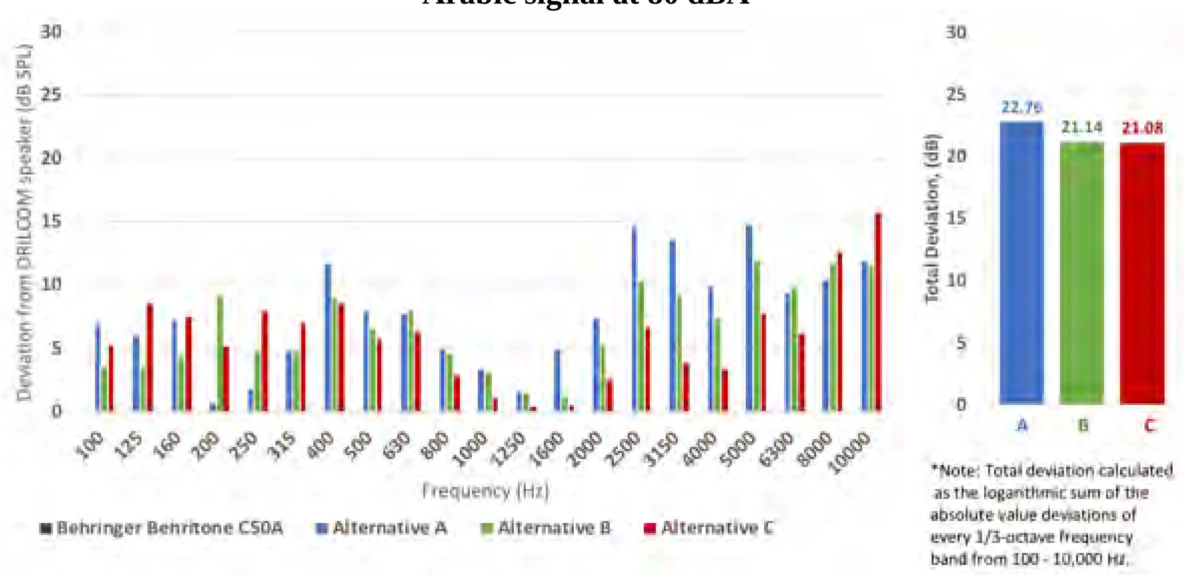


Figure 24. Sound pressure level deviations from DRILCOM loudspeaker for Arabic signal at 80 dBA.

AK-47 three-round burst signal at 55 dBA

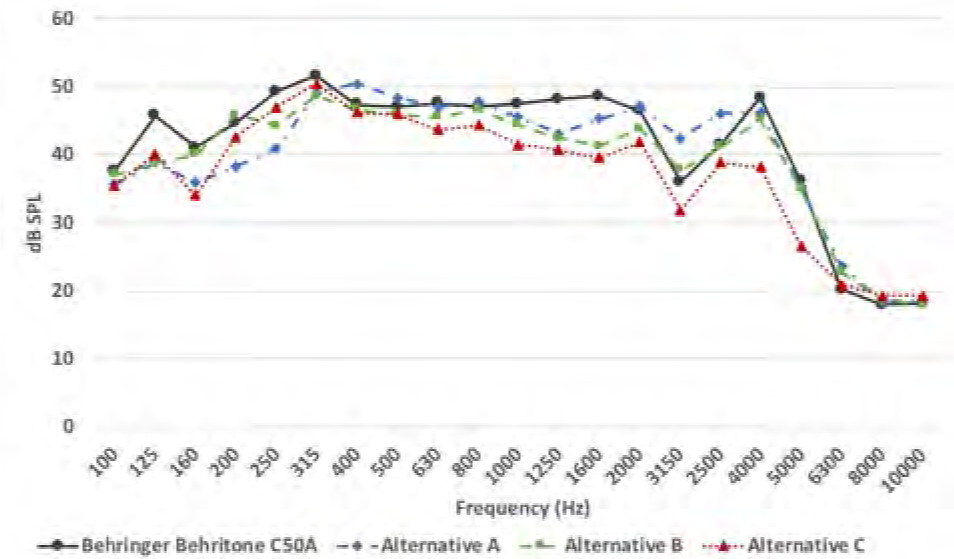


Figure 25. Sound pressure level of AK-47 three-round burst signal at 55 dBA.

AK-47 three-round burst signal at 55 dBA

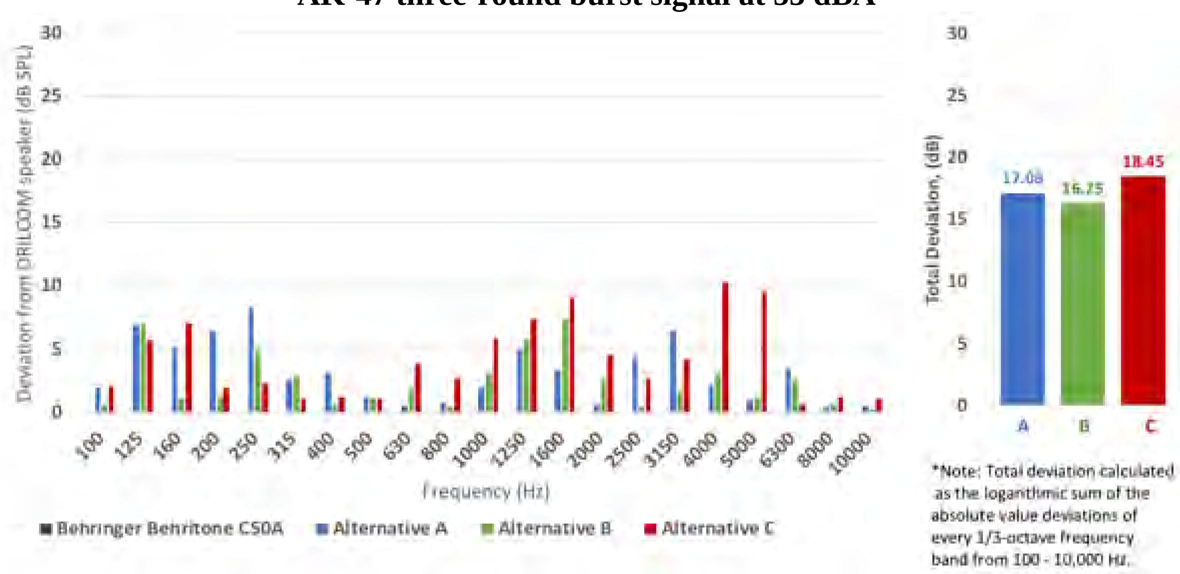


Figure 26. Sound pressure level deviations from DRILCOM loudspeaker for AK-47 three-round burst signal at 55 dBA.

AK-47 three-round burst signal at 80 dBA

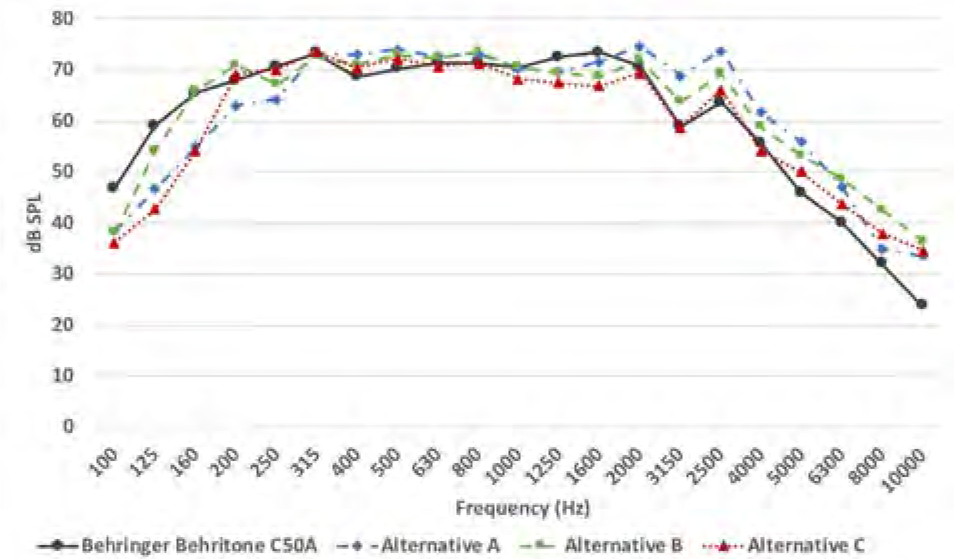


Figure 27. Sound pressure level of AK-47 three-round burst signal at 80 dBA.

AK-47 three-round burst signal at 80 dBA

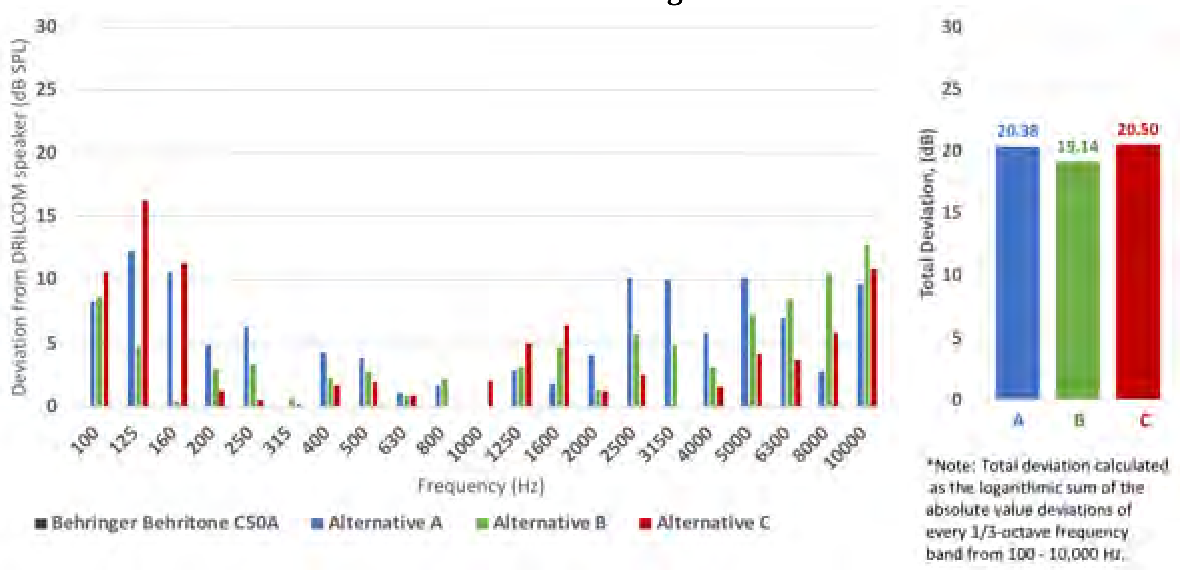


Figure 28. Sound pressure level deviations from DRILCOM loudspeaker for AK-47 three-round burst signal at 80 dBA.

For each pair, please **select** and **circle** your preferred alternative for the ability to reproduce DRILCOM signal design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		B
B		C
C		A

Sensitivity

In principle, the minimum sensitivity for the PALAT system loudspeaker must be approximately 75 dBA with an ample power rating to achieve 85 dBA at the listener's ear.

Table 3 below displays the manufacturer-reported sensitivity in dB SPL at 1 watt measured at 1 meter (Alt A, 2017; Alt B, 2017; Alt C, 2017). Note: Alternative A, a 6-ohm loudspeaker, requires less voltage to achieve 1 W. Standard acoustical measurement errors are typically within ±2 dB (IEEE Std 219-1975, 1975).

Table 3. Manufacturer-reported sensitivity.

	Alternative A	Alternative B	Alternative C
Sensitivity	84 dB SPL	86 dB SPL	85 dB SPL

For each pair, please **select** and **circle** your preferred alternative for the sensitivity design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		B
B		C
C		A

Recommended Power Rating

Higher power ratings are preferred in order to ensure that the loudspeaker can produce 85 dBA at the listener's ear without distorting the signal.

Table 4 below displays the manufacturer-reported recommended power rating in Watts (Alt A, 2017; Alt B, 2017; Alt C, 2017).

Table 4. Manufacturer-reported recommended power rating.

	Alternative A	Alternative B	Alternative C
Power rating	12 W (48 W peak)	25 – 200 W	10 – 100 W

For each pair, please **select** and **circle** your preferred alternative for the recommended power rating design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		B
B		C
C		A

Impedance

The impedance of the loudspeakers, along with the sensitivity ratings and power ratings, should provide enough headroom to produce signals at 85 dBA without distortions.

Table 5 below displays the manufacturer-reported impedance in Ohms (Alt A, 2017; Alt B, 2017; Alt C, 2017).

Table 5. Manufacturer-reported impedance.

	Alternative A	Alternative B	Alternative C
Impedance	6 Ohms	8 Ohms	8 Ohms

For each pair, please **select** and **circle** your preferred alternative for the impedance design criterion. *(Since Alternative B and C are both 8 Ohm loudspeakers, please rate your preference between Alternative A, 6 Ohms, verses either Alternative B or C, 8 Ohms.)*

Alternative	vs	Alternative
A		B (or) C

Weight

In principle, loudspeakers of minimal weight are preferred to increase portability of the system.

Table 6 below displays the manufacturer-reported weight for the loudspeaker in pounds (Alt A, 2017; Alt B, 2017; Alt C, 2017).

Table 6. Manufacturer-reported weight.

	Alternative A	Alternative B	Alternative C
Weight	1.9 lbs	0.95 lbs	1 lb

For each pair, please **select** and **circle** your preferred alternative for the weight design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		B
B		C
C		A

Physical dimensions

In principle, loudspeakers of minimal size are preferred to increase portability.

Table 7 below displays the manufacturer-reported dimensions for the loudspeaker in inches (Alt A, 2017; Alt B, 2017; Alt C, 2017).

Table 7. Manufacturer reported dimensions.

	Alternative A	Alternative B	Alternative C
Dimensions	3.0 x 3.0 x 4.0 inches	3.1 x 3.1 x 3.3 inches	3.7 x 4.3 x 4.5 inches

For each pair, please **select** and **circle** your preferred alternative for the physical dimensions design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		B
B		C
C		A

Driver type

The PALAT system loudspeaker driver type must consistently produce a signal that is localizable at a distance of 1 meter under open ear listening conditions and while wearing hearing protection devices.

Table 8 below displays a description of the driver type for each alternative (Alt A, 2017; Alt B, 2017; Alt C, 2017).

Table 8. Loudspeaker driver type.

	Alternative A	Alternative B	Alternative C
Driver size(s)	2.5-inch full-range	2.25-inch BMR	2.5-inch woofer 0.5-inch tweeter
Driver type	Full-range cone driver	Balanced Mode Radiator (BMR)	2-way, coaxial driver
	Typically consist of a wire voice coil inside a magnetic field attached to a funnel shaped diaphragm, or cone. Pistonic motion is used to create acoustical waves.	BMR is a flat loudspeaker that combines the pistonic motion of cone drivers with the vibration motion of flat panel loudspeakers.	The tweeter is mounted directly in front of the 2.5-inch woofer via a bridge mount. Crossover frequency of 5000 Hz.

For each pair, please **select** and **circle** your preferred alternative for the driver type design criterion. Please evaluate the loudspeaker driver types impacts on auditory localization. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		B
B		C
C		A

Wire terminal type

The PALAT system loudspeaker wire terminals must provide a secure electrical connection during repeated assembly, operation, and disassembly. In addition, the terminal polarity should be clearly marked.

Table 9 below displays a description of the wire terminal types (Alt A, 2017; Alt B, 2017; Alt C, 2017). Figures 29-31 show each alternatives wire terminal and details polarity markings.

Table 9. Loudspeaker terminal type and compatible connectors.

Wire terminal type	Alternative A Spring clip terminal	Alternative B 4-way Binding post	Alternative C Spring clip terminal
Compatible connectors	Bare wire Pin connectors	Bare wire Pin connectors Spade connectors Banana plugs	Bare wire Pin connectors

Alternative A

Figure 29. Alternative A wire terminals.

Alternative B

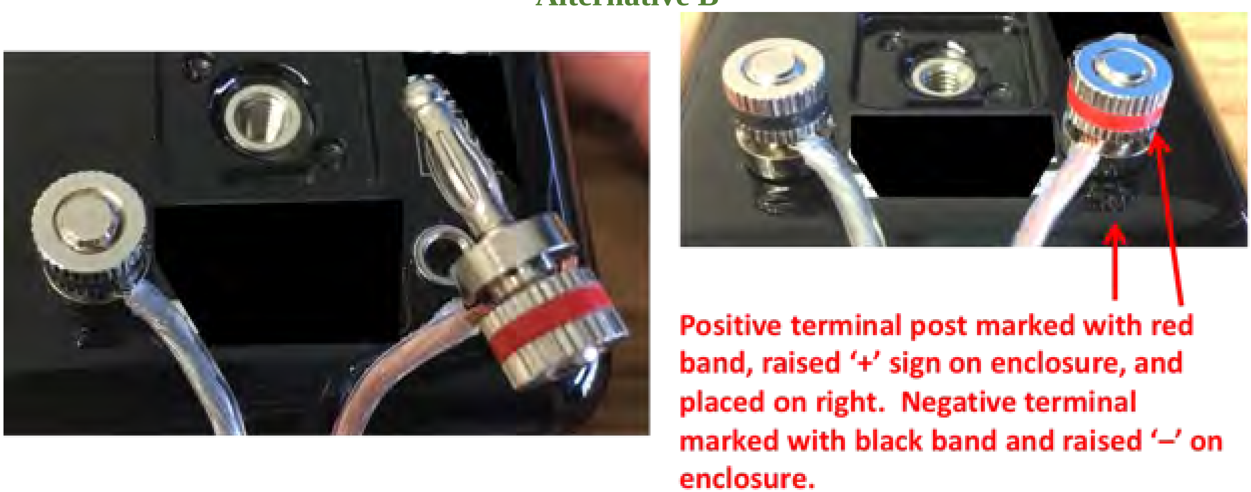


Figure 30. Alternative B wire terminals.

Alternative C

Figure 31. Alternative C wire terminals.

For each pair, please **select** and **circle** your preferred alternative for the wire terminal type design criterion. *(You must select a preference; no ties are allowed.)*

Alternative	vs	Alternative
A		**B**
B		**C**
C		**A**

Please feel free to add any additional comments or feedback that you feel will help in selecting the most effective loudspeaker for the PALAT system.

____

____

____

____

____

____

____

____

____

Appendix D. PALAT System Parts List and Costs

Item	Part	Quantity	Unit Price	Total Cost
PALAT Frame				
4 x 6 ft Canopy		3	$36.26	$108.80
	Telescoping Pole (4 per)	12		
	Accordion Scissor Support Joints	12		
1 x 2.5 in by 12 ft Aluminum Bar		1	$112.00	$112.00
	Bottom (stationary) Connector	12		
	Top (sliding) Connector	12		
Washer for connect	Washer	20	$0.14	$2.95
3/8-16 UNC-2A Thread, Stubby Pull Ring	Pull Pin	2	$2.98	$5.96
Low-Friction Surface-Protection Tape Made with Teflon® PTFE	Teflon Tape (Anti-Friction)	1	$42.00	$42.00
Bolts for bottom connector	10-24 x 1.5 in Phillips Machine Bolt	12	$7.50	$7.50
48 inch 1 x 1/8 Aluminum flat bar	L-bracket, speaker mount	2	$6.30	$13.27
48 inch 1 x 1/8 Aluminum flat bar	L-bracket, speaker mount	4	$6.30	$26.54
Philips Pan Head Screw for Loudspeaker mount	Loudspeaker bolt for L-bracket	100	$10.58	$10.58
6 feet 1/16 x 1 in Angled aluminum bar	3-speaker mount base	2	$11.69	$24.62
8 feet 1 in square aluminum tube	Vertical speaker mount	1	$19.19	$20.21
8 feet 1 in angled aluminum tube	3 speaker mount	1	$25.56	$25.56
8 feet 1 in square aluminum tube	Vertical speaker mount	1	$22.73	$22.73
Flat aluminum speaker mount	1 speaker mount	1	$6.63	$6.63
Door knob assembly for swivel mount	Swivel mount for 3 speaker gate	1	$8.40	$8.40
PALAT Hardware				
Speaker wire, 16 guage 100 feet	Speaker wire	1	$18.92	$18.92
Speaker wire, 16 guage 50 feet	Speaker wire - speaker to connector	2	$7.49	$14.98
Wire connectors Bullet male and female	Wire connectors	1		$9.60
Banana plugs	Banana plugs	64		$21.00
Switch, 32 relay	32 relay switch	1		$179.95
Wire connector for switch	Switch wire connectors	6	$0.22	$1.39
Speakr wire, 18 guage 500 feet	18 guage speaker wire	1	$49.00	$49.00
Cambridge Audio Loudspeaker	Minz Min 12	68	$50.00	$3,400.00
Pink-noise generator	Mystic Marvels PNG-400	1	$99.00	$99.00
Amplifiers				
Stewart Audio AV30MX-2 amplifier (30 W)	Stewart Audio	2	$215.00	$430.00
Gator Case	Audio Rack Case	1	$129.00	$129.00
Acrylic sheets for relay switch case	Relay switch case	4	$3.22	$12.88
Power cable for audio case	Short power cables	4	$2.50	$9.99
Power cable for amplifier	Short amp power cable	2	$5.95	$11.90
Power supply	Audio case power supply	1	$104.99	$104.99
Rocker switch for pink noise control	Rocker switch pink noise	1	$10.41	$10.41
Connectors for rocker switch	Wire connectors	1	$3.78	$3.78
Heat Shrink tubing				
Heat Shrink tubing 6 inch black	Heat shrink	3	$6.43	$19.30
Heat Shrink tubing 8 ft	Heat shrink	2	$1.98	$3.96
USB connector	USB connector	1	$5.26	$5.26
Vinyl tubing for wire protector	Cord protector	1	$0.80	$0.80
Acrylic sheets for relay switch, masking tape	Relay switch case	1	$22.72	$22.72
Heat Shrink tubing 6 inch black	Heat shrink	2	$15.22	$15.22
Heat Shrink tubing 6 inch black	Heat shrink	1	$4.08	$4.08
8 feet 1 in square aluminum tube, 1/16th inch Aluminum tube	Vertical speaker mount	3	$19.24	$57.71
Power cable for audio case and amplifier	Short power cables	5	$4.38	$21.89
Speaker wire	16 guage speaker wire	1		$11.77
Speaker wire	16 guage speaker wire	1		$7.52
Wiring connector	Wiring connector	1		$3.78
Microsoft Surface Pro Tablet with stylus	Tablet with stylus	1		$1,000.00
Surface Pro case	Tablet case	1		$75.00
			Total	**$6,193.55**

Appendix E. Human Subjects IRB Documents

Virginia Tech IRB Approval Letter #11-047

Investigator IRB Human Subjects Training Certificate

Informed Consent Form for Proposed Experiment
(Phase II Experiment)

Virginia Tech IRB Approval Letter #11-047

Date*	OSP Number	Sponsor	Grant Comparison Conducted?
04/24/2017	PBMVOHIB	Office of Naval Research (Title: Innovative Portable Auditory Localization Acclimation-Test (PALAT) System for Military Applications, including Full-Scale and Portable Versions, Leveraged on DRILCOM Test Battery and Validated via an In-Field Experiment)	Compared on 04/24/2017
09/07/2016	P5XV6K57	HEAR LLC (Title: 1st Installment: HEAR Subcontract for TCAPS Training)	Not required (Not federally funded)
09/25/2015	PQYQHPCQ	Office of Naval Research (Title: Objective Comparison of Protection Effectiveness, Situation Awareness, and User Preference of Foam vs. Deep Custom-Molded Earplugs, Inclusive of User Training Implications)	Compared on 09/28/2015
09/30/2013	14051309	US Department of Defense (Title: Objective Metric-Based Assesments for Efficient Evaluation of Auditory Situation Awareness Characteristics fo Tastical Communications and Protective)	Compared on 10/01/2013
09/30/2013	13169402	Office of Naval Research (Title: Addition to Wind Noise Effect on In-Lab Auditory Detection and Identification of Acoustic Threat Signatures and on Speech Intelligibility using Augmen)	Compared on 10/01/2013
08/10/2012	11150201	Office of Naval Research (Title: In-Field Human Factors Evaluation of the Effects of Augmented Hearing Protection/Enhancement Devices (HPEDs) on Auditory Detection and Identification)	Compared on 01/20/2011

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.

Investigator IRB Human Subjects Training Certificate

VIRGINIATECHVIRGINIATECHVIRGINIATECHVIRGINIATECHVIRGINIATECH

Certificate of Completion

This certifies that

Brandon Scott Thompson

Has completed

Training in Human Subjects Protection

On the following topics:

Historical Basis for Regulating Human Subjects Research

The Belmont Report

Federal and Virginia Tech Regulatory Entities, Policies and Procedures

on

August 11, 2017

David Moore, IRB Chair

Phase II Informed Consent Form

VirginiaTech
College of Engineering

Grado Department of Industrial and Systems Engineering
A University Exemplary Department*
250 Durham Hall (0118)
Blacksburg, Virginia 24061
540/231-6656 Fax: 540/231-3322
E-mail: ise@vt.edu
www.ise.vt.edu

Virginia Tech Auditory Systems Laboratory:
Informed Consent for Participants
in Research Projects Involving Human Subjects

This research is funded at Virginia Tech by the Department of Defense's Hearing Center of Excellence.

Title of Project: Objective Metric-Based Assessments for Efficient Evaluation of Auditory Situation Awareness Characteristics of Tactical Communications and Protective Systems (TCAPS)

Investigators: Kichol Lee, Ph.D. Research Assistant Professor and Ear Acoustics Specialist, Auditory Systems Lab of ISE and John G. Casali, Ph.D., CPE, Grado Professor of ISE and Director, Auditory Systems Lab. (Dr. Lee will serve as the "Experimenter" noted throughout this document.) *The principal investigators of this research project, Dr. Casali and Dr. Lee, are also a co-founder of HEAR, the company that developed the testing software, which is a product used in this research project.*

Participants: You will be one of at least 15 and up to 30 participants. All participants are 18 years old or older with normal hearing or some minor level of hearing loss. This research involves predominantly male participants since the research has implications for U.S. military operations, where males outnumber females by approximately 4 to 1.

I. Purpose of this Research

The purpose of this research study is to assess auditory (hearing) situation awareness influences of various military tactical communications and protective systems (TCAPS) and hearing protection devices. Auditory situation awareness will be measured in four realistic tasks that involve listening and hearing: detection, recognition/identification, localization, and communication. This experiment is designed to simulate various tasks that constitute auditory situation awareness required by a soldier in military service. You, as the participant, will be asked to detect, recognize/identify, and localize the source of the sound as accurately and quickly as possible. You will also be required to conduct a communications test to measure the voice communication capability provided by each device.

II. Procedures

There will be up to, but not more than, 10 experimental sessions (totaling less than 20 hours) for all participants. Experimental sessions will occur in the Virginia Tech Auditory Systems Laboratory (ASL) located on the fifth floor of Whittemore Hall.

Initial Qualification/Training Session:

Before qualification testing, you will be asked to fill out three forms. First, an informed consent form (this form) will be shown to you. After you read the informed consent form, you can ask any questions related to the experiment. Then you will be asked to fill out a demographic form and a hearing protection use history form.

Qualification (Eligibility) Testing: The first test session will begin with audiometric qualification testing. Audiometric qualification testing will include 1) a standard hearing test, to determine your hearing sensitivity, 2) a visual inspection of your ear canal using a lighted otoscope, to determine if there are any obstructions, and 3) a history of your hearing protection use for the last six months. If you have impacted earwax or other ear canal problems, you will be asked not to participate, and perhaps to visit an ear health professional such as an audiologist (hearing specialist) or otolaryngologist (ear physician) to have

**University Exemplary Department Awards recognize the work of departments that maintain, through collaborative efforts of dedicated colleges exemplary teaching and learning environments for students and faculty.*

your ear canals clinically checked and cleaned of earwax if needed, and perhaps return for a second screening for this experiment if you so desire.

Upon successful completion of all eligibility requirements, you will:

1. Receive training/orientation with the actual hearing protection devices (HPDs) (including HPEDs-hearing protection enhancement devices and TCAPS-tactical communications and protection systems) that will be used in the study, as well as training on the experimental procedures.
2. Undergo training sessions for the 4 situation awareness sub-tests: detection, recognition/identification, localization, and communication. You will be trained without any HPDs (i.e., with open ears) for the 4 tests. Once you are familiarized with each test, you will also practice using all HPDs.

Throughout the training sessions and the four sub-tests described below, your responses will always entail only a mouse click entry or an answer by your voice (speaking) At the conclusion of each sub-test, you will rate and rank your experiences and impressions with the various HPDs via rating and ranking scales, explained below.

The Experimenter will fit all HPDs on you to reduce any variability that might be due to individual differences in fitting and wearing HPDs. The purpose of this study is to determine each HPD's capability in maintaining auditory situation awareness, and *not* to assess the fit of HPDs.

On each of the 4 sub-tests below, where sound signals or noises are presented to you, these will always be kept to sound exposure levels that are less than half that allowed by the U.S. OSHA (1983) workplace health and safety laws for a U.S. worker in a single work day (8-hour shift), in order to minimize any risk of noise-induced hearing loss. In this experiment, the signal and noise levels are known to be at decibel and time duration values that do not pose a risk to your hearing, even when you are listening with open ears. Specific noises will be discussed below where they are to be applied under each sub-test.

Sub-test 1: Detection test

For the Detection test, you will be asked to listen to a test signal and press and hold a response switch whenever you can hear the test signal. Pressing and holding the button will lower the test signal volume and you should let go of the switch as soon as you cannot hear the sound. After several repetitions, the computer program will cause the test to advance to the next signal. You will repeat the test with open ears and several HPDs, and also with the signals coming from loudspeakers that are located at front, back, left, and right of you.

The Detection test will be similar to a standard Real Ear Attenuation at Threshold test (REAT) (already approved by Virginia Tech's Institutional Review Board (IRB) for Human Subjects as VT IRB-05-701) with modifications as follows: 1) the test signal will only be played via 1 loudspeaker located at front, back, left, and right of you, and 2) the sound signal will consist of 1/3-octave bands with center frequencies of 125Hz, 250Hz, 500Hz, 1000Hz, 2000Hz, 4000Hz, 8000Hz (same as the standard REAT test signals), plus two military-relevant signals that are low-intensity gunshots and gun-cocking sounds. Because this will be a detection test at or very near auditory threshold (just barely heard), the sound levels of the gunshots and gun-cocking presented to the you will be at very quiet levels, and you will have to be extremely attentive and listen carefully to hear them. Your response (pressing a response switch and holding it pressed at any time when you <u>can</u> hear the signal) will be tracked by a computer program and recorded for later analysis. The intensity level of the background masking noise (pink noise, which includes all pitches of sound) will always be played at 85 dBA or below. (As mentioned above, note that OSHA (1983) allows U.S. workers to be exposed to an average of 90 dBA for a full 8-hour workday, and throughout the subtests described herein, all noise exposures will be less than half of that amount per day of this experiment.)

Sub-test 2: Recognition/Identification (Re/Id) test

In this test, you will be asked to listen and identify a particular (target) sound clip from the 3 sound clips presented in rapid sequence to you, and to identify that target sound as accurately and quickly as possible.

The Re/Id test will consist of listening to series of sound signals presented in groups of three and identifying the requested target sound. The sound signals will be played via 1 loudspeaker located in front and to the right of you. You will be asked to click one of the four buttons that identifies the target sound on your monitor as accurately and quickly as possible. You will also determine, through the button on your monitor, when the next set of sounds is to be played. The background masking noise (pink noise) level will always be played at 85 dBA or below, as described previously, which is considered a safe level. The test sound signals will always be played at 60-80 dBA throughtout the test. Some examples of test sound signals are: speech, truck engine idling, gun-cocking, footstep, rifle firing, heavy truck breaking.

Sub-test 3: Localization test
In this test, you will be asked to listen to test sound signals presented to you while sitting in the center of a circle of loudspeakers and to identify where the test signal is coming from, in other words, which loudspeaker is playing that signal.

The exact location and number of loudspeakers will be hidden from your view. Possible loudspeaker locations (targets) will be shown in your monitor and you will identify (click) the loudspeaker icon which corresponds to where you think the test signal came from. You will listen for a test sound signal and identify the absolute location in either 360° azimuth (horizontally around you) or in frontal elevation (level or upward in front of you) as accurately and quickly as possible. You will also signal, through the button on your monitor, when the next test sound signal is to be played.

Your response will be recorded by a computer program for later analysis. The background masking noise (pink noise) level will always be played at 85 dBA or below, as described previously. The sound signal for the localization test will be a sound that includes both low and high frequency ranges that are well-within the pitches of sound that can be heard, and localized, by the human ear. The test sound signal will be played at a level slightly higher than the background masking sound (not more than 10 dB higher) in order to be audible. The duration of each test signal will be very short, that is, 4 seconds or less, so the total noise exposure will be well below that allowed by OSHA for U.S. industrial workers, as described previously.

Sub-test 4: Communication test
In this test, you will be asked to listen to sentences of prerecorded speech and verbally repeat the sentences exactly as you heard them. The Experimenter will notify you before each prerecorded sentence is played.

The Communication test will be a modified version of the "QuickSIN" test, which is a standardized test used for testing people's ability to understand "Speech in Noise" (SIN). The modified QuickSIN test will be presented through loudspeakers in the test room instead of through a headphone which the original test was designed to use. Your responses will be entered into a computer by the Experimenter for later analysis. Because the QuickSIN test includes masking sound as part of test presentation, there will not be an additional background masking sound as will be used in the other 3 tests.

III. Risks

Experimental purpose: This experiment is designed to measure levels of auditory situation awareness afforded by each military hearing protection device (HPD) when a listener is required to detect, recognize/identify, and localize a sound source as well as to hear and repeat sentences spoken to him/her. The sound sources (test signals) and communication sentences are not loud enough to be hazardous, and there is no known bodily danger associated with this study. Furthermore, all of the signals are recorded sounds, and thus there are no "real" sources of the signals in the experimental room, such as gunshots. However, if you feel that this experiment would make you uncomfortable or cause you emotional distress during or after the experiment, you may freely decide not to participate.

Hearing Protection Devices (HPDs): HPDs are designed to have a tight fit and you may experience some minor discomfort while wearing them. If you experience more than minor discomfort, tell the Experimenter immediately and he will assist you in adjusting or removing the hearing protector or will provide a different size of eartip. Also, electronic protectors may emit a squealing or whistling noise if not properly sealed; while not dangerous it can be annoying, if this occurs please notify the Experimenter so that he can adjust the seal of the device. This can also occur if you accidentally place your finger over the microphone, so do not reach up and touch the device. In all cases, the Experimenter will fit the devices in or over your ears, and adjust the gain-amplification setting to help avoid the possibility of a squealing or whistling noise.

Other Risks: If you feel tired, or become thirsty during the tests, please inform the Experimenter and you will be allowed to rest and have something to drink.

Your Responsibilities: If you consent to participate, and later, if, in the unlikely event that you seek medical or counseling services that you feel are a result of your participation, you will be responsible for the costs of such services.

IV. Benefits

Your participation in this experiment will provide information on the level of auditory situation awareness afforded by advanced hearing protection devices. This information will *primarily* be used to help the U.S. military to determine what hearing protection devices that an individual should use for situations in which they need to hear and communicate. This information may also be of use for selection of protectors for certain law enforcement, industrial, construction, or hunting applications. No promise or guarantee of benefits has been made to encourage you to participate; however, you will receive the monetary compensation that is covered below in Section VI.

V. Extent of Anonymity and Confidentiality

Your identity will be kept confidential. If you choose to participate in the experiment, you will be identified by only a participant number. This number will be used in data collection and analysis. At no time will the researchers release your identity to anyone other than individuals working on the project without your written consent. This Consent Form as well as the raw data from the experiment (for example, your detection, identification, localization, and communications responses, and your experience ratings) will be solely in the possession of the investigators. Data will be analyzed on the investigators' Virginia Tech computers, which are password-protected.

It is possible that the Virginia Tech Human Participants Institutional Review Board (IRB), or a U.S. Military IRB, may review this study's collected data for auditing purposes. An IRB conducts the oversight of the protection of human subjects involved in research.

VI. Compensation

Participants will be monetarily compensated for participation in the study at the rate of $10 per hour during test/training sessions. For any fraction of time less than 1 hour, you will be paid for the closest ¼-hour period, rounded up (in your favor). You will also receive a $20 bonus for successful completion of all experimental sessions. You will be paid at the conclusion of each experimental session. Military participants (active duty, national guard, or reserve) are not eligible for study-related payment unless they are on leave status.

VII. Freedom to Withdraw

You are free to withdraw from this study at any time without penalty. If you choose to withdraw from the study, you will be compensated for the portion of your time spent in the study. There may be circumstances under which the investigator may determine that you should not continue as a subject, and while this is a rare occurrence, you must abide by that decision if it occurs. Again, you would be paid for the time that you have spent in the experiment under any circumstances.

VIII. Subject's Responsibilities

I voluntarily agree to participate in this study.
I have the following responsibilities:

- To listen for, detect, identify, and localize the signals, and to listen for and repeat the communications sentences in the experiment to the best of my ability, and to provide accurate ratings of my impressions about the listening conditions.
- To inform the Experimenter if a hearing protector, or any other aspects of the test condition, becomes overly uncomfortable.
- To inform the Experimenter if I become tired or thirsty and wish to rest.
- To avoid biasing other potential participants, to not discuss the study with anyone until 6 months after the day of my participation.

IX. Participant's Permission

I have read the Consent Form and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent:

______________________ ________________________ _______
Participant's printed name Participant's signature Date

Age________

******Participant's Tear-Off Portion—Participant to Keep This******

Contact information for investigators:

John G. Casali, Ph.D. (Principal Investigator)
(540) 231-5073 email: jcasali@exchange.vt.edu

Kichol Lee, Ph.D.
(540) 231-3294 email: kichol@exchange.vt.edu

Should you have any questions or concerns about the study's conduct or your rights as a research participant, or need to report a research-related injury or event, you may contact the Virginia Tech Institutional Review Board at irb@vt.edu or (540) 231-3732.

Appendix F. Phase II Participant Questionnaire

Participant #: ________ Date: ________________

HPD/Listening Condition-DRILCOM: ______________________________

Instructions: Please circle a number to best describe your selection.

1. Rate how **training** using the DRILCOM system **impacted your confidence** in your ability to localize sounds, from **before to after** all the training you received using this system.

EXTREMELY LESS CONFIDENT	-3	-2	-1	0	1	2	3	EXTREMELY MORE CONFIDENT

2. Rate the **impact** you felt the **proximity (distance) of the loudspeakers** of the DRILCOM system contributed to your **ability to train to localize** sounds.

EXTREMELY NEGATIVE IMPACT	-3	-2	-1	0	1	2	3	EXTREMELY POSITIVE IMPACT

3. Rate how **easy it was to operate** the DRILCOM system hardware and software during your localization training.

EXTREMELY DIFFICULT	-3	-2	-1	0	1	2	3	EXTREMELY EASY

4. Rate the **impact** you felt the **room environment** of the DRILCOM system contributed to your **ability to train to localize** sounds.

EXTREMELY NEGATIVE IMPACT	-3	-2	-1	0	1	2	3	EXTREMELY POSITIVE IMPACT

5. Rate how much you feel your **ability** to determine sound location improved as a **result of training** with this system**.**

EXTREMELY LESS CAPABLE	-3	-2	-1	0	1	2	3	EXTREMELY MORE CAPABLE

6. Rate how **difficult** it was to judge the **location** of the sounds **using this system**.

EXTREMELY DIFFICULT	-3	-2	-1	0	1	2	3	EXTREMELY EASY

7. Rate how **training** using the DRILCOM system **impacted your reaction time** in determining sound location, from **before to after** all the training you received using this system.

EXTREMELY SLOWER REACTION TIME	-3	-2	-1	0	1	2	3	EXTREMELY FASTER REACTION TIME

8. Rate how much of an **impact** the DRILCOM system **user interface** (monitor, software, loudspeakers, wires, etc.) had on your **ability to train your sound localization skills**.

EXTREMELY NEGATIVE IMPACT	-3	-2	-1	0	1	2	3	EXTREMELY POSITIVE IMPACT

9. Rate how training in the **room environment** of the DRILCOM system **impacted your reaction time** in determining sound location.

EXTREMELY SLOWER REACTION TIME	-3	-2	-1	0	1	2	3	EXTREMELY FASTER REACTION TIME

10. Rate the **impact** you felt the **hidden loudspeakers** of the DRILCOM system contributed to your **ability to train to localize** sounds.

EXTREMELY NEGATIVE IMPACT	-3	-2	-1	0	1	2	3	EXTREMELY POSITIVE IMPACT

For the following questions, please compare the DRILCOM system with your previous training using the PALAT system.

11. Compared to the PALAT system, rate how **confident you are in your ability** to localize sounds **using the DRILCOM system**.

EXTREMELY LESS CONFIDENT	-3	-2	-1	0	1	2	3	EXTREMELY MORE CONFIDENT

12. Compared to the PALAT system, rate how much the **user interface** (monitor, software, loudspeakers, etc.) of the DRILCOM system **impacted** your **ability to train to localize** sounds.

EXTREMELY NEGATIVE IMPACT	-3	-2	-1	0	1	2	3	EXTREMELY POSITIVE IMPACT

13. Compared to the PALAT system, rate how much of an **impact training** with the DRILCOM system had on your **ability to localize sounds**.

EXTREMELY NEGATIVE IMPACT	-3	-2	-1	0	1	2	3	EXTREMELY POSITIVE IMPACT

14. Compared to the PALAT system, please rate how much the **room environment** of the DRILCOM system **impacted** your **ability to localize sounds**.

EXTREMELY NEGATIVE IMPACT	-3	-2	-1	0	1	2	3	EXTREMELY POSITIVE IMPACT

Rate your preference between the DRILCOM (large room) system and the PALAT (small room) system on the each of the following aspects. Ratings to the Left of 0 would indicate strength of preference for DRILCOM, and to the Right of 0 would indicate strength of preference for PALAT.

15. **Confidence** in **<u>accurately</u>** localizing the sounds:

DRILCOM	-3	-2	-1	0	1	2	3	PALAT

16. **Confidence** in making **<u>quick</u> decisions** (reaction time) about the location of the sounds:

DRILCOM	-3	-2	-1	0	1	2	3	PALAT

17. **Preference** as to the **room environment** for training for sound localization:

DRILCOM	-3	-2	-1	0	1	2	3	PALAT

18. Preference as to the loudspeaker configuration and proximity:

DRILCOM	-3	-2	-1	0	1	2	3	PALAT

19. **Preference** as to the **user interface** for responding to the location of the sound:

DRILCOM	-3	-2	-1	0	1	2	3	PALAT

20. **Confidence** in the benefits achieved with the training for sound localization:

DRILCOM	-3	-2	-1	0	1	2	3	PALAT

Please provide any additional comments about your localization training using the DRILCOM system. Think about strong points and weak points of the DRILCOM system in formulating your answer.

Appendix G. Phase II Screening Form

Initial Screening Questionnaire

Participant ID: _________ Sex: M F Age: ____________

1) Hearing level requirements:

Pass Fail	25 dBHL or better in both ears at 250, 500, 1000, 2000, 4000, 6000, and 8000 Hz
Pass Fail	No bilateral asymmetry of greater than 15 dB
Pass Fail	Otoscopic inspection

Yes No	2) Have you had any prior experience with any military, law enforcement, or industrial HPD or TCAPS which has a pass-through communication feature?
Yes No	3) Have you had any prior experience with military, law enforcement or similar "game" training in tactical localization, identification, and/or elimination of threats, specifically threats that are recognizable by the sound they make? If so, what experience did you have?

Appendix H. Phase II Recruiting Flyer

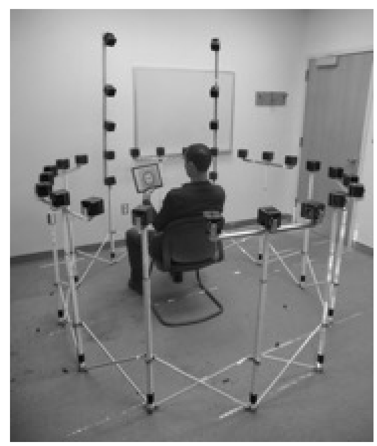
SOUND LOCALIZATION STUDY
PARTICIPANTS NEEDED

Title: Evaluation of a Portable Auditory Situation Awareness Training system.

Requirements:

- 18-45 years old
- No prior experience with sound localization studies
- Must pass a hearing test administered in the VT Auditory Systems Laboratory as part of study

Experiment Details:



- Participants will train to localize sounds in a lab setting
- Sound localization tests will measure performance using military-type sounds
- Participants may be trained and tested while wearing military Hearing Protection Devices
- 6 training sessions, 2 hours each
    - (Training can be completed within 1-2 weeks)
- Compensation: $10/hour + $25 completion bonus

Participation is voluntary and confidential. Research protocols in this experiment have been approved by the VT Institutional Review Board (IRB #11-047).

Appendix I. Phase II Script

PALAT instructions

Introduction to the study:
The purpose of the study is to test and evaluate the effectiveness of two sound localization training systems. You are going to hear an unusual sound similar to a buzz coming from one of 12 speakers. The speakers are arranged in a clock face, so the one directly in front of you is 12 o'clock. Every other speaker in the array corresponds to one of the positions on the clock. Only the 12 clock-face speaker locations will be used for training and testing. Your basic task is to use the tablet pen and touch screen to indicate where you think you heard the sound coming from. You will undergo one familiarization task, a pretest, and then 5 learning blocks with four sub-blocks in each: 1) Sound signals played in a sequential clockwise or counterclockwise pattern, 2) Sound signals played from speakers in a randomized order, 3) User-choice where you, the Participant, can select the speaker location to play the sound signal, and 4) a test.

Show them the hard copy of what a correct and incorrect answer looks like.

You may take a break or withdraw at any time.

For ITE listening condition
At this time, I will insert the TEP-100 hearing protection device in each ear and turn on the devices. The TEP-100 are rechargeable electronic earplugs that reduce loud noise and amplify low level sounds. You will be training and testing using a "unity" gain level setting on the TEP-100.
(Fit the TEP-100 and ensure the "unity" gain setting is selected)

For OTE listening condition
At this time, I will fit you with the COMTAC III hearing protection device. The COMTAC III is electronic tactical communication headset that reduces loud noise and amplifies low level sounds. You will be training and testing using a "unity" gain level on the COMTAC III.
(Fit the COMTAC III and ensure the "unity" gain setting is selected)

For ALL CONDITIONS

<u>Familiarization</u>

You're going to hear sounds coming from 12, 3, 6, and 9 in order. Go ahead and begin by selecting the green button in the center of the screen. This is self-paced, so the task today will pace according to when you select the button. **Please respond as quickly and accurately as possible.** You may select any of the 24 speaker locations that you think best represents the location where the sound signal originated. The first sound you will hear is coming from 12 o'clock.
*12, 3, after 6, tell them to respond to the last sound signal with the 2 o'clock position to demonstrate what an incorrect answer will look like.

Pretest

Good, now we are going to do a pre-test. You won't know where the sound is coming from and you won't receive feedback regarding your results. **Just respond as quickly and accurately as possible.**

LU1

Now you are going to do what we call Sequential clock-face presentations. You'll hear the sound coming from 12 o'clock then it will move in counterclockwise order to 11 o'clock.
-Then, the sound will come from 9 o'clock and move to 10 in a clockwise order.
-Then the signal will start at 3 o'clock and move counterclockwise to 2.
-Then the signal will start at 6 and move clockwise to 7.

Random

For this block, you will not know where the sound is coming from, but you will receive feedback regarding your results. Please respond as quickly and accurately as possible.

User-choice

This next session is for you to choose to re-play some of the sound locations you wish to hear again. You'll see a counter in the upper left corner that will tell you how many trials you have remaining (18 total).

Test

For this block you will not know where the sound is coming from and you will not receive feedback regarding your results. The screen is going to look a little different as you will have 24 instead of 12 choices that will look like this *(show them the hard copy)*

This completes the first learning unit. You will now complete 4 more learning units. You are free to take a break at any time during the training. However, there will be a planned break after Learning Unit 2 (LU2).

LU2-LU5

You will now complete 4 more learning units.

Following LU5 you will receive a questionnaire to rate your training experience.

*Give questionnaire after LU5.

Appendix J. Phase III IRB Approval

VIRGINIA TECH.

Division of Scholarly Integrity and Research Compliance
Institutional Review Board
North End Center, Suite 4120 (MC 0497)
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-3732
irb@vt.edu
http://www.research.vt.edu/sirc/hrpp

MEMORANDUM

DATE:	May 30, 2019
TO:	John Casali, Kichol Lee, Kara Meghan Cave, Brandon Scott Thompson
FROM:	Virginia Tech Institutional Review Board (FWA00000572, expires January 29, 2021)
PROTOCOL TITLE:	Innovative Portable Auditory Localization Acclimation-Test System (PASAT) for Tactical Communications and Protective Systems (TCAPS) in Military and Other Applications, including Full-Scale and Portable Versions, Leveraged on DRILCOM Test Battery and
IRB NUMBER:	**19-176**

The Virginia Tech Institution Review Board (IRB), acknowledges the Amendment request for the above-mentioned research protocol.

This acknowledgement recognizes the item(s) identified in the Special Instructions section.

NOTE: Amendments that must be submitted to BRANY for review and approval include changes to funding, conflict of interest, ANY and ALL changes to study procedures and study documents. If your study qualified for Not Human Subjects or for an Exemption please review the information at the end of your approval Letter.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

IRB Number 19-176 page 2 of 2 Virginia Tech Institutional Review Board

SPECIAL INSTRUCTIONS:

The Virginia Tech IRB acknowledges the transfer of IRB oversight from WIRB to BRANY for this protocol. Please read the information below for more details.

Dear Investigators:

This email serves as a notice that your protocol is under active transfer from WIRB to BRANY. We ask that you do not submit any further requests to WIRB or to the Virginia Tech IRB.

FAQ's:

Q. How will I know when my protocol has been accepted by BRANY?

A. BRANY IRB will send you a notification indicating your transferred study has been accepted.

Q. Has the Virginia Tech IRB drafted guidance?

A. Yes. We have created guidance and it is available on a PID protected website.

https://internal.research.vt.edu/sirc/hrpp/brany-transfer

This link will be provided on all Authorization Letters.

Q. How do I gain access to BRANY's IRBManager?

A. This section is very important. Not everyone listed as study personnel needs to have access to IRBManager. The PI, active Co-I(s), and study coordinators are the typical research team members that will need to have access. In order to gain access, each person will need to complete the Request for User Access form and sign it with wet ink. Digital signatures and script style font are not accepted. [http://www.brany.com/wp-content/uploads/2018/07/BRANY-User-Access-Form-complete-sign-return-2
0170323-V2_.pd]

Q. I need to submit an amendment to my protocol. What should I do?

A. Once your protocol has been accepted in the BRANY IRBManager system, you will be able to submit requests directly to BRANY for review in their IRBManager system. Refer to the guidance provided by the Virginia Tech IRB using the web link above.

Q. I need to revise my list of study personnel. What should I do?

A. You will no longer submit personnel changes to the Virginia Tech IRB. You will submit personnel changes to BRANY through their IRBManager system. You should follow the guidance provided by the Virginia Tech IRB using the web link above.

Q. I am actively working with research subjects (including recruiting, consenting, enrolling, collecting data). Do I need to alter my consent forms? Do I need to notify my participants of the change of IRB oversight?

A. This section is very important. When you receive the notification from BRANY that your study has been accepted, instructions will be included regarding consent needs.

Q. How will I know that I need to submit a Continuing Review request?

A. BRANY will send reminder emails 45, 30, and 15 days prior to the expiration date. The automated reminders will cease when either a continuing review or closure application is received and processed by BRANY IRB.

Date*	OSP Number	Sponsor	Grant Comparison Conducted?
03/14/2019	PBMVOHIB	Office of Naval Research (Title: Innovative Portable Auditory Localization Acclimation-Test (PALAT) System for Military Applications, including Full-Scale and Portable Versions, Leveraged on DRILCOM Test Battery and Validated via an In-Field Experiment)	Compared on 03/14/2019

* Date this proposal number was compared, assessed as not requiring comparison, or comparison information was revised.

If this protocol is to cover any other grant proposals, please contact the HRPP office (irb@vt.edu) immediately.

Appendix K. Phase III Questionnaire

IRB# 19-176

Participant #: P20 Date: 4/14/19

Questionnaire II

HPD: Open ear Session: Field test

Instructions: Please click on the button under the number to best describe your selection.

1. Rate how **confident you were** in your ability to locate sounds **under this listening condition.**

NO CONFIDENCE	1	2	3	4	5	6	7	EXTREMELY CONFIDENT

2. Rate your perceived **accuracy** to determine sound location **under this listening condition.**

HIGHLY INACURATE	1	2	3	4	5	6	7	HIGHLY ACCURATE

3. Rate how **difficult** it was to judge the location of the sounds **under this listening condition.**

EXTREMELY DIFFICULT	1	2	3	4	5	6	7	EXTREMELY EASY

4. Rate your perceived **reaction time** in determining the sound location **under this listening condition.**

EXTREMELY SLOW	1	2	3	4	5	6	7	EXTREMELY FAST

5. Please rate how **comfortable the open ear condition** was while performing this experiment.

EXTREMELY UNCOMFORTABLE	1	2	3	4	5	6	7	EXTREMELY COMFORTABLE

IRB# 19-176

6. How likely would you be to **keep your ears open (no hearing protector)** during a task similar to this experiment that required sound localization if you had access to a hearing protection device.

EXTREMELY UNLIKELY	1	2	3	4	5	6	7	EXTREMELY LIKELY

7. Rate the degree of preparedness you felt as a result of the training on the localization system (ring of loudspeakers) compared to the **task of localizing .22 blank gunshot sounds.**

EXTREMELY UNPREPARED	1	2	3	4	5	6	7	EXTREMELY PREPARED

Please provide any additional comments about your localization testing using the localization system.

Appendix L. Phase III Screening Form

Localization Questionnaire- IRB 19-176

6/6/18

Initial Screening Questionnaire

Subject ID: ____________ Sex: M F Age: ____________

1) Hearing level requirements:

Pass Fail 25 dBHL or better in both ears at 250, 500, 1000, 2000, 4000, 6000, and 8000 Hz

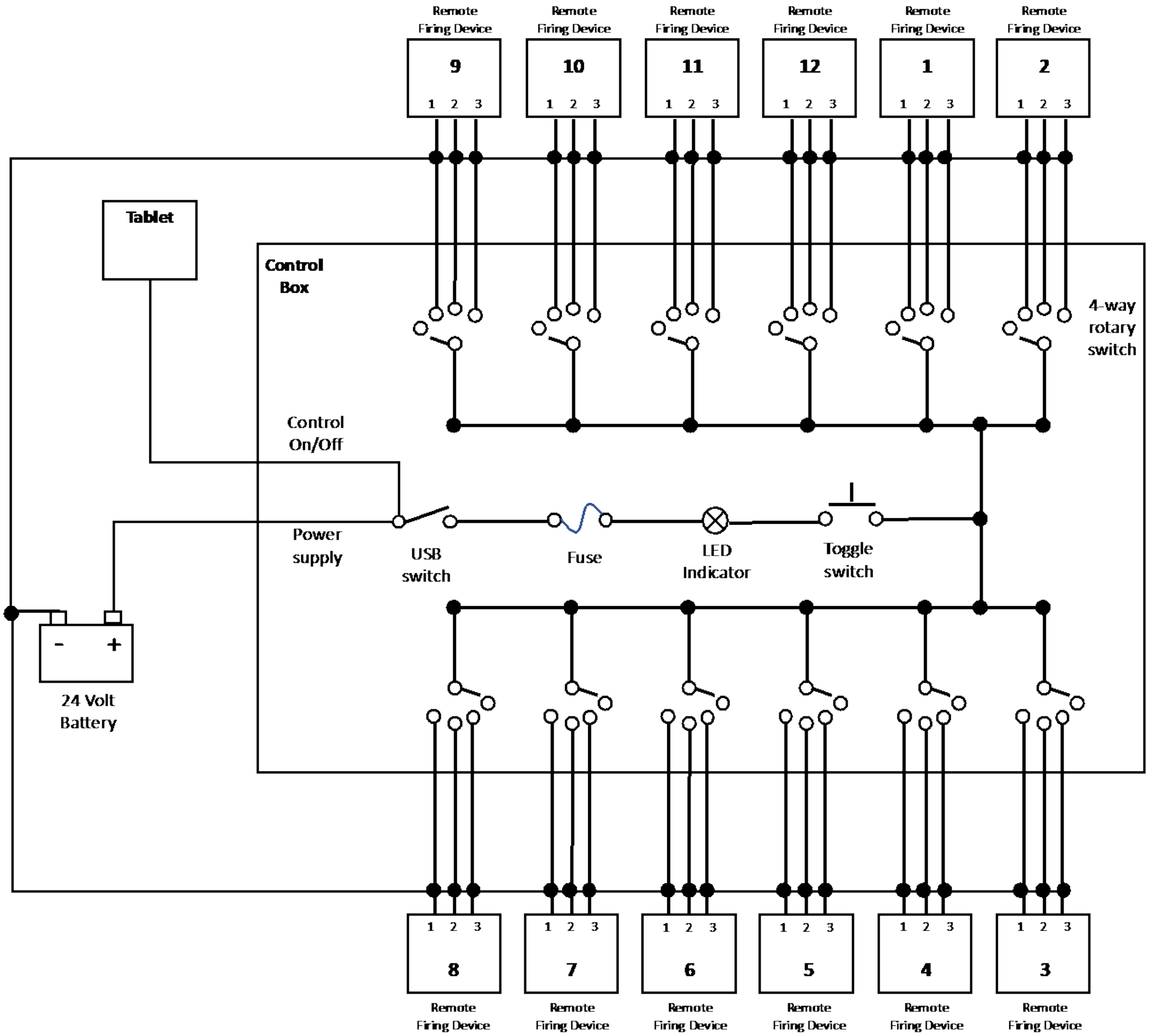
Pass Fail no bilateral asymmetry of greater than 15 dB

Pass Fail Otoscopic inspection

Yes No 2) Have you had any prior experience with any military, law enforcement, or industrial HPD or TCAPS which has a pass-through communication feature?

Yes No 3) Have you had any prior experience with military, law enforcement or similar "game" training in tactical localization, identification, and/or elimination of threats, specifically threats that are recognizable by the sound they make? If so, what experience did you have?

Appendix M. Remote Firing Device Wiring Diagram



Appendix N. Phase III Participant Flyer

#24088302.0

IRB# 19-176

IRB Approved at the
Study Level
Mar 22, 2019

PARTICIPANTS NEEDED

In a sound localization learning study for military-relevant sounds.

Requirements:

- 18-45 years old
- No prior experience with sound localization studies

Experiment Details:

- Must pass a hearing test in the VT Auditory Systems Laboratory
- Participants will be required to perform sound localization tasks of military-type sounds in a lab setting and in a field setting (about 45 minutes away from VT campus)
- ~6-9 hours total, spread over two to five days
- Compensation: $10/hour for all time spent and $30 bonus for completion of all experimental sessions

Please email Kara Cave if interested:
karacave@vt.edu

Appendix O. Phase III Informed Consent

IRB APPROVED
Mar 22, 2019

RESEARCH SUBJECT CONSENT FORM

Title:	Innovative Portable Auditory Localization Acclimation-Test System (PALAT) for Tactical Communications and Protective Systems (TCAPS) in Military and Other Applications, including Full-Scale and Portable Versions, Leveraged on DRILCOM Test Battery and Validated via an In-Field Experiment (VT OSP #450489)
Protocol No.:	19-176 WIRB® Protocol #20190789 19-176
Sponsor:	Virginia Polytechnic Institute and State University
Investigator:	John G. Casali, PhD, CPE 250 Durham Hall (0118) Blacksburg, Virginia, 24061 USA
Daytime Phone Number:	540-231-5073
24-hour Phone Number:	(904) 307-8144

RESEARCH CONSENT SUMMARY

You are being asked for your consent to take part in a research study. This document provides a concise summary of this research. It describes the key information that we believe most people need to decide whether to take part in this research. Later sections of this document will provide all relevant details.

What should I know about this research?

- Someone will explain this research to you.
- Taking part in this research is voluntary. Whether you take part is up to you.
- If you don’t take part, it won’t be held against you.
- You can take part now and later drop out, and it won’t be held against you
- If you don’t understand, ask questions.
- Ask all the questions you want before you decide.

How long will I be in this research?

We expect that your taking part in this research will last up to approximately nine hours spread over five sessions, spread out over the span of a week (seven days). The experiment could be as short as approximately six hours spread over two days, spanning a week.

Why is this research being done?

The purpose of this research is to assess the benefits of training using a portable auditory, or hearing, localization training system on localization performance in an outdoor environment. The experiment is designed to simulate a scenario where a military service member is listening for gunshots from a long distance. You, the participant, will need to locate, via hearing, where the simulated gunshot is located. Another purpose to the experiment is to determine if training affects localization accuracy without hearing protection (that is, with open ears) and with certain types of hearing protection.

What happens to me if I agree to take part in this research?

If you decide to take part in this research study, the general procedures include at least two stages and possibly a third. In the first stage, a screening session will occur that involves filling out a demographic questionnaire, looking in your ears with an ear microscope, and taking a brief hearing test. The first session will also involve a test to determine how well you locate sounds. The sound used in this session will seem similar to a beep, played at a moderate (not loud) level. Some participants will be asked to take part in three additional training sessions in the same location as the first session. All participants will be asked to participate in a sound location test in a field environment about 45 minutes from the Virginia Tech Campus. This field session will take place five to seven days after the first session. The sound used in the field session is a blank simulated gunshot. No "live" ammunition will be used. The blanks will be fired from a fabricated device specifically made for this purpose and located at least 150 feet away from you. Should the fabricated device fail, a starter pistol, used at sporting events to mark the start of a competition such as a foot race, will be used. You will be asked to listen with your open (uncovered) ears and with two different types of hearing protection.

Could being in this research hurt me?

The most important risks or discomforts that you may expect from taking part in this research include discomfort from hearing blank simulated gunshots. The experiment is designed to simulate a scenario where a soldier must listen for shots coming from a long distance away. If you feel this scenario would make you uncomfortable, please do not participate. Exposure to the noise from the blanks will not be loud enough to create a risk to your hearing, in that they will be well below that which is governed by the Occupational Safety and Health Administration (OSHA). OSHA allows peak exposures to be up to 140 dB peak sound pressure level. The blanks in this study will not exceed 113 dB peak sound pressure level. OSHA also allows average sound pressure levels to be up to 90 dBA for up to an 8-hour day, at which level the use of hearing protection becomes mandatory. The average levels in this experiment are at less than 85 dBA, and will not be presented for more than 4 hours per day. You will be given an opportunity to observe the blank ammunition and the firing device (starter pistol or fabricating device as applicable).

Hearing protectors are designed to have a tight fit and you may experience some minor discomfort while wearing them. If you experience more than minor discomfort, please tell the experimenter and he/she will assist you in adjusting or removing the hearing protector.

If you feel tired or become thirsty during the test, please inform one of the experimenters. You may take a break at any time. Water is will be made available to you.

Will being in this research benefit me?

The most important benefits that you may expect from taking part in this research include information on the ability to learn to localize sounds with the open ear and while wearing different hearing protectors. No promise or guarantee of benefits have been made to encourage you to participate. It is not expected that you will personally benefit from this research.

Possible benefits to others include assisting the military and law enforcement to determine the effects of auditory localization training and testing using hearing protection devices. The data in this study will also be used in fulfillment of two dissertations in human factors engineering at Virginia Tech.

What else should I know about this research?

Other information that may be important for you to consider so you can decide whether to take part in this research is that one session will take place in an outdoor field, located about 45 minutes away in Pulaski County. The session that will take place in the field will span about 3-4 hours, including transportation time.

DETAILED RESEARCH CONSENT

You are being invited to take part in a research study. A person who takes part in a research study is called a research subject, or research participant.

What should I know about this research?

- Someone will explain this research to you.
- This form sums up that explanation.
- Taking part in this research is voluntary. Whether you take part is up to you.
- You can choose not to take part. There will be no penalty or loss of benefits to which you are otherwise entitled.
- You can agree to take part and later change your mind. There will be no penalty or loss of benefits to which you are otherwise entitled.
- If you don’t understand, ask questions.
- Ask all the questions you want before you decide.

Why is this research being done?

The purpose of this research is to assess training effects of an indoor Portable Auditory Localization Acclimation-Test System (PALAT) by testing participants’ ability to determine the directions from which sounds are coming (sound localization) in an outdoor field setting. Furthermore, the experimenters are attempting to determine if sound localization ability is affected by the wearing of certain types of hearing protection. The experiment was designed to evaluate a training method for sound localization for later use in a military population.

About 24 subjects will take part in this research.

How long will I be in this research?

We expect that your taking part in this research will last approximately 9 hours. Each participant will complete at least two sessions. The first session will last approximately two hours. All participants will be asked to complete the field session that will last from three to four hours. Some participants will be randomly selected to complete training. Training take place in three sessions lasting approximately an hour each for each session, totaling three hours spread over no more than five days. After no more than seven days from the first session, a field test will be conducted (three-four hours). Therefore, the time commitment is expected not to exceed a week.

What happens to me if I agree to take part in this research?

Before audiometric testing, you will be asked to review and sign an informed consent form. Two copies of the informed consent form (this form) will be provided to you upon arrival at the initial screening and training session (as applicable): one signed copy will be maintained by the researchers and one copy is for you, the participant. This informed consent form is the same form that was provided to you, and each participant, via email upon volunteering for the study. After you read the informed consent form, you can ask any questions related to the experiment. You will be put into a study group by chance (like a coin toss/ like drawing straws). You have a 1 out of 2 chance of being placed in each group. You cannot choose your study group.

<u>Audiometric Qualification (Eligibility) Testing</u>
The screening and training stage will begin with audiometric qualification testing. Audiometric qualification testing will include 1) a standard hearing test, to determine your hearing sensitivity, 2) a visual inspection of your ear canal using a lighted otoscope, to determine if there are any obstructions, and 3) a history of your hearing protection use and localization study/training experience. If you have impacted earwax or other ear canal problems, you will be asked not to participate, and perhaps to visit an ear health professional such as an audiologist (hearing specialist) or otolaryngologist (ear physician) to have your ear canals clinically checked and cleaned of earwax if needed, and perhaps return for a second screening for this experiment if you so desire. The audiometric test and visual inspection will be conducted by an Active Duty U.S. Army Audiologist. You will be informed if you met the hearing eligibility requirements to complete the experiment.

<u>Screening session</u>
All participants will be asked to complete this session. Upon successful completion of all eligibility requirements, you will:

1. Receive familiarization training/orientation with the actual hearing protection devices (HPDs) known as Tactical Communication and Protective Systems (TCAPS) that will be used in the study.
2. Undergo a auditory localization pretest using the Portable Auditory Situation Awareness Training (PALAT) system. During the pretest, you will be asked to listen to a series of 36 sound signals (beeps) presented to you while sitting in the center of a circle of 12 loudspeakers and to identify and respond as accurately and quickly as possible with the direction you perceived the sound. Your responses will be recorded on a computer tablet by a

computer program for later analysis. The background masking noise (pink noise) level will always be played at 55 dBA or below, which is a quiet level. The sound signal for the localization test will be a sound that includes both low and high frequency ranges that are well-within the pitches of sound that can be heard and localized by the human ear. The test sound signal will be played at 70 dBA for one second in length (this is not loud, but well below sound levels allowed by the Occupational Safety and Health Administration (OSHA) for U.S. industrial workers).

Training sessions

Half of the participants will undergo training sessions. The investigator will tell you if you are assigned to this group. Auditory localization training will occur over a period of 3 sessions, each session lasting approximately 1 hour. You will be asked to perform the auditory training and testing under three listening conditions: open ear (no hearing protection device), wearing an in-the-ear TCAPS, and an over-the-ear TCAPS. The experimenter will fit the hearing protectors in (earplugs) or on (earmuffs) your ears. After each training session you will be asked to fill out a questionnaire rating your confidence in ability to localize the sound signal.

Field session:

All participants will be asked to complete this session. The experimental stage will take place in a rural field in which you will stand in surrounded by a wood forest in which two to three experimenters will be located. The experimenters will initiate blank gunshots from at least 150-feet away, that is, ammunition that is not "live" and has no bullet, from a device designed specifically for this purpose -- that is, it is not an actual weapon or gun. This equipment is similar to perimeter alarms used to contain livestock. You will be able to inspect the device that fires the blanks, as well as the blanks, before the start of the experiment. After each shot, you will be asked to verbally identify one of 24 numbered signs that corresponds most closely to the direction (location) you think the shot was fired from. There will be three listening conditions: open ear (no hearing protection device), wearing an in-the-ear TCAPS, and an over-the-ear TCAPS. The experimenter will fit the hearing protectors in (earplugs) or on (earmuffs) your ears. After each localization test you will be asked to fill out a questionnaire rating your confidence in ability to localize the blank gunshot signals and your impressions about the TCAPS.

What are my responsibilities if I take part in this research?

If you take part in this research, you will be responsible to:

- Listen for and localize the signals in the experiment to the best of my ability
- Furnish accurate ratings of my impressions about my ability to localize under all listening conditions
- Inform the Experimenter if a protector, or any other aspects of the test condition becomes uncomfortable
- Schedule multiple sessions within the allotted time and adhere to scheduled appointments with the experimenter
- Inform the Experimenter if you are unable to make your scheduled time
- Inform the Experimenter if you become tired, thirsty or wish to rest
- Inform the Experimenter if you wish to withdraw from the study

Could being in this research hurt me?

- This experiment involves localizing blank simulated gunshot sounds. No live ammunition or weapons will ever be present or used. This scenario could pose an emotional risk to the participant. If think this sound could be distressing you are asked decline participation.
- You will be exposed to the impulse (pop) sounds from blanks that will be moderately loud, but not of a level that is hazardous to hearing, even in conditions where hearing protection is not applied (i.e., the open ear condition).
- You will be wearing hearing protection as part of this study. They are designed to be snug-fitting and may occasionally emit a whistle or squeal if the microphones are covered. These are not hazardous conditions, but any reports of discomfort will be met with an offer of a rest breaks, re-inspection, and re-fitting of the devices.
- You may become thirsty during this study, especially in the field. Please let the experimenter know and water and rest breaks will be offered.
- Given the unknown availability obstetric emergency care in the field location located 45 minutes from Virginia Tech, this research has an unknown risk for pregnant females. Due to this distance, and because taking part in this research may harm a pregnancy in unknown ways, pregnant females cannot participate in this research. Please notify the investigator if you think you may be pregnant.

Will it cost me money to take part in this research?

No, it will not cost you money. Should you choose to drive yourself to the field location in Pulaski County, you will be reimbursed for mileage at the rate Virginia Tech currently uses. Additionally, you will be paid for your travel time.

Will being in this research benefit me?

We cannot promise any benefits to you or others from your taking part in this research. However, possible benefits to you include information on the ability to learn to localize sounds while wearing different hearing protectors.

What other choices do I have besides taking part in this research?

This research is not designed to diagnose, treat or prevent any disease. Your alternative is to not take part in the research.

What happens to the information collected for this research?

Your private information will be shared with individuals and organizations that conduct or watch over this research, including:

- The investigators listed on this study
- The Institutional Review Board (IRB) that reviewed this research
- Representatives of the Department of Defense
- Your personal information will not be shared with the research sponsor, but the data you produce will be shared

We may publish the results of this research. However, we will keep your name and other identifying information confidential and you cannot be identified in any manner in these publications.

We protect your information from disclosure to others to the extent required by law. We cannot promise complete secrecy.

Data collected in this research will be deidentified and used for future research or distributed to another investigator for future research without your consent.

Who can answer my questions about this research?

If you have questions, concerns, or complaints, or think this research has hurt you or made you sick, talk to the research team at the phone number listed above on the first page.

This research is being overseen by an Institutional Review Board ("IRB"). An IRB is a group of people who perform independent review of research studies. You may talk to them at (800) 562-4789, help@wirb.com if:

- You have questions, concerns, or complaints that are not being answered by the research team.
- You are not getting answers from the research team.
- You cannot reach the research team.
- You want to talk to someone else about the research.
- You have questions about your rights as a research subject.

What if I am injured because of taking part in this research?

If you are injured or get sick because of being in this research, you will be responsible for the medical care and costs incurred. If an emergency arises, emergency medical care (911) will be called. Your insurance may be billed for this treatment.
If you are injured as a result of this study, you do not give up your right to pursue a claim through the legal system.

Can I be removed from this research without my approval?

The person in charge of this research can remove you from this research without your approval. Possible reasons for removal include:

- You are unable to keep your scheduled appointments

We will tell you about any new information that may affect your health, welfare, or choice to stay in this research.

What happens if I agree to be in this research, but I change my mind later?

If you decide to leave this research, contact the research team so that the investigator can reimburse you for your time. No adverse consequences will exist if you withdraw.

Will I be paid for taking part in this research?

For taking part in this research, you may be paid up to a total of $ 120.00. Your compensation will be broken down as follows:

- $10.00 per hour during training/test sessions/travel time
- For any fraction of time less than 1 hour, you will be paid for the closest ½ hour rounded up in your favor
- You will be paid at the conclusion of each screening, training, and experimental session
- You will be paid a $30.00 bonus for successful completion of all experimental sessions

Statement of Consent:
Your signature documents your consent to take part in this research.

____
Signature of adult subject capable of consent Date

____
Signature of person obtaining consent Date

Appendix P. Phase III Participant Instructions

IC process:

1) Furnish participant with 2 copies of the Informed Consent (one is for participant to keep)

2) Instruct participant to review the informed consent.

Highlight the following:

Purpose

Introduction to the study:
The purpose of the study is to better understand how people determine sound location. You will be asked to locate a sound with and without hearing protection. Two types of electronic hearing protection will be used. If at any time the hearing protector becomes uncomfortable, please let the experimenter know. If the hearing protector is moved, you may hear a squeal. Please let the experimenter adjust the device to avoid this.

In the first session today, you are going to hear an unusual sound similar to a buzz coming from one of 24 speakers. The speakers are arranged in a clock face, so the one directly in front of you is 12 o'clock. Your basic task is to select on the screen via touchscreen to indicate which loudspeaker emitted the sound. You will undergo one familiarization task and a pretest. We are attempting to learn how individuals localize sound without any previous training.

(show graphic of display)

Everyone in the study will be asked to complete field testing. Please note that gunfire-like sounds will be used. These sounds have been tested repeatedly and pose no risk to your hearing. Blanks will be used, but no weapons will be firing these blanks. You will have an opportunity to examine the firing device should you wish to do so.

For the training condition
In between the pretest and the field test, some personnel will be asked to conduct training. The task is very similar to the pretest, only practice sessions will be incorporated to the training session. Each training session will take about an hour, spread over three sessions.

You may take a break or withdraw at any time.

Procedures section: Pending signature of this form, the next step would be to look in your ears and then conduct a hearing test. Provided the results are acceptable for continuation in this study, we will proceed with the pretest to see how well people localize sounds. Some personnel in this study will proceed to a training session comprised of 3 sessions of about an hour each. All personnel will take a second test in a field environment 45 min away from here in an open field in Pulaski County (show picture of site). The field site will take about 3-4 hours to complete including transportation. The field session will occur in the next 3-5 days.

Show them the hard copy of what a correct and incorrect answer looks like.

Please let the experimenter know if you would like to take a break, water and coffee is available in room 513. At the field site as bathroom facilities are located in a cabin about ½ mile from the testing site and we can drive you to the location.

Before they sign, ask:

____Do you understand the information provided?
____Do you feel like you are deciding without the pressure of time or other factors to make a decision?
____Do you understand that there is a voluntary choice to make?
____Are you capable of making and communicating an informed choice?

What are your questions?

Familiarization

You're going to hear sounds coming from 12, 3, 6, and 9 in order. Go ahead and begin by selecting the green button in the center of the screen. This is self-paced, so the task today will pace according to when you select the button. Please respond as accurately and quickly as possible. The first sound you will hear is coming **from 12 o'clock**
*3, 6, after 9, tell them that the last one will answer incorrectly to demonstrate what an incorrect answer will look like.

Pretest
*****_change ear condition (ITE, OTE, open) on tablet_
Now we are going to do a pre-test. You won't know where the sound is coming from and you won't receive feedback regarding your results. Just respond as accurately and quickly as possible. I emphasize that it is very important that you are accurate, but also respond as quickly as you can.

Administer the questionnaire (PASAT desktop folder, open subject's folder, then condition, press "save" in the upper left, and submit in the upper right).

LU1
Now you are going to do what we call sequential presentations. You'll hear the sound coming from 12 o'clock then it will move in clockwise order to 11 o'clock.
-Then, the sound will come from 9 o'clock and move to 10 in a counterclockwise order.
-Then the signal will start at 3 o'clock and move clockwise to 2.
-Then the signal will start at 6 and move CCW to 7.
-Once the sequential presentations have finished, the program will auto-advance and the screen will change slightly.
-You will move into the random session.

Random
For this block, you will not know where the sound is coming from like in the sequential, but you will receive feedback regarding your results. Please respond as accurately and quickly as possible.
-Once this block is over, the screen will change slightly and you will proceed to the "choose" session.

Choose
The choose session is for you to choose to re-play some of the sound locations you wish to hear again. You'll see a counter in the upper left corner that will tell you how many trials you have remaining (18 total). Your task is to touch the black circle of the speaker where you would like to hear more presentations.

Test
For this block you will not know where the sound is coming from and you will not receive feedback regarding your results

LU2-LU5 Sequential
Now you are going to do one sequential presentation. You will know where it's coming from (listed on the screen) and you will receive feedback on your results. Please respond as accurately and quickly as possible.

Random
For this block, you will not know where the sound is coming from like in the sequential, but you will receive feedback regarding your results. Please respond as accurately and quickly as possible.
-Once this block is over, the screen will change slightly and you will proceed to the "choose" session.

Choose
of the sound locations you wish to hear again. You'll see a counter in the upper left corner that will tell you how many trials you have remaining (18 total). Your task is to touch the black circle of the speaker where you would like to hear more presentations.

Test
For this block you will not know where the sound is coming from and you will not receive feedback regarding your results.

*Administer questionnaire after LU5

Field test

Familiarization
We are going to walk through four examples from 12, 3, 6, and 9 o'clock. This task is a little different from back in the lab. The experimenter will tell you when to start, you will select the "click to start" button. The sound will play, you will respond on the tablet, and then the screen

will prompt you to speak your response so we can write it down. Then the experimenter will tell you it’s okay to start the experiment.

The tablet response is the same as before. You have 24 options, even the in-between responses are considered valid. For example, when the sound is between 1 and 2 o’clock, please say that. We will tell you when to select the green button, a sound will play.

If at any time the hearing protector becomes uncomfortable, please let the experimenter know. If the hearing protector is moved, you may hear a squeal. Please let the experimenter adjust the device to avoid this.

If you need a break at any time, please let the experimenter know. Water, sunblock, and bug spray are available. Also, bathroom facilities are located about ½ mile down the hill. One of the experimenters can drive you to the facility.

Appendix Q. Figures of statistically non-significant findings included in qualitative analysis.

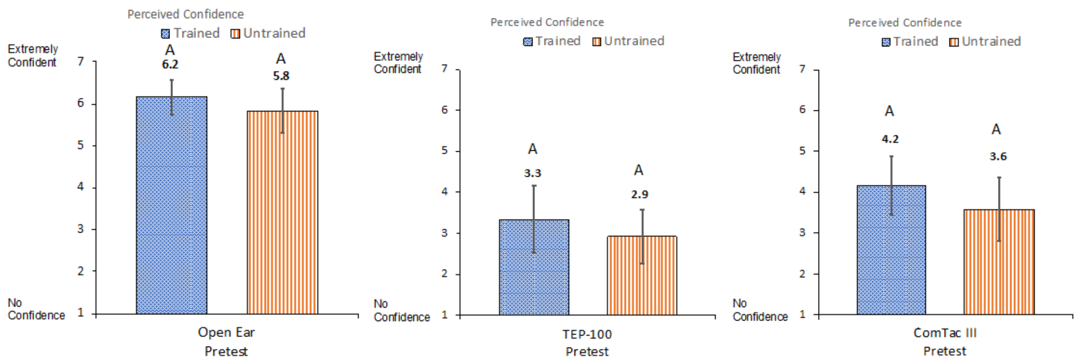


Figure 221. Mann-Whitney U results comparing mean ratings of confidence for each listening condition at pretest for trained versus untrained groups for Question 1, Perceived Confidence.

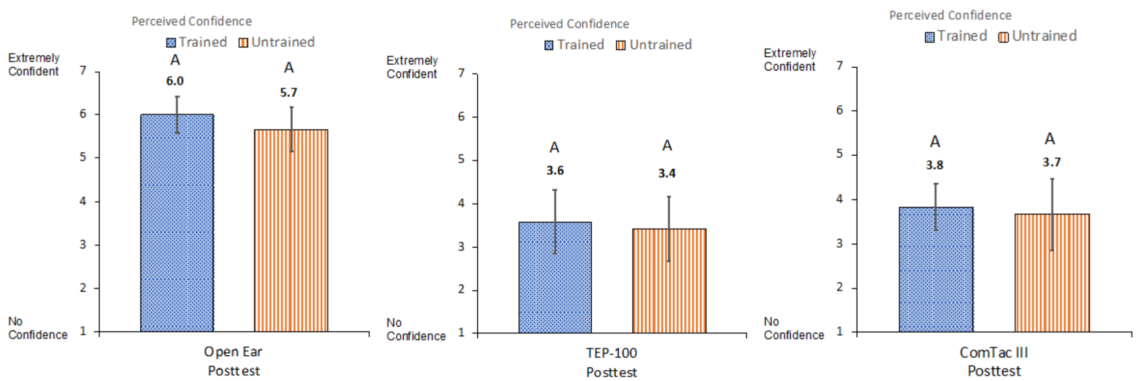


Figure 222. Mann-Whitney U results comparing mean ratings of confidence for each listening condition at posttest for trained versus untrained groups for Question 1, Perceived Confidence.

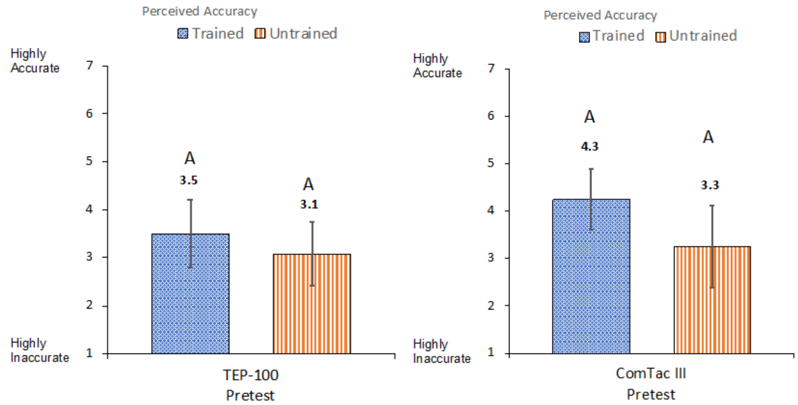


Figure 223. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at pretest in the TEP-100 and ComTac™ III conditions for Question 2, Perceived Accuracy.

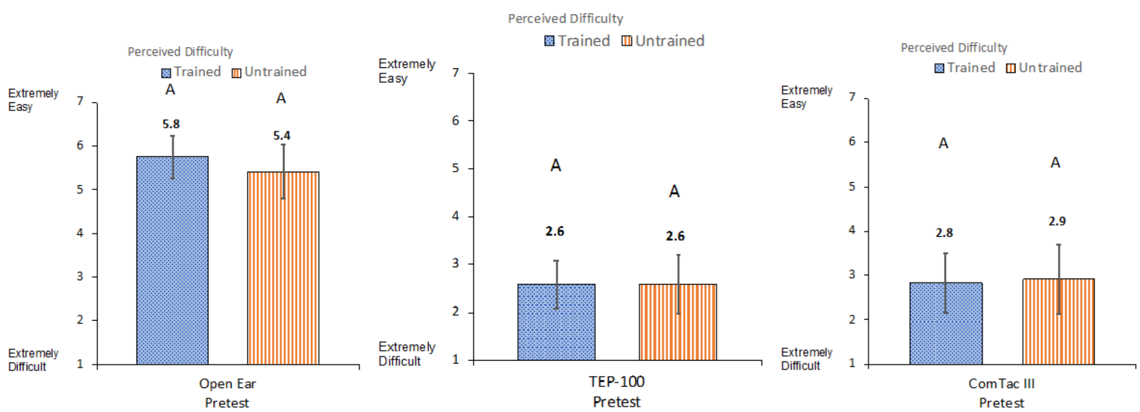


Figure 224. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at pretest for all listening condition for Question 3, Perceived Difficulty.

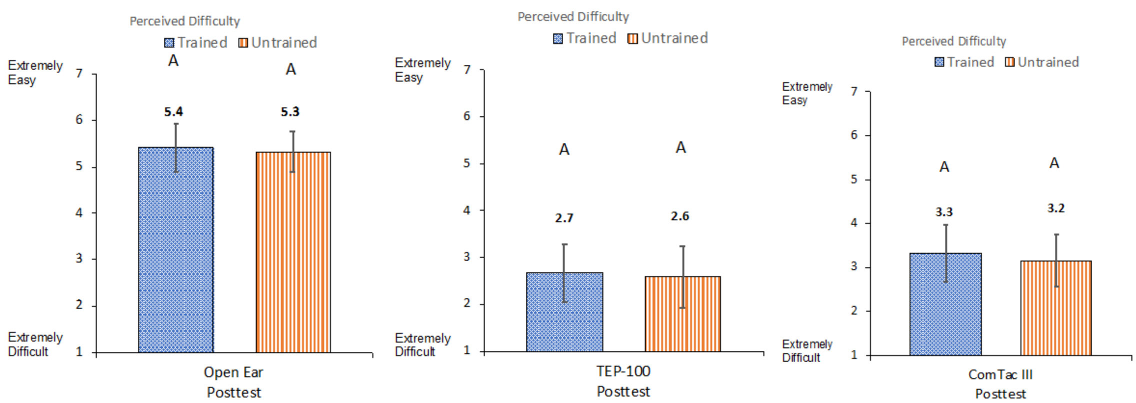


Figure 225. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at posttest for all listening condition for Question 3, Perceived Difficulty.

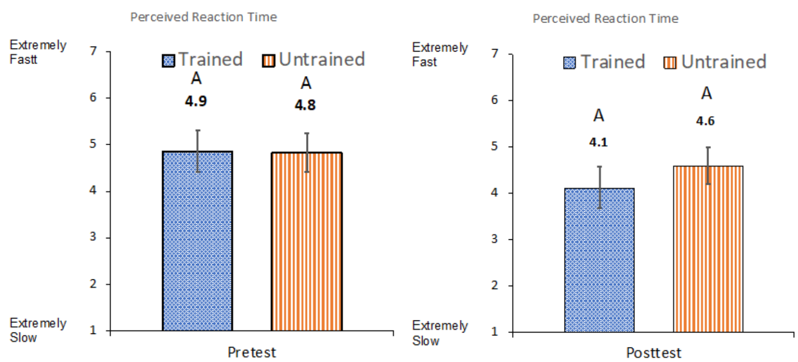


Figure 226. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at pretest and posttest collapsed across listening conditions for Question 4, Perceived Reaction Time.

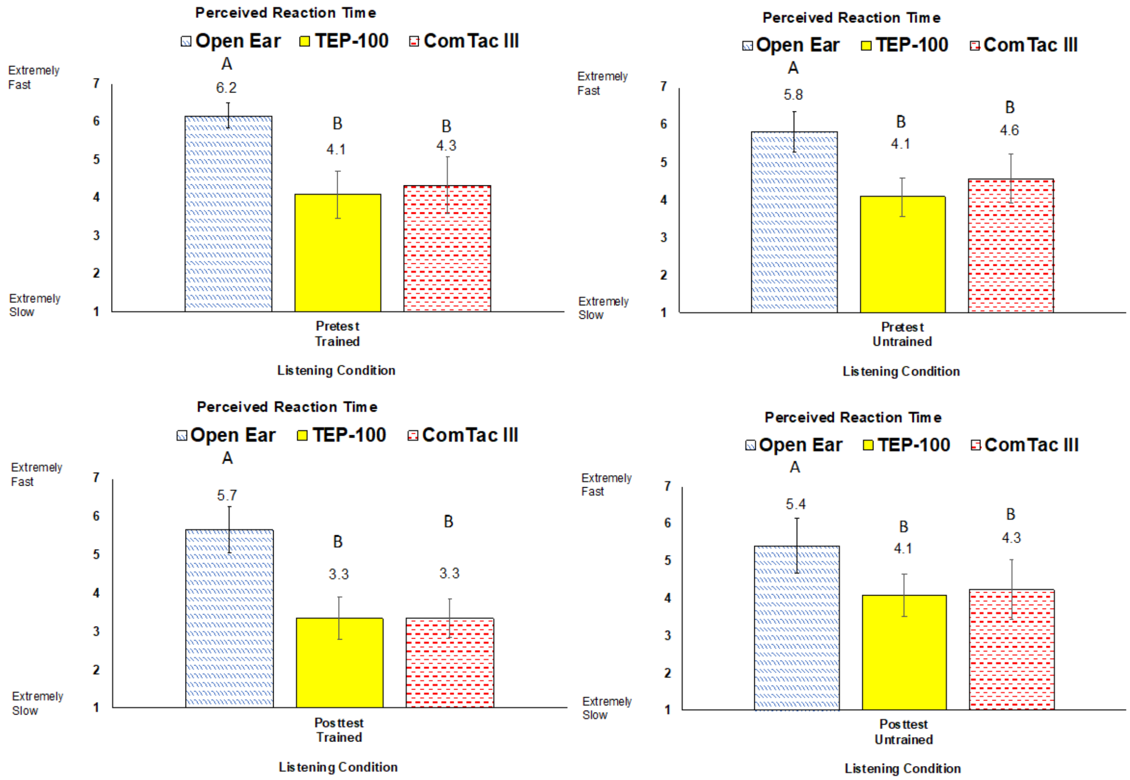


Figure 227. Mann-Whitney *U* results comparing trained and untrained ratings of perceived reaction time at pretest for all listening condition for Question 4, Perceived Reaction Time.

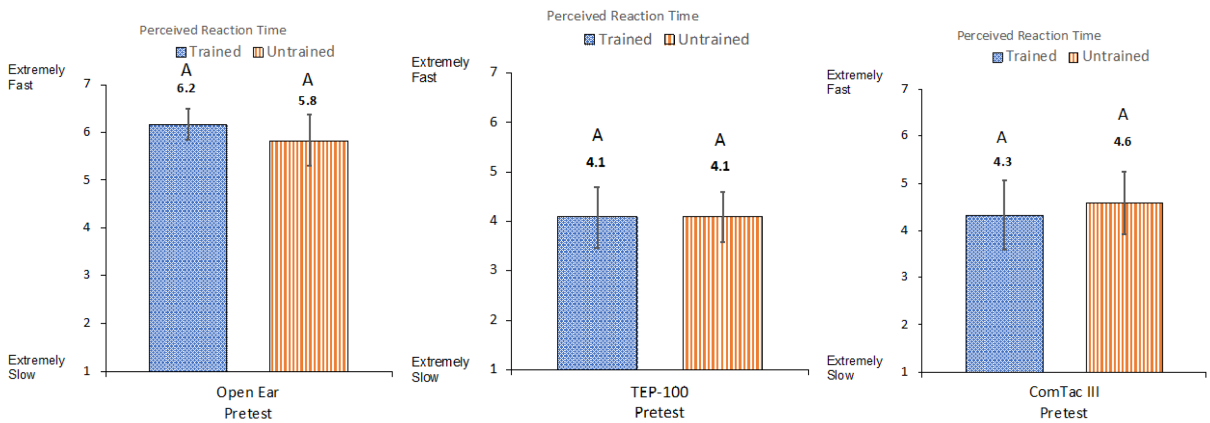


Figure 228. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at pretest for all listening condition for Question 4, Perceived Reaction Time.

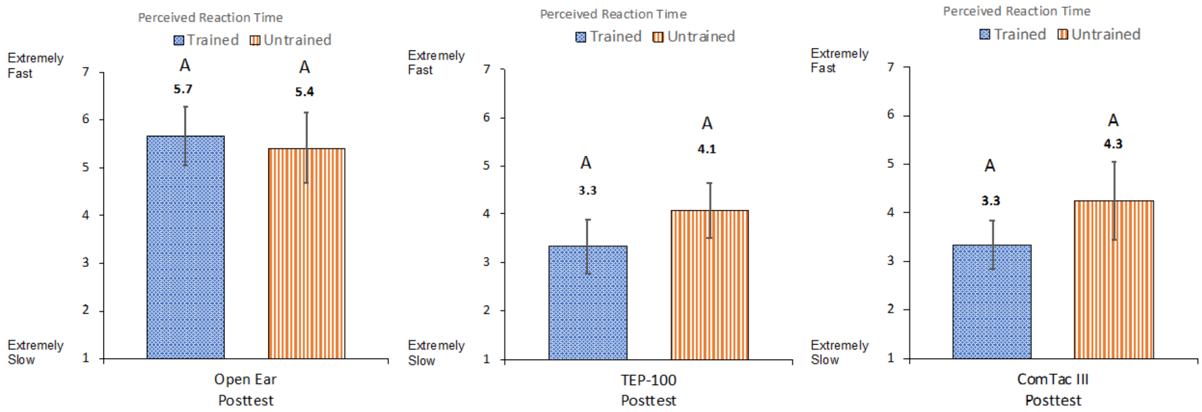


Figure 229. Mann-Whitney *U* results comparing trained and untrained ratings of perceived accuracy at posttest for all listening condition for Question 4, Perceived Reaction Time.

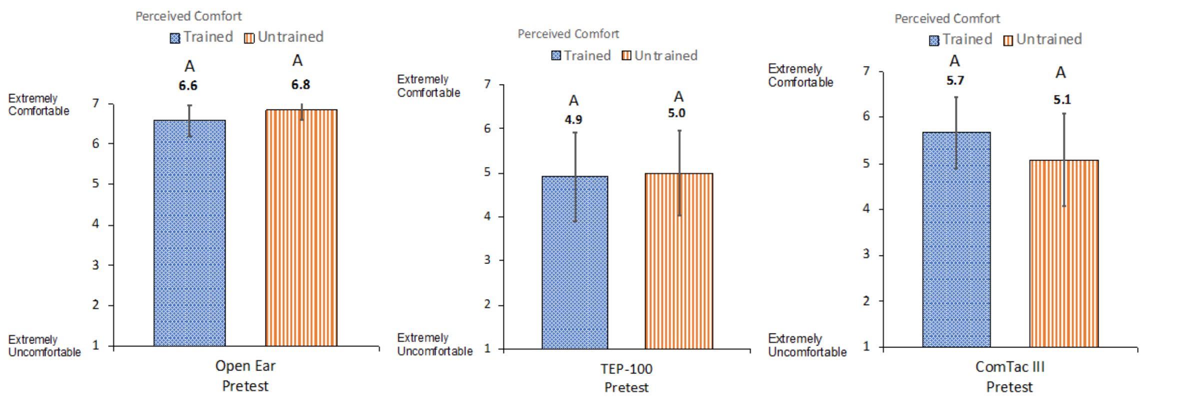


Figure 230. Mann-Whitney U results comparing mean ratings of confidence for each listening condition at pretest for trained versus untrained groups for Question 5, Perceived Comfort.

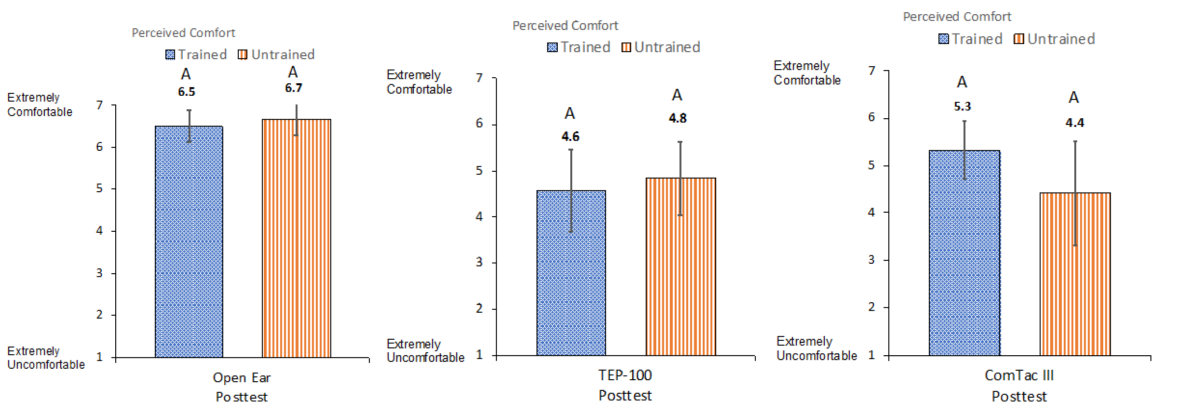


Figure 231. Mann-Whitney U results comparing mean ratings of confidence for each listening condition at posttest for trained versus untrained groups for Question 5, Perceived Comfort.

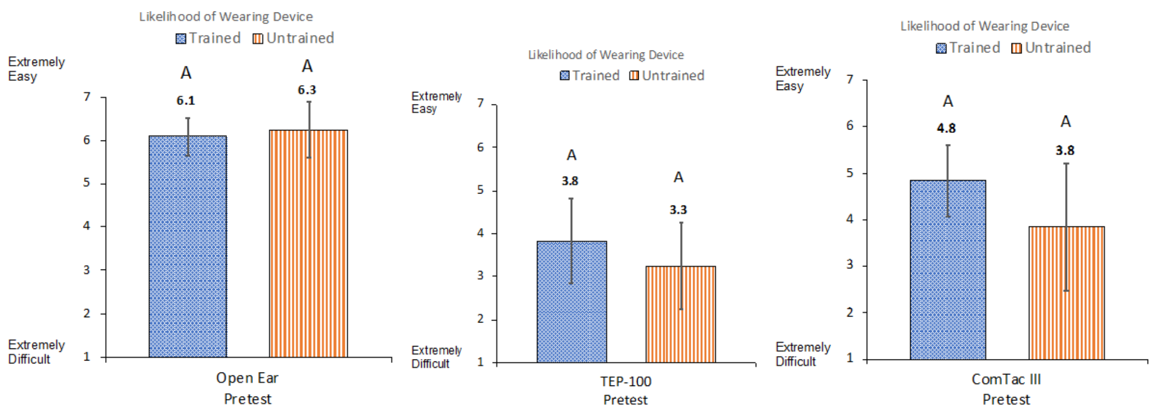


Figure 232. Mann-Whitney U results comparing mean ratings of likelihood of wearing device for each listening condition at pretest for trained versus untrained groups for Question 6, Likelihood of Wearing Device.

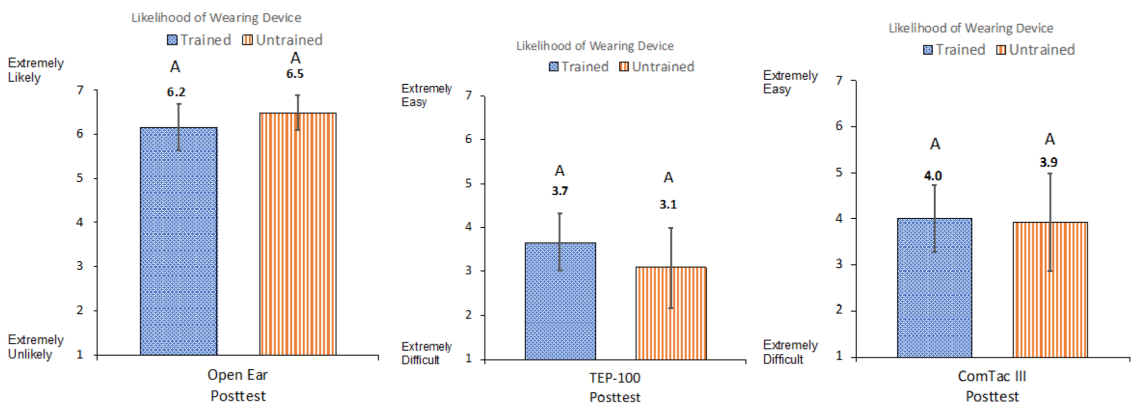


Figure 233. Mann-Whitney U results comparing mean ratings of likelihood of wearing device for each listening condition at posttest for trained versus untrained groups for Question 6, Likelihood of Wearing Device.