CONNECTED ELECTRIC VEHICLES: PERSONAL LISTENING DEVICE VEHICLE-TO-PEDESTRIAN COMMUNICATION SYSTEMS IN VIRTUAL REALITY

FINAL REPORT
SEPTEMBER 2024

AUTHORS

RAFAEL N.C. PATRICK

ESHA MAHENDRAN

MIRANDA BROWN

TANNER UPTHEGROVE

HUMAN IMPAC-T, VIRGINIA TECH

US DEPARTMENT OF TRANSPORTATION GRANT 69A3551747125

DISCLAIMER

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Connected Electric Vehicles: Personal Listening Device Vehicle-to-Pedestrian Communication Systems in Virtual Reality		5. Report Date September 2024
		6. Source Organization Code
7. Author(s) Patrick, Rafael; Mahendran, Esha; Brown, Miranda; Upthegrove, Tanner		8. Source Organization Report No. CATM-2024-R10-2024
9. Performing Organization Name and Address Center for Advanced Transportation Mobility Transportation Institute 1601 E. Market Street Greensboro, NC 27411		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. 69A3551747125
12. Sponsoring Agency Name and Address University Transportation Centers Program (RDT-30) Office of the Secretary of Transportation–Research U.S. Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590-0001		13. Type of Report and Period Covered by Final Report: April 2021 – September 2024
		14. Sponsoring Agency Code USDOT/OST-R/CATM
15. Supplementary Notes:		

16. Abstract

The last decade has witnessed a rapid evolution in the automotive industry, primarily due to the advent of electric vehicles. This shift has unintentionally reduced noise emissions, leading to increased pedestrian safety concerns. With the widespread acceptance of personal listening devices (PLDs), new methods of communicating vital traffic information to vulnerable road users (VRUs) have been introduced. The current research systematically investigated critical factors associated with VRU navigation while using PLD vehicle-to-pedestrian (V2P) systems during high-fidelity simulated crossings. Results indicated comparable reaction times through air and bone conduction modalities with differences based on bone location and frequency. However, alert signals played through open-ear PLDs resulted in quicker detection. In situations with limited visual cues, verbal guidance yielded faster crossing performance. Findings also suggest that closed-ear PLDs could isolate users from critical environmental cues, leading to increased uncertainty and hesitation during crossings. Consequently, the use of open-ear PLDs should be considered, as they offer the benefits of intelligible perception of navigational aids and environmental awareness. In sum, based on the findings, an overarching recommendation would be to create components of PLD V2P systems that consider user preferences, the content of information, and the timing of when it is presented to VRUs.

17. Key Words Personal listening devices; vulnerable road users; vehicle-to-pedestrian	18. Distribution Statement Unrestricted; Document is available to the public through the National Technical Information Service; Springfield, VT.		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 22	22. Price …

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Drivers, environmental factors, perceptual limitations, and distractions typically dictate how vulnerable a pedestrian is while walking alongside or attempting to cross roads at signalized and unsignalized crosswalks. In response, recent Virginia Law § 46.2-924 states that the driver of any vehicle on a highway shall stop when any pedestrian is crossing at any clearly marked crosswalk, while also stating that no pedestrian shall enter or cross an intersection in disregard of approaching traffic. This aligns with the National Highway Safety Administration's (NHTSA) expectation that all road users will remain aware of their surroundings, although the vehicle operator is considered the primary party responsible for traffic safety.

Despite these considerations, the NHTSA found that in 2018, 80% of pedestrians killed in traffic collisions were in urban cities while reports from the Centers for Disease Control and Prevention (CDC) state that in the United States alone, 173,887 pedestrians were involved in unintentional, non-fatal emergencies related to traffic collisions with pedestrians in 2018. Considering the rapid evolution in the automotive industry due to the advent of electric vehicles and their unintended reduction in noise outputs, pedestrian safety concerns are expected to rise. As EVs become commonplace, decreased safety for vulnerable road users (VRUs) is to be expected, therefore, it is easy to see why vehicle-to-pedestrian (V2P) alert systems are imperative in areas where walking is a primary form of transportation.

Considering the cumulative acceptance of personal listening devices (PLDs), new ways to communicate with VRUs have been introduced. Intending to emit non-intrusive alerts directly to head-worn audio interfaces, PLD V2P systems offer novel solutions to pedestrian safety concerns by providing traffic navigational aids based on real-time information. The current project sought to systematically investigate critical factors associated with how VRUs navigate crosswalk situations using various PLD V2P interface modalities and alert signal configurations using a safe high-fidelity pedestrian crosswalk environment.

The main findings were that in a controlled setting with a direct connection, listeners respond comparably, in terms of reaction times, through air and bone conduction modalities with some differences based on bone location and frequency. In situations with limited visual cues such as nighttime road usage, verbal guidance yields faster crossing performance; however, the use of an alert signal played through open-ear PLDs yields faster detection. This indicates the value of using PLDs that afford awareness of the surrounding environment either through passthrough or unoccluded means. Findings show that the use of closed-ear PLDs could isolate users from critical cues in the environment, especially in low-vision situations resulting in increased uncertainty and crossing hesitation which could lead to hazardous crossing predicaments. Consequently, the use of open-ear PLD capabilities offers benefits of intelligible perception of navigational aids and environmental awareness.

Electric vehicles (EVs) are being continuously developed to detect VRUs and other vehicles as obstacles to be avoided. Due to the increased demand and design of EVs and their reduced sound outputs compared to traditional engines, the safety of VRUs should be a vital consideration in the pursuit of transportation technological advancement. Therefore, it is recommended that PLD V2P communication systems be designed with VRUs in mind, accounting for individual differences and preferences, and be done in such a way that prioritizes early warning of unsafe situations allowing VRUs ample time to detect and react to hazardous crossing situations.

DESCRIPTION OF PROBLEM

The last decade has witnessed a rapid evolution in the automotive industry. With the advent of the electric motor, car manufacturers worldwide are focusing on developing proprietary electric vehicles (EVs) in response to consumer demand [1]. While electric motors prove more efficient and affordable than their internal combustion engine counterparts, concerns for pedestrian safety are linked to the reduced noise output of EVs. This concern is strongly reinforced in high-density, high-traffic urban environments with high ambient noise levels, leading to riskier crossing decisions for pedestrians [2].

The National Highway Traffic Safety Administration (NHTSA) found that in 2018, 80% of pedestrians killed in traffic collisions were in urban cities [3]. Reports from the Centers for Disease Control and Prevention (CDC) states that in the United States alone, 173,887 pedestrians were involved in unintentional, non-fatal emergencies related to traffic collisions with pedestrians in 2018 [4]. That same year, 6,704 pedestrian-vehicle crashes resulted in fatal injuries, marking an increase of over 10% from the previous year [5]. As EVs become more commonplace, decreased safety for vulnerable road users (VRUs) is to be expected. Based on this, it is easy to see why vehicle-to-pedestrian (V2P) alert systems are imperative in large cities or university campuses where walking is a primary form of transportation.

The issue of pedestrian safety regarding vehicle noise was initially addressed by the NHTSA in the Pedestrian Safety Enhancement Act (PSEA) of 2010, subsequently tasking the NHTSA to issue a performance standard that ensures all EVs "emit a sound that meets certain minimum requirements to aid visually-impaired and other pedestrians in detecting vehicle presence, direction, location, and operation" [2]. Current regulations developed by the NHTSA in the Federal Motor Vehicle Safety Standard (FMVSS) No. 141 require EVs to equip alert systems to produce sounds that are recognizable to pedestrians while moving at slow speeds utilizing loudspeaker systems integrated into the body of the EV [6].

Considering the cumulative acceptance of personal listening devices (PLDs), new ways to communicate with VRUs have been introduced. Intending to emit non-intrusive alerts directly to head-worn audio interfaces, PLD V2P systems offer novel solutions to pedestrian safety concerns by providing traffic navigational aids based on real-time information. While studies reveal the latency of signals communicated in various forms [7], there has not been a study quantifying the detection and performance capabilities while engaging in street crossings with PLD V2P systems. With that as motivation, the current research complements a previous project supported by CATM titled "Acoustic Situation Awareness and Its Effects on Pedestrian Safety within Virtual Environments." The series of studies aims to understand how college students interact with traffic at signalized and unsignalized crosswalks, and how they detect and respond to alert signals (both warnings signals and navigational aids) through naturalistic observation and empirical investigation. The goal is to provide design recommendations for developing effective PLD V2P traffic communication systems.

The PLD ASA study gathered contextually representative observations and subjective feedback of pedestrian-vehicle behaviors on a university campus [8] to inform the design of a virtual environment (VE) for evaluating pedestrian auditory situation awareness (ASA) while performing safe and controlled simulated street crossings [9]. Insights from technology development determined that it is feasible to develop a high-fidelity testbed environment – representative of a university campus – in which a pedestrian can engage in simulated street crossing scenarios, while findings from the research informed the selection of signal presentation modalities (i.e., air conduction (AC) and bone conduction (BC)) and signal types.

Traffic data obtained from the NHTSA revealed that in urban environments, 73% of traffic-pedestrian collisions occur at locations other than intersections, with alcohol consumption being present in nearly half of the traffic incidents [3]. Although the NHTSA identifies vehicle operators as the primary party responsible for traffic safety, pedestrians would benefit from an alert system that provides vehicle information when walking nearby roadways [2]. Involving the pedestrian as a necessary part of the V2P communication system provides an alternative means of protection, especially if the driver is under the influence or distracted by an in-vehicle source (i.e., in-dash computer/display, smartphone, passengers, etc.). *With this in mind, the reported research sought to investigate PLD V2P interface modality and alert signal configurations using a high-fidelity pedestrian crosswalk environment.*

RESEARCH APPROACH & METHODOLOGY

It should be noted that the implementation of a V2P system is two-fold. As such, for it to be a reliable and effective system, the transmission must be intelligible (i.e., the extent to which an audio signal can be heard and understood), as well as accurate (i.e., correct in all details) in its presentation of the alert signals to the intended receiver (e.g., pedestrian) for a timely and meaningful response. The research followed a three-phased process consisting of (**I**) an in-lab auditory reaction time study, (**II**) immersive environment development, and (**III**) simulated PLD V2P street crossings. The overall research objectives were to:

1. Compare auditory reaction times for signal presentation modalities and types.
2. Develop a mobile 1:1 multimodal immersive pedestrian crosswalk testbed.
3. Compare reaction and crossing times for various PLD V2P communication configurations during simulated street crossings.

Results from the combined studies will aid in identifying and developing a much-needed alert system that enables safe mobility for all vulnerable pedestrians by determining the ideal configuration for alerting pedestrians via PLD V2P systems so that pedestrians can initiate crossings during safe gaps in vehicle traffic at signalized and unsignalized crosswalks.

Phase I: In-Lab PLD Auditory Reaction Time

To evaluate human auditory capabilities and limitations associated with PLD conduction mode, a signal presentation modality reaction time study was conducted in a controlled laboratory setting. When considering requirements for developing an effective PLD V2P alert system, human operators are typically more variable in their performance compared to the technology; therefore, it is important to know the delay in auditory perception. As a result, this phase sought to address the following research questions:

- RQ1 – What is the difference in auditory reaction time based on conduction mode?
- RQ2 – What is the difference in auditory reaction time based on frequency?
- RQ3 – What bone conduction location yields the fastest auditory reaction time?

To compare auditory reaction times for signal presentation modalities and types (Objective 1), the study utilized a 5 x 7 within-subjects design with independent variables consisting of signal presentation modality ($AC_{closed\text{-}ear}$, $BC_{Condyle}$, $BC_{Mastoid}$, $BC_{Forehead}$, BC_{Inion}) and signal frequency (250, 500, 1k, 2k, 4k, 6k, 8k Hz), with the dependent variable being subjective reaction time. A variety of BC locations were selected for potential PLD communication mediums due to the affordance of open-ear perception via BC and the inconspicuous nature of the technology which would be appealing for non-invasive PLD V2P alert communication systems.

A total of thirty-one participants (*19 M; 12 F*) were recruited, however, three were omitted as outliers; therefore, twenty-eight (N = 28) individuals with an age range of 18 - 35 years old measured to have normal hearing qualified for the study. Upon successful completion, typically two hours, participants were compensated in the form of a $20 Amazon gift card. The study was approved by the Institutional Review Board (IRB #22-521).

The experiment began with participants reading and signing the informed consent. Thereafter, they were escorted to the experimental environment and given a brief audiometric hearing evaluation to screen for normal hearing indicated by 50% perception of a pure tone signal across the test frequencies at ≤ 20 dB HL. Once passed, participants were briefed on the experimental procedure – fully depress the response button when the signal is heard. Before beginning the experimental trials, participant signal intensities were set to 45 dB sensation level (SL) based on the results of their hearing evaluation. Note, that dB SL was used to establish the presentation signal intensity because it allows for all signals to be presented to each listener at the same perceived level (i.e., signal intensities are reduced to standard for listeners with better hearing while increased to standard for listeners with worse hearing). Lastly, signals were delivered to the ears with better hearing to measure optimal performance.

All measurements (pre-screening and experimental trials) were conducted within a sound-treated environment with background interference that did not mask the presentation of the test stimuli. Participation was strictly voluntary, and participants were encouraged to inform the experimenter if they felt fatigued or needed a break.

During experimental trials, signals were presented in a predetermined frequency order (1k, 2k, 4k, 6k, 8k, 250, 500 Hz) across modality conditions to reduce hearing fatigue, while modality presentation was presented in a pseudo-randomized order. Test signals which were pure tones at a given frequency were presented to listeners through a certified clinical audiometer (AD629 Interacoustics), fed through a BioPac system (BioPac MP160, BioPac AMI100D) to measure the difference between signal presentation and subjective response using a BioPac 8-channel marker box, then presented to the listener through either a Telephonics TDH-39 Audiometric Headphones for the AC conditions or a Radioear B71 Bone transducer at the appropriate contact location for the BC conditions. Figure 1 illustrates the apparatus schematic while Figure 2 shows how responses could be monitored by the experimenter using the computer interface. BioPac software, AcqKnowledge, was used to record the signal presentation and listener response timing while Excel and SPSS were used to analyze the resulting dataset.

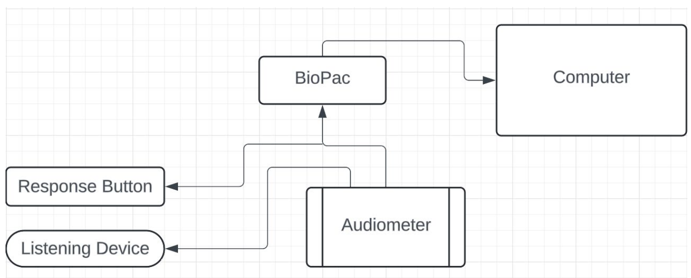

Figure 1. Signal presentation modality experimental apparatus schematic.

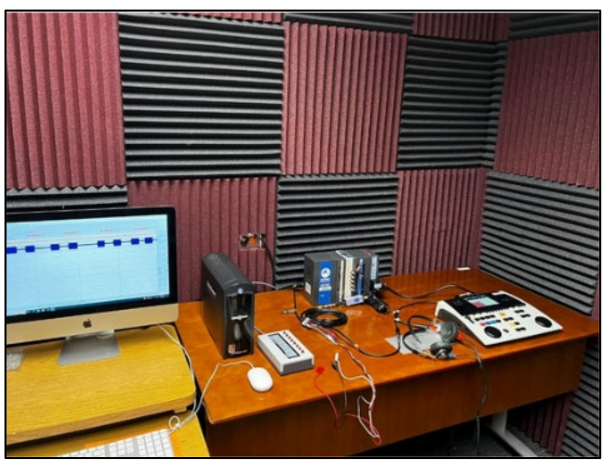

Figure 2. In-lab auditory reaction time experimental setup.

Starting with the 1k Hz condition (and continuing in the pre-determined order), listeners were fitted with the appropriate PLD and then presented with the test tone, pure tone, at random intervals for 30 seconds (~ 10 – 15 presentations each lasting ~ 2 seconds) while they responded as fast as possible, via response button, each time the signal was heard. Pauses were introduced to reduce listener expectations. This process was completed for each frequency before changing the PLD. Once all conditions were completed, participants were debriefed, compensated, and then dismissed. Data was then saved and prepared for analysis.

Phase II: Tesseract Immersive Audio Interface

To engage in safe, controlled, and realistic simulated street crossings amidst naturalistic environmental noise and vehicle traffic, a mobile 1:1 multimodal immersive pedestrian crosswalk environment was developed (Objective 2). A detailed report was published in the conference proceedings of the *30th International Congress of Sound and Vibration* titled The Tesseract: A Portable High-Density Loudspeaker Array for Spatial Sonification and Auralization [10]. The following is a summary of the development process.

State-of-the-Technology

Multi-loudspeaker spatial audio systems exist primarily in research and entertainment settings. These systems are typically standardized based on physical architecture and content production environments for fixed-media playback. However, arbitrary and nontraditional loudspeaker arrays that deliver real-time spatial audio could offer distinct advantages in several industrial domains [11], such as transportation research. Utilizing spatial audio systems to study VRU safety is an unrealized advantage of the technology. High-density loudspeaker arrays (HDLAs) allow for realistic immersion in an acoustic-rich environment essential for high-fidelity naturalistic pedestrian performance studies.

Permanent HDLAs are found in a handful of universities worldwide and in many performing arts venues. Large-scale facilities like The Curtis R. Priem Experimental Media and Performing Arts Center (EMPAC) at Rensselaer Institute of Technology demonstrate modular loudspeaker arrays of over 500 discrete speakers used for a variety of experimental acoustics studies [12]. The AlloSphere at the University of California Santa Barbara can house over 54 speakers arranged in hemispherical rings behind a large, curved display [13]. The Cube is a similar multilevel facility at Virginia Tech that houses 140 speakers in a black-box theater [14]. These academic facilities are state-of-the-art; however, it is difficult to replicate the experience they can achieve elsewhere, limiting accessibility.

Regarding research and training, Casali and Lee [15] developed an ASA test battery that inspired the development of a portable speaker array [16]. This important work demonstrated the power of incorporating a portable HDLA in a real-world situation for perceptual limitation training. While this system demonstrates the benefits of an HDLA which can be rapidly deployed in many types of non-permanent settings, it is limited in that it only routes audio (i.e., stimuli) to one loudspeaker at a time for specific research use cases (i.e., localization training). Therefore, more research was needed to establish multi-stream audio output. This would allow for the eventual study of their effects on perception and immersive realism. Moving in this direction, Upthegrove and Roan developed several temporary immersive audio systems for presenting spatial audio in small and medium-sized venues. A3, one of the largest, is a system co-designed with and implemented by Meyer Sound at the Durham Armory for Moogfest 2018 [17]. Despite technological advancement, psychoacoustic capabilities and limitations remain unknown.

Tesseract Iterative Development and Design

The desire to easily assemble and disassemble a modular speaker array is appealing for maneuverability and layout flexibility. In addition, if a system can self-calibrate, it would serve as a robust and dynamic tool that can be transported and adjusted to fit a wide range of operational environments. In early 2021, the first iteration of the Tesseract was designed and assembled by Upthegrove et. al. The Tesseract, a portable modular HDLA utilizes network protocols for both audio transmission and loudspeaker power using one connector. The Tesseract (Gen. 1) was developed as a flexible HDLA design to be used for a variety of applications. The initial development, supported by the Commonwealth Cyber Initiative (CCI), sought to explore the development and use of multichannel spatial audio displays for the sonification of cybersecurity data. Since then, the Tesseract has been used as an adaptable experimental apparatus and an artistic immersive audio interface (with the option of visual displays and tracking systems for multi-modal presentation).

The Tesseract (Gen. 1) was constructed using an off-the-shelf modular box truss, which can be mixed and matched with standard theatrical truss and pipe elements. Between truss members, aluminum pipes act as support arms for speakers that are attached by clamps. The common structural parts afford predictable designs for most locations, are easy to source, and offer structural flexibility for additional display components and measurement equipment.

MiniDSP SPK-4P and SPK-4 loudspeakers were selected due to their relatively low cost, two-channel amplifier, AVB data input, integrated Analog Devices Super Harvard Architecture Single-Chip Computer (SHARC) digital signal processors, and Power over Ethernet (PoE) power distribution. Each SPK-4P requires only one Ethernet cable for both data and power and can drive a passive companion SPK-4 (i.e., two loudspeakers driven by one Ethernet connection). An AVB-enabled Extreme Network switch routes all loudspeakers to a computer terminal (Mac Mini) which is lightweight, reasonably durable, and portable. The built-in AVB audio output from the MacOS operating system was originally selected due to it being built into Mac hardware, but channel limitations and routing challenges necessitated the addition of RME Digiface AVB for signal routing. The RME Digiface AVB can route 16 streams of two to eight channels each, connecting to a total of 32 miniDSP loudspeakers. MacOS Core Audio offers aggregation of sound cards, enabling rapid expansion when additional outputs are needed. However, only one RME Digiface can be addressed by a MacOS computer at a time.

AVB software and hardware are used to deliver audio over Ethernet to each speaker. This method allows for uncompressed 48kHz multi-channel audio to be routed with extremely low latency. Cycling '74 Max and Pure Data are the most common signal-generating software suites used with The Tesseract. Digital audio workstations are also used to present fixed-media playback. Spatial audio software from the Institute for Computer Music and Sound Technology (ICST) [18] provides ambisonic encoding and decoding. However, any other software platform capable of generating a signal that can access the sound card should be viable.

Additionally, an external input, such as a USB input device or Audinate Dante Virtual Soundcard, can provide expanded input options for software non-native to MacOS. As the physical layout of loudspeakers can vary, ambisonics format is often used to create content for the Tesseract. When installed, the physical layout of the loudspeakers is measured and provided for a third-order ambisonic decoder. The miniDSP SPK-4P loudspeakers have software to control the onboard SHARC processors. This software can be accessed on the network for one loudspeaker pair at a time when Dynamic Host Control Protocol (DHCP) is enabled, providing control for gain, equalization, delay, and other signal processing features. Note, that when DHCP is enabled, the AVB stream reliability predictably degrades and fails, so this must be disabled when producing audio.

In terms of configuration malleability and portability, the network-based approach permits variability in the number of speakers needed to achieve a decent modular output count and to afford arbitrary speaker layouts. Since being developed, the Tesseract has been assembled several times with a team of three taking anywhere from 30 minutes to 6 hours, depending on the use case. In addition, the Tesseract has also been deployed using different truss systems and dimensions to fit specific venues which serves as the motivation for the Gen. 2 iteration.

Tesseract Gen. 2 supports the same software suites as Gen. 1; however, the additional Dante Controller provides quick routing and management tools, while lightweight trusts were used for the new structure affording quicker assembly (~2 – 4 hours with two individuals). Figure 3 and Figure 4 diagram the signal flow while the physical structure can be seen in Figure 7.

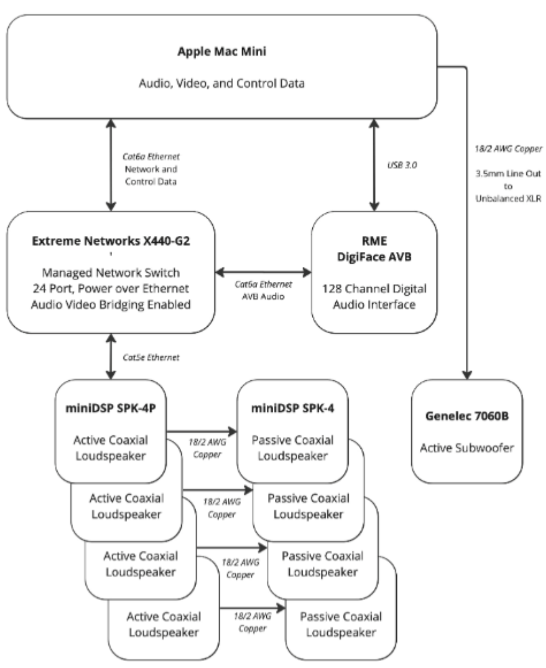

Figure 3
32-Channel Tesseract Gen. 1
Signal Flow

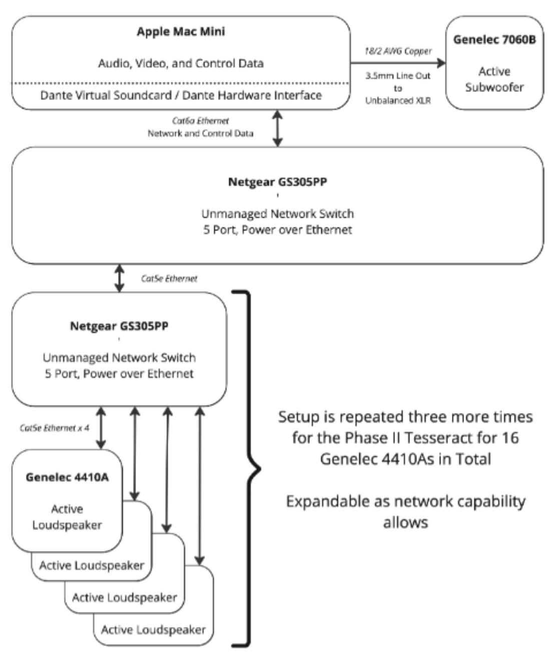

Figure 4
16-Channel Tesseract Gen. 2
Signal Flow

Several acoustic measurements were taken of the Tesseract with Gen. 1 in a 62-channel configuration and Gen. 2 in a 16-channel configuration to validate its acoustic performance. Note that no subwoofers were used nor were signal processing techniques to correct for distance or individual differences between speakers. Measurements were taken with a miniDSP UMIK-2 calibrated microphone and Room EQ Wizard (REW) in an acoustically treated environment with Gen. 1 measured in the Cube and Gen. 2 measured in the Perform Studio both within the Moss Art Center at Virginia Tech. The Tesseract was assembled in the center of the respective environment with the measurement microphone placed at the geometric center of the Tesseract based on the measured loudspeaker locations. Gen. 1 was measured to have a round-trip latency of 8.6ms from signal generation to presentation at a loudspeaker and recorded back to the sound-generating software, while Gen. 2 was measured to have an 11ms round-trip latency indicating a 2.4ms improvement in speaker performance. The frequency responses of both generations are shown in Figure 5 and Figure 6, derived from a -12dBFS and -48dBFS sine sweep impulse response measurement in REW; respectively.

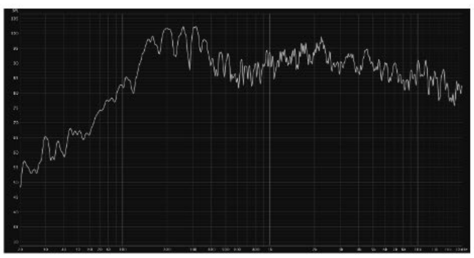

Figure 5
62 Channel Tesseract Gen. 1
Frequency Response with 1/48 Smoothing

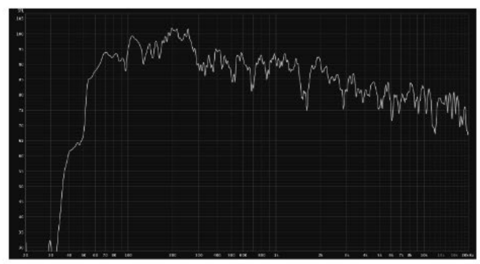

Figure 6
16 Channel Tesseract Gen. 2
Frequency Response with 1/48 Smoothing

As an experimental tool, Tesseract Gen. 2 was used as the immersive acoustic apparatus for the PLD V2P crossing reaction time study described in Phase III. To highlight its accessibility and educational potential, the Tesseract Crosswalk Module, in its 16-channel configuration, was demonstrated at the Joint CATM and CR^2C^2 Symposium in 2024 as seen in Figure 7.

Figure 7. 16-Channel Tesseract Gen. 2 Demonstration

Phase III: PLD V2P Simulated Street Crossings in Virtual Reality

With an empirically based understanding of how pedestrians respond to PLD-based alert signals, the focus of Phase III is to identify a means of creating an effective PLD V2P alert system targeted for VRUs performing unsignalized street crossings. This phase aims to evaluate the efficacy of signaling methods, signal presentation modality, signal type, and virtual environment elements on crossing performance to enhance VRU safety at unsignalized crosswalks. This phase sought to address the following research questions:

- RQ1 – What is the effect of visual information on reaction and crossing time?
- RQ2 – What is the effect of alert signal type on reaction and crossing time?
- RQ3 – What is the effect of signal presentation modality on reaction and crossing time?
- RQ4 – What is the effect of the signaling method on reaction time?

To understand pedestrians' reactions and crossing times for various PLD V2P communication configurations during simulated street crossings (Objective 3), the study utilizes a 2 x 2 x 2 x 3 "fully-factorial" within-subjects design with the independent variables consisting of the visual environment as VE (visual vs. non-visual), signaling method as SM (M1 and M2), signal type as ST (dissonance and verbal), and signal presentation modality as SP (AC earphones ($AC_{open\text{-}ear}$), AC headphones ($AC_{closed\text{-}ear}$), and BC headphones located at the condyle location ($BC_{open\text{-}ear}$)). Dependent variables in the study are participant's crossing time, trigger reaction time, and crossing reaction time, as well as their subjective opinion scores regarding condition preferences. Thirty-one (N = 31) individuals (*23 M; 8 F*) with an average age of 27 years old (S.D. 6.16) were recruited and reported normal/corrected vision and hearing. Upon successful completion, typically two hours, participants were compensated in the form of a $20 Amazon gift card. The study was approved by the Institutional Review Board (IRB #23-1188).

The experiment began with participants reading and signing the informed consent. Thereafter, they were escorted to the experimental environment and given a brief overview of the experimental procedure – cross between two cars when there is a safe crossing gap. Once prepared, they were seated and given instructions on how to wear the VR headset (Meta Quest 2), PLD audio interfaces ($AC_{open\text{-}ear}$ – Apple Air Pods Pro 1st Generation, $AC_{closed\text{-}ear}$ – Sony WH-1000XM4, and $BC_{open\text{-}ear}$ – AfterShokz OpenRun), and presented with the verbal and dissonance signals so that they could become familiar with the audio stimuli. Next, participants were familiarized with the 1:1 naturalistic crossing behavior (VE movement mimicked real-world movement) and were permitted to train until they felt comfortable.

The experimental environment consisted of a VE (technology transfer component of the ASA PLDs in the VR study), the Tesseract Gen. 2 (as described in Phase II), and the physical crossing environment (a framed crosswalk inside the Tesseract located within the Human IMPaC-T Lab at Virginia Tech). The VE was a modified version created in Unreal Engine 4.26 featuring a realistic street crossing scene with two vehicles, separated by a safe gap, traveling at a speed representative of university traffic (~25mph), realistic vehicle engine emissions, and recorded environmental background noise [9].

First, two VE modes were used throughout the study: visual and non-visual. The visual condition consisted of visual aids/cues such as clearly visible vehicles and a complete view of the road and crosswalk, while the non-visual mode represented nighttime crossings with low visibility of the vehicles and crosswalk.

Next, audio was presented to pedestrians in two forms: environmental noise and alert signals. Cycling ’74 Max, a visual programming language for music and multimedia, was used to drive audio content to both the Tesseract (see Phase II) and PLDs, as seen in Figure 8. Pedestrians were immersed in background and vehicle noise (~73 & 80 dBA SPL; respectively) which were recorded in naturalistic form using an Ambisonic microphone at a real campus crosswalk location while alert signals (ST) were presented through PLDs based on experimental trials. Note, that signal-to-noise was set so that the environmental noise did not interfere with the perception of the ST. Alert signals (verbal and dissonance) were presented through PLDs based on the signaling method (SM) using trigger boxes placed in the VE. Trigger boxes were invisible indicators that could be programmed to execute gameplay logic (i.e., send an alert signal to PLD based on SM). Note, that participants were instructed to respond to the ST while the background noise served as a source of distraction and the vehicle noise as information.

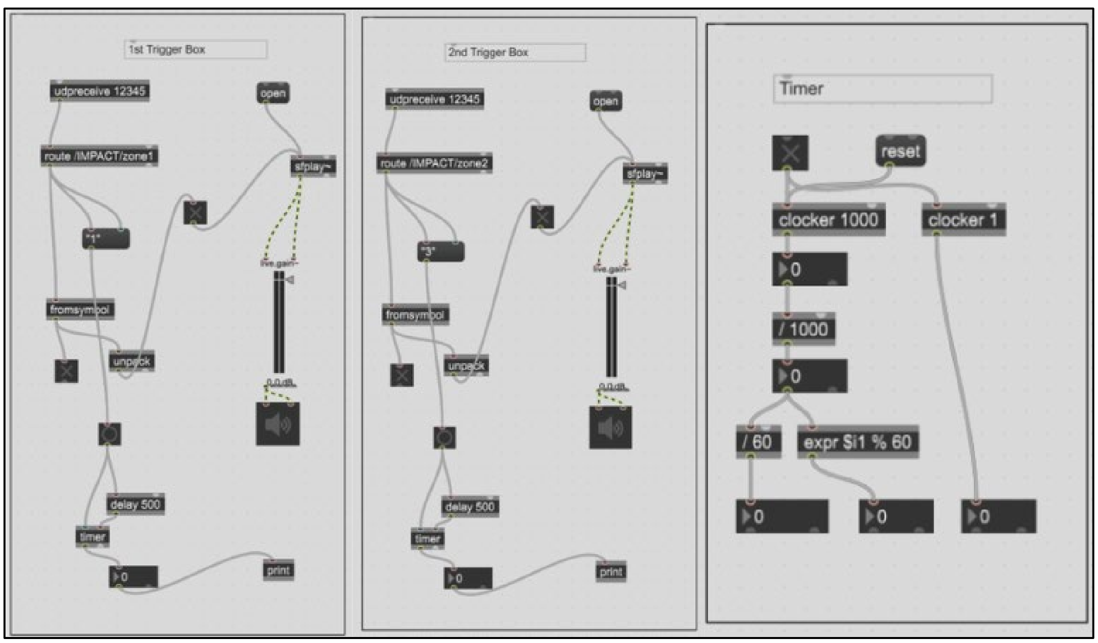

Figure 8. PLD V2P Cycling '74 Max Trigger Box Patch.

In terms of the alert signals (ST), crossing information to be acted upon by engaging the trigger and crossing was presented in two formats: a verbal command and a dissonance signal. First, the dissonance signal consisted of a combination of pure tones which was designed to capture the attention of a listener due to its unpleasant nature. Designed by Casali and Lee in 2016 for military localization training [19], the signal avoids the use of consonance by merging and shifting several pure tone signals with frequencies spanning 100 – 8000 Hz (104, 295, 450, 737, 2967, 4959, 7025 and 7880 Hz) which enables listeners to better recognize interaural time differences cues (below about 1000 Hz) and interaural level differences cues (above about 3000 Hz). The unique characteristics of the signal allow listeners to distinguish it amidst environmental interference. During trials, the dissonance signal duration was about eleven seconds. In contrast, the verbal signal presented speech content using simple assertive sentences such as “Do not cross the crosswalk” or “You can cross the crosswalk now.” The dissonance and speech signals are associated with caution and positive reinforcement.

Two signaling methods (SM) were used to inform pedestrians of vehicle information: M1 and M2. The M1 signaling method provided alert signal information to pedestrians when there was not a safe gap in the traffic suitable for crossing, while the M2 signaling method informed pedestrians that there was a safe gap in traffic encouraging crossing. M1 uses the first and second trigger boxes to initiate the verbal signal “Do not cross the crosswalk” and “You can cross the crosswalk now” and to play the dissonance signal before the lead and trailer vehicle. In contrast, M2 only makes use of the second trigger box to initiate playing the verbal signal “You can cross the crosswalk now” and the dissonance signal when there is an acceptance gap.

In sum, as seen in Figure 9, pedestrians were tasked with performing 24 trials of simulated crossings within the experimental environment, while performance metrics were automatically recorded by Cycling ’74 Max in the form of crossing time and signal reaction time. Pedestrians engaged in simulated crossing using every configuration of the PLD V2P communication system. At the end of each trial, participants were asked to respond to a short user preferences questionnaire which would help determine their preferred configuration of an ideal street crossing navigational aid using a PLD V2P alert system. Lastly, participants were encouraged to take breaks between trials to reduce fatigue and the occurrence of virtual cybersickness.

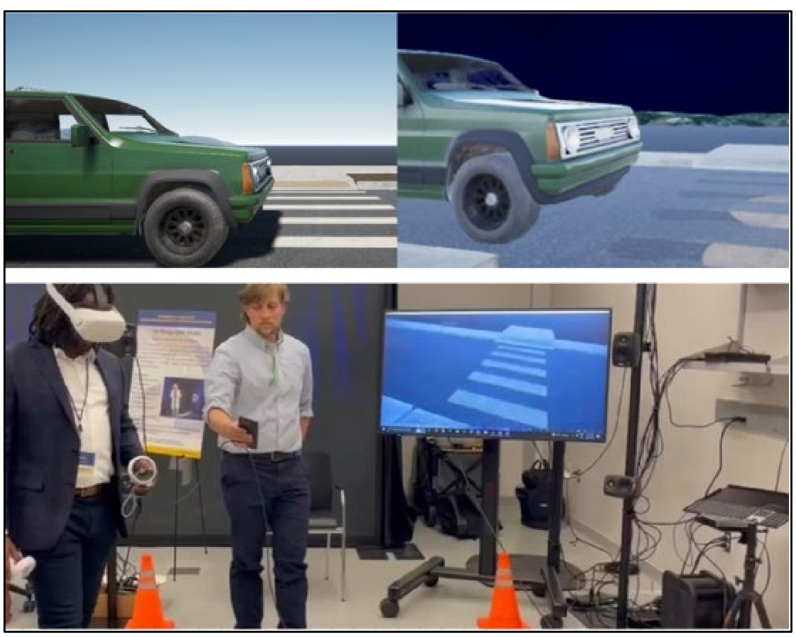

Figure 9. PLD V2P Experimental Street Crossing Environment

FINDINGS

Phase I: In-Lab PLD Auditory Reaction Time

To design effective PLD V2P alert systems, the human response latency to auditory stimuli must be known, therefore, to address this knowledge gap, PLD auditory reaction time was investigated using presentations of pure tones across a range of frequencies through air and bone conduction pathways. As seen in Figure 10, the data recorded by BioPac AcqKnowledge included a graph of the voltage readings from the audiometer and the listener response button. Once collected in its entirety, the raw data needed to be manually manipulated to extract the time between signal generation and perception which equated to auditory reaction time. To accomplish this with high accuracy, a cursor tool within the AcqKnowledge interface was used to measure the distance between two points: the start of the signal and the subjective button response of listeners. Additionally, due to human error and bias associated with coding, three independent coders reviewed the raw data and produced independently refined datasets.

Figure 10. BioPac software, AcqKnowledge, data coding.

Preliminary data analysis revealed that listeners had a consistently slower reaction time during the first response for each frequency regardless of the PLD presentation, which is typical for reaction time studies due to the acclimation period. Since frequencies were presented continuously without breaks, the first subjective response for each condition was not interpreted. Additionally, experimenter observations noted that some participants struggled to remain silent and focused throughout the trial; therefore, visible instances of distractions, speaking, or natural disturbance were recorded resulting in that data point marked for exclusion.

The average auditory reaction time across modalities and frequencies was 0.235 seconds with the fastest response (0.215 seconds) occurring with the 1k Hz signal played through the AC modality and the slowest response (0.257 seconds) occurring with the 250 Hz signal played through the AC modality. When considering reaction time regardless of frequency, the fastest response (0.228) was from the $BC_{Condyle}$ location, and the slowest response (0.242) was from the $BC_{Forehead}$. This difference of 0.014 seconds resulted in a significant difference ($p < 0.001$)

between the $BC_{Condyle}$ and $BC_{Forehead}$ locations. The $BC_{Condyle}$ was also found to significantly differ ($p = 0.004$) from the BC_{Inion} by 0.013 seconds. Lastly, 0.010 seconds between the $BC_{Mastiod}$ and $BC_{Forehead}$ yielded a significant difference ($p = 0.024$). Regarding frequency, the average response time was 0.234 seconds averages over locations with 4k Hz and 1k Hz reporting the fastest and slowest responses of 0.231 and 0.238 seconds; respectively.

Phase III: PLD V2P Simulated Street Crossings in Virtual Reality

Crossing performance based on trigger reaction time, crossing time, and crossing reaction time while receiving traffic information through each PLD V2P configuration was measured. Trigger reaction time was a continuous variable, measured in seconds, indicating the pedestrian's response to perceiving the alert signal which began once the signal was played and stopped once the pedestrian activated the trigger button. Crossing time was a continuous variable, measured in seconds, indicating the total time to traverse the crosswalk beginning once the pedestrian reacted to the alert signal by initiating their walk and ending once they completed the cross. Lastly, crossing reaction time was created as a composite variable derived from the difference between crossing time and trigger reaction time. All time variables were measured automatically via Cycling '74 Max software and displayed on the experimenter's interface. The following is a summary of the findings:

Subjective Configuration Preferences and Feedback

Subjective opinion scores of PLD V2P alert system configurations from thirty-one participants are as follows: 39% reported being familiar with the use of VR whereas 61% had no experience with VR. Regarding visual vs non-visual street crossing preferences, 74% reported that they preferred crossing in the daytime compared to 26% who preferred the nighttime crossings. Thirty-five percent (35%) of respondents preferred the $AC_{open-ear}$ (Apple Air Pods Pro 1st Generation), 32.5% preferred the $BC_{open-ear}$ (AfterShokz OpenRun), and 32.5% preferred the $AC_{closed-ear}$ (Sony WH-1000XM4) PLD. Lastly, since signaling method and signal type design recommendations are typically based on performance and developer intentions, participants were not systematically asked to rate these conditions. Although anecdotally, most participants indicated that they preferred the verbal commands compared to the dissonance signal, albeit salient, due to it being annoying and ambiguous, and they preferred when the alert signal was presented when it was not safe to cross (M1) compared to when it was safe to cross (M2).

PLD V2P Crossing Performance

Trigger reaction time, representing pedestrians' detection – initial stage in information processing – averaged 20.38 seconds. As seen in Table 1, the fastest performance (19.52 seconds) was when the dissonance signal was presented during unsafe crossing conditions through the $AC_{open-ear}$ PLD during the visual crossing environment. In contrast, the slowest performance (21.45 seconds) was when the dissonance signal was presented for safe crossing conditions through the $AC_{closed-ear}$ PLD in the non-visual crossing environment. A significant difference ($p = 0.016$) was observed between the $AC_{open-ear}$ and $AC_{closed-ear}$ conditions when presenting not-safe crossing information with a mean difference of 0.801 seconds. A mean difference of 0.939 seconds resulted in a significant difference ($p = 0.049$) in signaling methods when crossing information was presented through the $AC_{open-ear}$ PLD. During the verbal presentation of unsafe crossing information, a mean difference of 1.026 seconds was observed to be significant ($p = 0.025$) between $AC_{open-ear}$ and $AC_{closed-ear}$ PLDs. Significance ($p = 0.031$) was also observed between the presentation of safe crossing information as a verbal command rather than a dissonance signal when presenting through the $AC_{closed-ear}$ PLD based on a mean difference of 1.109 seconds. Lastly, $AC_{open-ear}$ and $AC_{closed-ear}$ presentation of unsafe crossing information in verbal form during visual crossing conditions was found to yield a significant ($p = 0.018$) mean difference of 1.603 seconds in trigger reaction time response.

Table 1. PLD V2P configuration trigger reaction time means and standard deviations.

Descriptive Statistics of Trigger Reaction Time

	N	Mean	S.D.
Visual. AC Open-ear. M1. Verbal	31	19.62	6.15
Non-Visual. AC Closed-ear. M2. Dissonance	31	21.45	5.49
Visual. BC Open-ear. M1. Verbal	31	20.31	6.43
Non-Visual. AC Open-ear. M2. Dissonance	31	20.55	5.54
Visual. AC Closed-ear. M1. Verbal	31	21.22	4.99
Non-Visual. BC Open-ear. M2. Dissonance	31	20.07	6.48
Visual. AC Open-ear. M1. Dissonance	31	19.52	6.20
Non-Visual. AC Closed-ear. M2. Verbal	31	19.97	6.03
Visual. BC Open-ear. M1. Dissonance	31	19.82	6.39
Non-Visual. AC Open-ear. M2. Verbal	31	20.78	5.99
Visual. AC Closed-ear. M1. Dissonance	31	20.34	5.97
Non-Visual. BC Open-ear. M2. Verbal	31	20.73	6.55
Visual. AC Open-ear. M2. Verbal	31	20.79	6.04
Non-Visual. AC Closed-ear. M1. Dissonance	31	20.22	5.93
Visual. BC Open-ear. M2. Verbal	31	19.96	6.06
Non-Visual. AC Open-ear. M1. Dissonance	31	19.88	5.55
Visual. AC Closed-ear. M2. Verbal	31	20.22	5.64
Non-Visual. BC Open-ear. M1. Dissonance	31	20.10	6.40
Visual. AC Open-ear. M2. Dissonance	31	20.78	6.51
Non-Visual. AC Closed-ear. M1. Verbal	31	20.58	6.54
Visual. BC Open-ear. M2. Dissonance	31	20.71	5.78
Non-Visual. AC Open-ear. M1. Verbal	31	20.13	6.44
Visual. AC Closed-ear. M2. Dissonance	31	20.95	6.62
Non-Visual. BC Open-ear. M1. Verbal	31	20.34	6.77

Crossing time, representing pedestrians' action – final stage in information processing – averaged 26.62 seconds. As seen in Table 2, the fastest performance (25.94 seconds) was when unsafe crossing information was presented as a dissonance signal through $BC_{open-ear}$ PLDs in visual crossing environments, while the slowest performance (27.53 seconds) was when unsafe crossing information was presented as a verbal command through $AC_{closed-ear}$ PLDs during visual crossing environments. When safe crossing information was presented through the $AC_{closed-ear}$ PLD, a significant ($p = 0.034$) mean difference of 1.033 seconds was observed between verbal and dissonance signal presentation. During the non-visual crossing environment, safe crossing information presented as a dissonance signal was found to yield a significant difference in crossing time (*p = 0.014*) between $AC_{closed\text{-}ear}$ and $BC_{open\text{-}ear}$ PLDs. Lastly, a significant difference (*p = 0.031*) was observed between verbal and dissonance presentation of safe crossing information through $AC_{closed\text{-}ear}$ PLDs during non-visual crossing conditions.

Table 2. PLD V2P configuration means standard deviations.

Descriptive Statistics of Crossing Time

	N	Mean	S.D.
Visual. AC Open-ear. M1. Verbal	31	26.37	5.76
Non-Visual. AC Closed-ear. M2. Dissonance	31	27.46	5.01
Visual. BC Open-ear. M1. Verbal	31	26.65	5.68
Non-Visual. AC Open-ear. M2. Dissonance	31	26.73	5.73
Visual. AC Closed-ear. M1. Verbal	31	27.53	4.40
Non-Visual. BC Open-ear. M2. Dissonance	31	26.11	5.78
Visual. AC Open-ear. M1. Dissonance	31	26.48	5.06
Non-Visual. AC Closed-ear. M2. Verbal	31	26.05	5.84
Visual. BC Open-ear. M1. Dissonance	31	25.94	5.86
Non-Visual. AC Open-ear. M2. Verbal	31	26.75	5.70
Visual. AC Closed-ear. M1. Dissonance	31	26.50	5.64
Non-Visual. BC Open-ear. M2. Verbal	31	27.07	5.64
Visual. AC Open-ear. M2. Verbal	31	27.09	5.93
Non-Visual. AC Closed-ear. M1. Dissonance	31	26.25	5.92
Visual. BC Open-ear. M2. Verbal	31	26.17	5.66
Non-Visual. AC Open-ear. M1. Dissonance	31	26.57	5.29
Visual. AC Closed-ear. M2. Verbal	31	26.54	4.87
Non-Visual. BC Open-ear. M1. Dissonance	31	26.66	5.60
Visual. AC Open-ear. M2. Dissonance	31	27.02	6.25
Non-Visual. AC Closed-ear. M1. Verbal	31	26.51	5.81
Visual. BC Open-ear. M2. Dissonance	31	26.53	5.24
Non-Visual. AC Open-ear. M1. Verbal	31	26.35	6.10
Visual. AC Closed-ear. M2. Dissonance	31	27.19	6.46
Non-Visual. BC Open-ear. M1. Verbal	31	26.45	6.28

Crossing reaction time, representing the difference between action and detection, averaged 6.27 seconds. As seen in Table 3, the fastest performance (5.94 seconds) was when the dissonance signal presented safe crossing information through the $BC_{open\text{-}ear}$ PLD within the visual crossing environment. In contrast, the slowest performance (6.84 seconds) was when the dissonance signal presented unsafe crossing information through $AC_{open\text{-}ear}$ PLDs within the visual crossing environment. A significant difference ($p = 0.028$) was observed between alert presentation to the $AC_{open\text{-}ear}$ compared to the $AC_{closed\text{-}ear}$ PLD when the dissonance signal was used to present unsafe crossing information.

Table 3. PLD V2P configuration crossing reaction times and standard deviations.

Descriptive Statistics of Crossing Reaction Time

	N	Mean	S.D.
Visual. AC Open-ear. M1. Verbal	31	6.7271	3.09
Non-Visual. AC Closed-ear. M2. Dissonance	31	6.0132	1.51
Visual. BC Open-ear. M1. Verbal	31	6.4042	2.03
Non-Visual. AC Open-ear. M2. Dissonance	31	6.2023	1.67
Visual. AC Closed-ear. M1. Verbal	31	6.3026	2.23
Non-Visual. BC Open-ear. M2. Dissonance	31	6.0552	1.96
Visual. AC Open-ear. M1. Dissonance	31	6.8432	2.78
Non-Visual. AC Closed-ear. M2. Verbal	31	6.2132	1.71
Visual. BC Open-ear. M1. Dissonance	31	6.0548	1.95
Non-Visual. AC Open-ear. M2. Verbal	31	5.9997	1.91
Visual. AC Closed-ear. M1. Dissonance	31	6.1481	1.61
Non-Visual. BC Open-ear. M2. Verbal	31	6.2981	2.95
Visual. AC Open-ear. M2. Verbal	31	6.2661	1.72
Non-Visual. AC Closed-ear. M1. Dissonance	31	6.0287	1.33
Visual. BC Open-ear. M2. Verbal	31	6.2984	1.97
Non-Visual. AC Open-ear. M1. Dissonance	31	6.7206	2.60
Visual. AC Closed-ear. M2. Verbal	31	6.2987	2.05
Non-Visual. BC Open-ear. M1. Dissonance	31	6.6477	2.63
Visual. AC Open-ear. M2. Dissonance	31	6.3639	2.18
Non-Visual. AC Closed-ear. M1. Verbal	31	5.9632	1.79
Visual. BC Open-ear. M2. Dissonance	31	5.9358	1.74
Non-Visual. AC Open-ear. M1. Verbal	31	6.2465	2.24
Visual. AC Closed-ear. M2. Dissonance	31	6.2597	2.44
Non-Visual. BC Open-ear. M1. Verbal	31	6.1065	1.75

CONCLUSIONS

The reported studies built upon previous findings to first investigate auditory reaction time, in a controlled setting, through air and bone conduction modalities and then use an immersive environment to study PLD V2P alert signal configurations during simulated street crossings.

The in-lab auditory reaction time study determined that listeners respond comparably through air and bone conduction modalities with some differences based on bone location and frequency. The condyle location yielded the fastest response compared to the forehead as the slowest. This is to be expected based on proximity to the ear, however, further post hoc investigation may reveal insights among conduction modalities based on frequency.

Regarding the PLD V2P crosswalk study, it was observed that in situations with limited visual cues, such as reduced visibility of the crosswalk or an approaching vehicle, verbal guidance yields faster crossing times. This improvement could be attributed to the clarity and reliability of verbal signals which are easier to comprehend and act upon. Conversely, non-verbal alert

signals played to open-ear PLDs led to faster reaction times and crossing times due to their ability to effectively draw attention and convey information without the need to comprehend language. This finding could suggest that non-verbal open-ear designs help preserve time and environmental awareness for effective decision-making in dynamic environments.
The use of the dissonance signal played through open-ear PLDs was observed to facilitate quicker detection as reflected through faster trigger reaction times, crossing times, and crossing reaction times when visual cues in the environment were reduced. In contrast, signals played through closed-ear PLDs yielded slower detection and action. This finding implies that closed-ear PLDs could isolate users from external cues in the environment, especially in low-vision environments, resulting in uncertainty and crossing hesitation. Consequently, the use of open-ear PLDs offers the benefit of intelligible perception of navigational alert signals and environmental awareness.

Furthermore, the optimal timing for alerting VRUs to achieve an effective reduction in crossing time and reaction time occurs when the information conveyed pertains to vehicle information rather than attempting to inform the pedestrian it is safe to cross. Therefore, the design of V2P systems should focus on the type and content of signals and how they are presented but also prioritize the timings of the alert signals to maximize VRU safety. The interplay of signal types, signal timings, preferred devices, cultural context, and the environment is what can create universally applicable and effective PLD V2P systems that enhance VRU safety at signalized and especially unsignalized crosswalks.

RECOMMENDATIONS

Based on the findings of the research, an overarching recommendation would be to create components of PLD V2P systems that consider user preferences, the content of information, and the timing of when it is presented to VRUs. Regarding the signal and its time, *designing signaling methods that play either a salient alert during unsafe crossing gaps or a verbal signal during safe crossing gaps, while prioritizing early warning of unsafe situations* to allow VRUs ample time to react and make safe crossing decisions. Considering accessibility, *non-verbal signals in PLD V2P systems can enhance accessibility for diverse populations due to their inherent ability to capture the attention of their users and deliver information without relying on language.* Therefore, future research could examine relationships between components of language comprehension as it relates to providing navigational aids for traffic avoidance during PLD V2P-assisted street crossings, and whether non-verbal signals can vary in their interpretations across different languages and cultures. Designing PLD V2P systems with open-ear and active noise-canceling features strikes the perfect balance between comfort and preserving acoustic situation awareness. This is because while closed-ear PLDs provide the most comfort to users, their ability to become aware is impaired due to reduced acoustic environmental information. Hence, open-ear PLDs must be considered especially valuable in complex urban settings where multiple sensory input integration is necessary for VRU safety.

In sum, electric vehicles are being continuously developed to detect vulnerable road users and other vehicles as obstacles to be avoided. Due to the increased demand and design of EVs and their reduced sound outputs compared to traditional vehicles, the safety of VRUs should be a vital consideration in the pursuit of transportation technological advancement.

REFERENCES

1. Hertzke, P., Müller, N., Schaufuss, P., Schenk, S., & Wu, T. (2019, August 26). Expanding Electric- Vehicle Adoption Despite Early Growing Pains. Retrieved July 26, 2020, from https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/expanding- electric-vehicle-adoption-despite-early-growing-pains.
2. National Highway Traffic Safety Administration. (2016, November). Minimum sound requirements for hybrid and electric vehicles: Final environmental assessment (Document submitted to Docket Number NHTSA-2011-0100. Report No. DOT HS 812 347). Washington, DC: Author.
3. National Center for Statistics and Analysis. (2019, October). 2018 fatal motor vehicle crashes: Overview. (Traffic Safety Facts Research Note. Report No. DOT HS 812 826). Washington, DC: National Highway Traffic Safety Administration.
4. Non-Fatal Injury Reports 2000-2018. (2020, July 01). Retrieved July 25, 2020, from https://www.cdc.gov/injury/wisqars/nonfatal.html
5. Fatal Injury Reports 2000-2018. (2020, July 01). Retrieved July 25, 2020, from https://www.cdc.gov/injury/wisqars/fatal.html
6. United States Department of Transportation, (2013), Minimum sound requirements for hybrid and electric vehicles. Federal Motor Vehicle Safety Standard, Docket No. NHTSA-2019-0085, Notice of proposed rulemaking (NPRM).
7. Smith, S. F., Rubinstein, Z. B., Dias, B., Marousek, J., Sarkhili, S., Ashqar, H., ... & Harper, C. (2019). Connecting Pedestrians with Disabilities to Adaptive Signal Control for Safe Intersection Crossing and Enhanced Mobility: Final Report [2019] (No. FHWA-JPO-19-754). United States. Department of Transportation. Intelligent Transportation Systems Joint Program Office.
8. Dam, A., Oberoi, P., Pierson, J., Jeon, M., & Patrick, R. N. (2023). Technological and social distractions at unsignalized and signalized campus crosswalks: a multi-stage naturalistic observation study. *Transportation research part F: traffic psychology and behaviour*, 97, 246-267.
9. Dam, A., Duff, C., Jeon, M., & Patrick, R. N. C. (2022). Effects of personal listening devices on pedestrians' acoustic situation awareness in a virtual reality environment. In Proceedings of the 27th International Conference on Auditory Display.
10. Patrick, R. N. C. & Letowski, T.R. (2023). Bone conduction equal-loudness: methodological issues. *Proceedings of the International Congress on Sound and Vibration.* Prague, Czech Republic.
11. Rönnberg, N. (2019). Towards interactive sonification for monitoring of dynamic processes. *In Proceedings of the Interactive Sonification Workshop, Nordic Sound and Music Computing Conference.*
12. Goebel, J. (2009). The Experimental Media and Performing Arts Center (EMPAC) at Rensselaer Polytechnic: Defining, building, and using highest-quality spaces for

hearing, seeing, and moving in space with integrated multi-modal media technology. *The Journal of the Acoustical Society of America*, 126(4_Supplement), 2155-2155.

13. Cabrera, A., Kuchera-Morin, J., & Roads, C. (2016). The evolution of spatial audio in the allosphere. *Computer Music Journal*, 40(4), 47-61.
14. Lyon, E., Caulkins, T., Blount, D., Ico Bukvic, I., Nichols, C., Roan, M., & Upthegrove, T. (2016). Genesis of the cube: The design and deployment of an hdla-based performance and research facility. *Computer Music Journal*, 40(4), 62-78.
15. Casali, J. G., & Lee, K. (2016). An objective, efficient auditory situation awareness test battery for advanced hearing protectors and tactical communications and protective systems (TCAPS): DRILCOM (detection-recognition/identification-localization-communication). *Journal of the Acoustical Society of America*, 140(4_Supplement), 3200-3200.
16. Thompson, B. S., Casali, J. G., Lee, K., & Cave, K. M. (2021). Design and Development of a Portable Auditory Situation Awareness Training (PASAT) System. *In IIE Annual Conference. Proceedings* (pp. 566-571). Institute of Industrial and Systems Engineers (IISE).
17. Ellison S., Bugg R., Upthegrove T., and Roan M. J., "Moogfest overview," [Online.] available: https://meyersound.com/video/moogfest-overview/.
18. Schacher, J. C. (2010, May). Seven years of ICST Ambisonics tools for maxmsp–a brief report. *In Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics* (Vol. 1).
19. Casali, J. G., & Lee, K. (2016). 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*.

APPENDIX

Publications, presentations, posters resulting from this project:

1. Patrick, R., Upthegrove, T., Mahendran, E., & Stankovic, N. (2023, November 14). Auditory Situational Awareness for Vehicle-Pedestrian Communication Systems: Tesseract Crosswalk Module. Transportation Webinar for CATM. https://www.ncat.edu/cobe/transportation-institute/catm/pedestrian-auditory-situational-awareness-tesseract-crosswalk-module.php
2. Mahendran, E., & Patrick, R. (2024, April 17). Immersive environments for vulnerable road user safety using personal listening devices at unsignalized crosswalks. CATM | VRU-Themed Talks.
3. Stankovic, N. Patrick, R., & Upthegrove, T. (2024, April 17). The Tesseract: An immersive audio crosswalk testbed. Oral Presentation at CATM and CR2C2 symposium | VRU-Themed Talks.
4. Patrick, R., Upthegrove, T., Mahendran, E., & Stankovic, N. (2024, April 18). Immersive environments for vulnerable road user safety using personal listening devices at unsignalized crosswalks. Research Demonstration at CATM and CR2C2 symposium | VRU-Themed Talks.
5. Upthegrove, T., Rafael, P., Hale, B., Duff, C., Stankovic, N., & Roan, M. (2023). Tesseract: A Portable High-Density Loudspeaker Array for Spatial Sonification and Auralization. *Proceedings of the 30th International Congress on Sound and Vibration*. Amsterdam, Netherlands.