ARL-TN-1019 • May 2020

Human-Autonomy Teaming Essential Research Program Project 2: Transparent Multimodal Crew Interface Designs
Technical Note 3: Multimodal Cueing for Transparency in Mobility Operations

by David Chhan, Angelique Scharine, and Brandon S Perelman

NOTICES

Disclaimers

ARL-TN-1019 ● MAY 2020

Human-Autonomy Teaming Essential Research Program Project 2: Transparent Multimodal Crew Interface Designs
Technical Note 3: Multimodal Cueing for Transparency in Mobility Operations

David Chhan, Angelique Scharine, and Brandon S Perelman
Human Research and Engineering Directorate, CCDC Army Research Laboratory

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
May 2020	Technical Note	1 November 2018–30 September 2019

4. TITLE AND SUBTITLE	
Human-Autonomy Teaming Essential Research Program Project 2: Transparent Multimodal Crew Interface Designs Technical Note 3: Multimodal Cueing for Transparency in Mobility Operations	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
David Chhan, Angelique Scharine, and Brandon S Perelman	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
CCDC Army Research Laboratory ATTN: FCDD-RLH-FC Aberdeen Proving Ground, MD 21005	ARL-TN-1019
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

ORCID ID(s): Brandon S Perelman, 0000-0001-6558-0286; Angelique Scharine, 0000-0001-8230-1824

14. ABSTRACT

This technical note, the third in its series, describes the use of multimodal cueing concepts to improve mission execution. This work is a subcomponent of the US Army Combat Capabilities Development Command (CCDC) Army Research Laboratory's (ARL's) second project in the Human-Autonomy Teaming Essential Research Program, transparent multimodal crew interface designs, in support of the US Army modernization priority—Next Generation Combat Vehicle (NGCV). The goal of this project was to improve crew members' understanding of autonomous systems' actions, intentions, goals, and general reasoning by at least 25%. To achieve this goal, a team of CCDC Army Research Laboratory scientists and contractor engineers developed iterative improvements to the NGCV Warfighter Machine Interface aimed at addressing multiple stages of the mission, from planning to execution. Interface improvements included a multimodal (tactile and auditory) cueing system designed to improve crew members' local situation awareness while commanding simulated autonomous robotic combat vehicles, a display for evaluating tradeoffs in Multi Domain Operations, and improvements to existing route-planning functionality. These interface improvements were implemented and tested in ARL's Information for Mixed Squads Laboratory, which permits the simulation of up to platoon-level operations. This note describes the scientific background for, and ideation, implementation, and testing of a multimodal cueing system.

15. SUBJECT TERMS

Next Generation Combat Vehicle, human-robot interaction, human-autonomy teaming, autonomous systems, multimodal interfaces, tactile cueing, auditory cueing

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					David Chhan
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified	UU	49	410-278-5985

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

1. Introduction

This technical note, the third in the series, describes the integration of interface transparency concepts into the Next Generation Combat Vehicle (NGCV) Warfighter Machine Interface (WMI). The WMI is designed to be platform agnostic, allowing for the control of both manned and robotic combat vehicles (RCVs). This effort, along with others by the US Army Combat Capabilities Development Command (CCDC) Army Research Laboratory (ARL) Human-Autonomy Teaming (HAT) Essential Research Program, supports the US Army's Modernization Priority NGCV. Specifically, this technical note describes the implementation of HAT Project 2, Multimodal Transparent Crew Interface Designs, Iteration 2, aimed at improving Soldiers' understanding of unmanned vehicle actions, intentions, goals, and general reasoning during the mission execution phase of mobility operations (for more discussion on these constructs, refer to ARL-TN-1003 [Perelman et al. 2020a]). Importantly, the cueing system described herein is platform agnostic like the WMI, and is intended for use in both manned and unmanned future vehicle platforms. For a detailed description of the Information for Mixed Squads (INFORMS) laboratory and the HAT Project 2 program, please refer to the first technical note in this series (ARL-TN-1003).

2. Multimodal Displays for Alerts and Warnings

According to contemporary human factors practices, information should be mapped to the modality most suited to convey it to allow the user to maintain situation awareness (SA). In high-demand environments, information should be presented redundantly to reduce the risk of signal loss, especially for the most critical warnings (Haas and van Erp 2014). The advantage of displaying information via redundant, multimodal channels is that it increases the probability that important information will be received without distracting the recipient from other concurrent tasks (Sarter 2006, 2013). By presenting the message via two or three modalities, it reduces the chance that perceptual noise will mask the information in any one modality (Partan and Marler 2005). For example, if information is presented through any one visual, tactile, or auditory channel, the probability that the message will be masked by a distractor or noise is high; armored vehicle interiors are noisy, bumpy environments where it is necessary to attend to multiple visual displays. However, the probability that it will be masked in all three modes is considerably lower (Haas and van Erp 2014). Use of multimodal cues support synergy, redundancy, disambiguation, and increase information transfer (Sarter 2013).

Each modality has advantages and disadvantages with respect to the kind of information it displays. Visual displays offer permanence, allowing for the presentation of complex information that can be retrieved as needed, while tactile and auditory displays offer immediacy, in that their cues are perceptible regardless of the orientation of the user (Haas and van Erp 2014). In addition, visual map displays offer a holistic perception of a large area, which is not possible with auditory or tactile cues. In contrast, auditory, speech, and tactile cues can offer immediate direction and distance information that bypasses the need to interpret map-based information. It has been shown that a combination of tactile direction cues with speech-based distance cues resulted in faster and more accurate responses (Hartnett et al. 2018). Auditory and tactile cues are best for capturing attentions providing short simple messages that are easily understood and remembered. Torso-based tactile cues are ideal for direction cueing, as they are omnidirectional, and associated with faster reaction times compared to auditory cues (van Erp 2007), while eliminating any front-back confusion. The torso provides a platform that is both stationary and intuitive. Used in combination with visual displays, auditory alerts can be used to draw attention to visual displays (Sarter 2013; Haas and van Erp 2014). Meta-analytic studies that compared audio, visual, and tactile displays have shown advantage to multimodal displays, particularly under conditions of high workload, and particularly when the tactile or audio cues are used to improve attention management (Prewett et al. 2012). Meta-analytic analyses also showed that tactile direction cues were particularly helpful for direction, orientation, and as augmentation to visual cues, particularly under high workload. (Elliott et al. 2009)

Speech, as a form of auditory cue, can convey more complex information—an advantage in a highly visual environment, as information can be conveyed without requiring the user to switch focus from one display to another. However, as a form of auditory warning, speech signals should be brief and only to instruct the user of immediate actions to be taken or when the visual environment is poor (Sarter 2013). This is to avoid their being partially masked or interfering with other events and to give the user an opportunity to react quickly.

Tactile or haptic displays use the sense of touch to convey information by applying forces or vibrations through a small vibrator, like that in a cell phone, to the user's skin. The vibrations provided by the display can be used to give feedback in response to commands, to signal navigational directions, and to draw attention to signals presented in other modalities. Tactile signals have been found to be especially beneficial in that they convey information and provide feedback without interfering with other tasks (Fitch et al. 2007; Ferris and Sarter 2008; Murata et al. 2017).

The Army has long been interested in using multimodal information displays (tactile, audio, and visual) to reduce cognitive workload and provide Soldiers the adequate SA required to successfully execute their mission. Prior research provided evidence that addition of tactile cues to traditional information displays that use visual or/and auditory modalities significantly reduced reaction times and improved performance in visually demanding tasks (Forster et al. 2002; van Erp and van Veen 2004; Calhoun et al. 2005). In particular, tactile cues have been found to be effective in providing alerts for attention management in simulated driving environments and providing spatial information for navigation (van Erp 2007; Elliott et al. 2007, 2009, 2010).

Based on these extensive research studies on multimodal information displays, we hypothesized that implementation of a combined tactile and auditory cueing system in NGCV would significantly enhance vehicle crew's SA and their performance.

2.1 Auditory Signal Design Considerations

We can use auditory cues in the NGCV to convey information to users and draw their attention to important events. It is important to have a clearly defined hierarchy of cues. The most critical events should be paired with warning signals that intuitively convey their urgency. Those of lesser importance should not interfere with the user's task or hearing other signals. A number of factors have been shown to affect the perceived urgency of auditory signals. In general, the perceived urgency of an auditory signal is a function of the pitch (frequency), intensity (amplitude), repetition speed, and spectral envelope of the signal (Patterson 1982; Edworthy et al. 1991, 1995; van Erp et al. 2015). The contributions of each of these features are discussed here.

2.1.1 Spectral Content

The spectral content of a signal is perceived as pitch and timbre but more importantly, it can affect the perceived urgency of the signal as well as determine its detectability. It is important to ensure that the signal is audible to all users. Generally, the human ear is most sensitive to frequencies between about 1 and 4 kHz. However, Soldiers often suffer from a phenomenon called the shooter's notch that results in a loss of hearing sensitivity somewhere in the range between 3 and 5 kHz (Nondahl et al. 2009). Because the frequencies affected by this notch vary, it is important to create signals composed of multiple frequencies. Patterson (1982) specifies that signals should have frequencies in the 0.5–5 kHz range, a fundamental below 1000 Hz, and contain at least four or more harmonically related components in the 1–4 KHz range. Doing so reduces the probability that the signal will be masked by an environmental sound event or contain frequencies in a range where the user is less sensitive (Vause et al. 1999).

Two features of spectral content alter a signal's perceived urgency: dynamic change of the fundamental and the relative harmonicity of the tonal components of the signal. Signals with rising fundamentals convey more urgency than stationary ones (Baldwin et al. 2012; Scharine and McBeath 2019). The harmonicity of the tonal components of a signal can also signal urgency. As inharmonicity increases, the signal is perceived as more dissonant and more urgent (Edworthy et al. 1995).

2.1.2 Intensity (Loudness)

Intensity is what people perceive as the loudness of a sound. Not surprisingly, more intense sounds are perceived as more urgent (Haas and Casali 1995). However, there are several reasons to limit the use of intensity to convey urgency. First, auditory signals should be intense enough to be heard over any environmental sounds (Laroche et al. 1991) but not so much as to cause hearing damage (Patterson 1982). MIL-STD-1472G (2012) gives specific guidance for level that varies as a function of the criticality of the warning. As a rule, the level of the signal should be at least 15 dB above the ambient noise level in the critical band centered on each major component frequency of the signal and should not exceed 115 dB sound pressure level (SPL) as measured at the ear of the listener. Second, one should consider the environment and whether there is a risk of lower priority signal masking more important signals and events in the environment (Haas and van Erp 2014). Finally, regardless of urgency, one should avoid startling the user as it tends to reduce the user's ability to respond appropriately to the warning. For example, Graham (1999) asked participants to rate urgency of several warning signals (e.g., beep, car horn, skidding tires) while performing a driving simulator task. Although participants consistently rated the car horn and skidding tires sounds as more urgent, they also were more likely to react inappropriately to those warnings, resulting in a 50% increase in errors.

2.1.3 Rhythm and Repetition

A warning signal can be a single tone or multiple tones with a distinct temporal pattern. Short patterns of tones with varying pitch content and rhythm are preferable for the distinction of signals, as they are easier to learn and recall (Stephan et al. 2006). Rhythm can be used to convey urgency, with faster tempo (shorter inter stimulus interval [ISI]) of successive signal elements conveying more urgency (Edworthy et al. 1991; Haas and Casali 1995; Baldwin et al. 2012; van Erp et al. 2015). Less critical events should have brief signals. Signals that convey a warning or a caution for a critical event should persist until addressed (Patterson 1982; MIL-STD-1472G 2012).

2.1.4 Priority

The current MIL-STD-1472G(1) (2019) has identified four levels of priority for auditory signals: warning, caution, alerting, and advisory (Fig. 1). They differ in terms of the immediacy of action required as well as in whether the information is critical or advisory in nature. Each of these priority levels carries with it requirements about the intensity, persistence, and content of the signal (Fig. 2). Signals of lower priority should not mask those of higher priority and signals of higher priority should convey greater urgency. Designing signals to be easily identified is expected to reduce the cognitive load of operators and decrease errors (Haas and van Erp 2014).

Priority	Immediate Action	Awareness
Critical	Warning	Caution
Informational	Alerting	Advisory

Fig. 1 Levels of warning criticality

	Duration	Type
Warning	Persists until addressed or terminated manually	2-element: Nonspeech cue followed by spoken instructions
Caution		Single element: Nonspeech cue followed by visual instructions
Alerting		Single element: Nonspeech cue
Advisory	< 2.5 s	Single element: Nonspeech cue

Fig. 2 Duration and type of auditory cue suggested for each level of cue priority

All signals, except the advisory signal, should persist until addressed, but this persistence should not interfere with other operations. Persistence can take various forms and does not necessarily need to be continuous. For example, a brief warning that remains unaddressed might be repeated periodically, perhaps with increased loudness until it is addressed. For a less critical event, a visual notification might be shown on the display. This persistence should always be in the control of the operator to ensure that they can complete their mission. There should always be the option of manually reducing the volume or shutting off the signal (Patterson 1982). However, signals should reset automatically and presented at the full level for subsequent events.

MIL-STD-1472G(1) defines two types of auditory signals and specifies when each should be used (2019). Two-element signals consist of a brief, less than 0.5-s nonspeech cue followed by spoken (<2 s) or written instructions. Single-element signals consist of the brief nonspeech cue alone. Two-element signals are only to be used for the most critical events, like warnings, that require immediate action. Single-element signals suffice for other events, but visual instructions can be provided when helpful.

Auditory signals should be easily distinguished from each other. Cues may be distinguished by differences in intensity, pitch, temporal pattern, or harmonics. At least two of these features should be different to ensure that memory for a feature like pitch or loudness is not required to distinguish them. The set of cues should be kept small to ensure that the most important cues are easily recognized. There should be no more than four to six distinct cues for critical events requiring immediate action (MIL-STD-1472G(1) 2019).

2.2 Tactile Signal Design Considerations

Similar to auditory signals that use spectral content, intensity, and rhythm to convey auditory information, tactile signals employ vibration patterns that use a combination of stimuli parameters of various waveforms, amplitudes, durations, frequency, and interpulse intervals (Brown et al. 2005; Jones and Sarter 2008). The major difference between the two sensory modalities is that they both operate at a different frequency range (tactile signal is around 250 Hz, whereas auditory signal is 1–4 kHz). Tactile information displays typically use an array of tactors between 4 and 16 tactors to provide spatial and nonspatial information to the user by applying vibration forces on their skin. Spatial information is conveyed by activating one single tactor out of the tactor array worn on the body, usually on the torso, to indicate the direction at which the user is alerted to. This method appears to be the simplest and most intuitive way to provide orientation cues (Elliot et al.

2009). For nonspatial information, a combination of tactors could be used simultaneously or in sequence to form a tactile pattern or message (Brewster and Brown 2004).

Tactile signals can be used to convey different levels of priority the same way as the auditory signals. Tactile signal parameters such as repetition rate and interpulse interval can be modified to create various vibration patterns, thus providing users different sensations and the ability to decode and distinguish between alert, caution, and warning cues. For nonspatial tactile information displays to be effective, tactile patterns not only need to be distinctive and intuitive, but also need to be limited to a small number, preferably fewer than five, to eliminate the need for users to learn and memorize (Fitch et al. 2011). Following this finding, it is logical that in any implementation of cueing systems that use multimodal information displays, tactile signals (limited to three: alert, caution, and warning) would serve to capture user's attention followed by auditory cues/signals that would provide detailed information or explanation of what the cues mean.

2.3 MMCS Implementation

The broader goals of this project are to improve crew members' understanding of the unmanned vehicle's actions, intentions, goals, and general reasoning processes (Perelman et al. 2020a). While other work under this project has served to target crew member comprehension during mission planning (Perelman et al. 2020b), the work described herein specifically targets the execution phase of the mission. In order to improve crew comprehension during mission execution, we focused on designing and implementing a multimodal cueing system (MMCS) that includes tactile and audio displays to provide warnings, cautions, and alerts aimed at enhancing crew SA and their understanding of the unmanned vehicle's actions and intentions as well as the environment in which it is operating. The MMCS prioritized information to be presented to the crew based on the attention required and the urgency of the information.

The MMCS was designed to help NGCV crews perform under circumstances of high workload and divided attention, while simultaneously helping to resolve the spatial cognitive challenges of maintaining mobility and security over remotely operated robotic vehicles. Here we provide a general description of how the MMCS was implemented in the INFORMS laboratory with information obtained from the WMI (Fig. 3). First, information such as the vehicle's speed and coordinates as well as the threats' coordinates and type were obtained from the WMI, logged and streamed in real time or near real-time using Lab Streaming Layer (LSL) protocol. Second, the MMCS used the vehicle and environment information in the LSL to determine whether 1) the vehicle was experiencing mobility difficulty and 2) a threat is within a predetermined vicinity or close proximity. Detailed descriptions of the two cues are discussed in the following sections. Through a parsing agent, these cues were prioritized based on whether they were a warning, a caution, or an alert such that if more than one cue were to occur simultaneously, the most critical cue would be presented first.

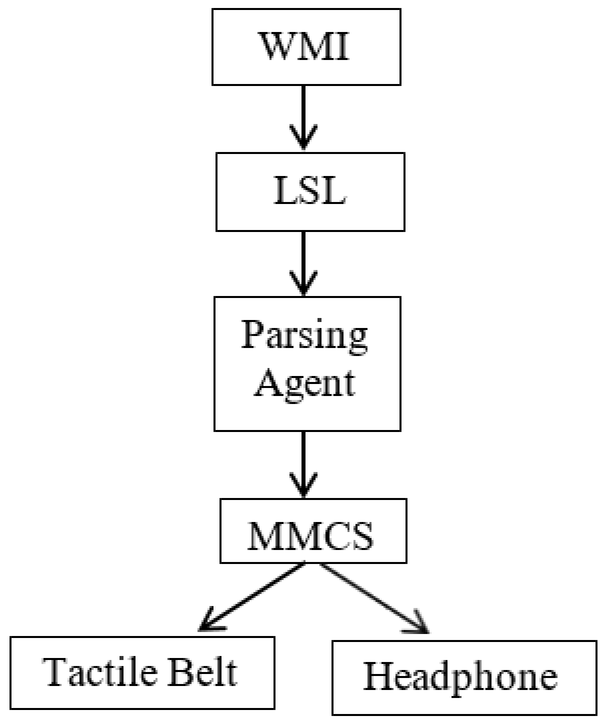


Fig. 3 Flowchart of the multimodal cueing system implementation

As a capability demonstration of the MMCS, this experiment focused on two types of cues to serve as proofs of concept. The first concept was a cue to problem conditions in vehicle mobility functionality. This mobility cue activated when the vehicle encountered mobility challenges during autonomous navigation and was intended to prompt a crew member to intervene via temporary teleoperation. This vibrotactile and auditory cue supplemented visual information that was available to the crew and was intended to improve crew members' ability to keep the vehicles moving effectively under conditions of elevated workload or divided attention. The second concept was a cue to threats in the environment. This threat warning cue activated when the WMI overhead map, referred to herein as the WMI Common Operating Picture (COP), populated environmental threats at certain distances from the vehicle under the operator's control, and provided distance, direction, and description of the threat. This cue was intended to improve crew members' ability to maintain local security of the vehicle under their control, under conditions of high workload and divided attention, as well as to reduce spatial reasoning challenges associated with the remote teleoperation of robotic vehicles.

2.3.1 Mobility Cues

The simulated vehicle in this experiment is equipped with autonomous mobility capabilities and outfitted with sensors that enable it to move independently in the environment. However, at times the vehicle may encounter perturbations in normal mobility (due to either the vehicle's mobility capabilities or environmental factors) and require assistance from the crew via teleoperation to circumvent obstacles. Should this occur, the vehicle operator will be notified via the MMCS. To detect when the vehicle is experiencing mobility difficulty requiring teleoperation, the MMCS continuously monitors the vehicle speed when it is in autonomous (waypoint) navigation mode and applies simple rules to trigger mobility cues. If the vehicle speed drops below 1 m/s for 5 s, the MMCS triggers a tactile and auditory warning, sending tactile signals to the belt and a two-element auditory warning consisting of a nonspeech auditory signal combined with a speech warning, "teleop required", played through the headphones. This multimodal warning will be repeated every 5 s until either a crew member takes over teleoperation of the vehicle or the mobility challenge resolves itself (i.e., the speed exceeds 1 m/s).

2.3.2 Threat Warning

During waypoint navigation among the rally points, the vehicle encountered a number of dynamic threats represented by Battle Space Object symbols (BSOs; Fig. 3 and Section 2.3), as depicted in MIL-STD-2525D (2014), that would update on the WMI COP; this process is identical to how other vehicles and intelligence, surveillance, and reconnaissance (ISR) assets mark targets using the WMI. In the Baseline WMI, the only information provided to the vehicle crew that a potential threat has been added to the WMI COP is the addition of a BSO. The MMCS enhances crew members' SA of the WMI COP by alerting them to the addition of new threats, as well as potentially providing the three Ds (distance/direction/description) following prioritization rules aimed at maximizing information flow while minimizing information overload. These rules are available in Table 1.

Table 1 Cue prioritization rules

Cue type	Priority level	Implementation
Mobility issue	Warning	Tactile and auditory cue repeat every 5s until addressed
Threat within 20 m	Warning	Tactile and auditory cue repeat every 5s until addressed
Threat within 40 m	Caution	One time tactile and auditory cue to avoid auditory overload
Addition of a BSO, threat greater than 40 m	Alert	One time alert chirp sound notification

Each time a BSO is added to the WMI COP, the crew member receives a simple auditory alert chirp notifying them that they COP has been updated (visually) but that no action is critical at this time. In cases where the COP update indicates a potential threat to the vehicle, the MMCS issues a caution or warning depending on the distance to the threat. In cases where more information is required, the MMCS computes the relative direction and distance between the location of the vehicle and the threat. This is more specific than how threats are called out in contemporary operations; rather than the threat's distance and direction being called out in relation to the observer or formation, the MMCS provides this information tailored to each individual vehicle to help resolve some of the difficulty in allocentric to egocentric localization associated with closed hatch operations or remote control. Once a threat has been added, this distance is checked and updated constantly in real time. The MMCS specifies three distance thresholds as parameters—one to provide alerts on COP updates, another to provide cautions, and another to provide persistent warnings as described in Table 1. In an operational system, we could expect that these distances change depending upon the nature of the threat relative to the vehicle's survivability against it. For the purpose of this concept test, the entire map was set as the threshold for pushing alert chirps to COP updates, while the caution and warning thresholds were set to 40 and 20 m, respectively. This number was selected based upon the parameters of the code that dynamically deployed the threats and, while not realistic in terms of real-world operations, would provide a useful proof of concept by virtue of allowing for more observations.

2.3.3 Tactile and Auditory Cue Implementation

2.3.3.1 Tactile Belt

For tactile cueing, we integrated a system developed by Engineering Acoustics, Inc. (EAI) that consists of a tactile belt and a control unit. The belt, which is worn about the torso, contains eight EAI-C2 tactors and is driven by the control unit (Fig. 4). This tactor is designed with a primary resonance in the 200–300 Hz range that coincides with peak sensitivity of the Pacinian corpuscle, the skin's mechanoreceptors that sense vibration (http://bdml.stanford.edu/twiki/pub/Haptics/VibrationImplementation/C2_tactor.pdf). This eight-tactor torso arrangement was initially designed to provide spatial information for navigation purposes. The eight tactors are arranged in the belt at 45° intervals, which can be associated with cardinal (N, E, S, W) and intercardinal (NE, SE, SW, NW) directions assuming that the crew member was facing N. In order to map these eight tactors to the 12 clock positions that are frequently used to indicate rough azimuth in military settings, we mapped the tactors located at the intercardinal positions to two clock positions each. For example, the NE intercardinal tactor would trigger to cue the crew member to the presence of a threat at either the 2 or 3 o'clock position. The auditory component of the MMCS provided the precise clock position.

Fig. 4 Tactile belt and C2 tactor (lower right hand panel) used in the experiment. A schematic depicting locations of the tactors on the belt is shown in the lower left-hand panel.

The tactile control unit (TCU) required to run the tactile belt is connected to the computer that is used to run the WMI via a USB interface. Modification of the TCU driver was necessary for the PC to be able to function properly. We used Zadig software to install a new driver (libusk) on the belt controller and used libusb1.0 library to write to and read from the belt controller.

2.3.3.2 Tactile Signals

Two tactile signals were designed and chosen to represent two levels of information type: warning and caution. Warning and caution tactile waveforms are shown in

Figs. 5 and 6, respectively. The short ISI and increased number of pulses simulate urgency of the information. We used a single tactor to display spatial information while multiple tactors activated simultaneously were used to represent nonspatial information. These tactile signal parameters were designed and crowd-source tested for effectiveness and intuitiveness before their implementation in the actual study (see Table 2 for details).

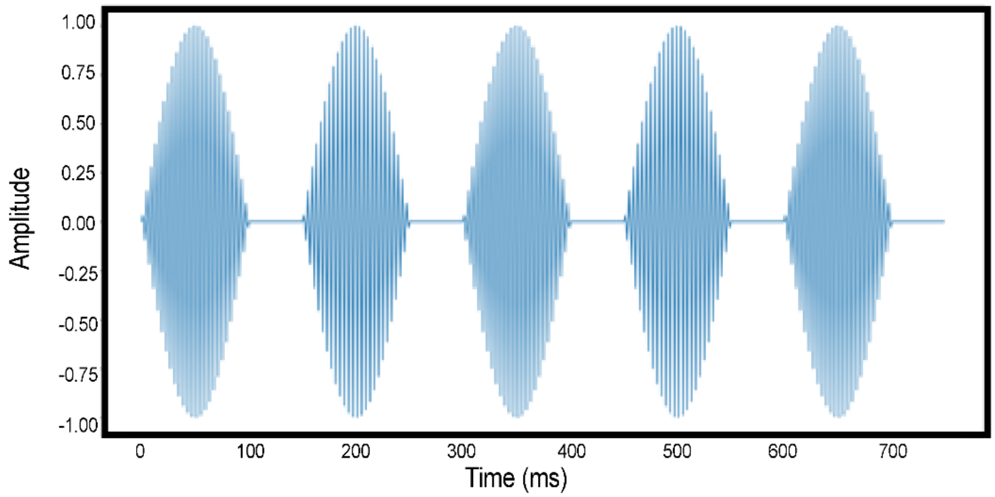


Fig. 5 Warning tactile waveform: 100-ms pulse-on-time and 50-ms ISI. Total signal duration is 750 ms.

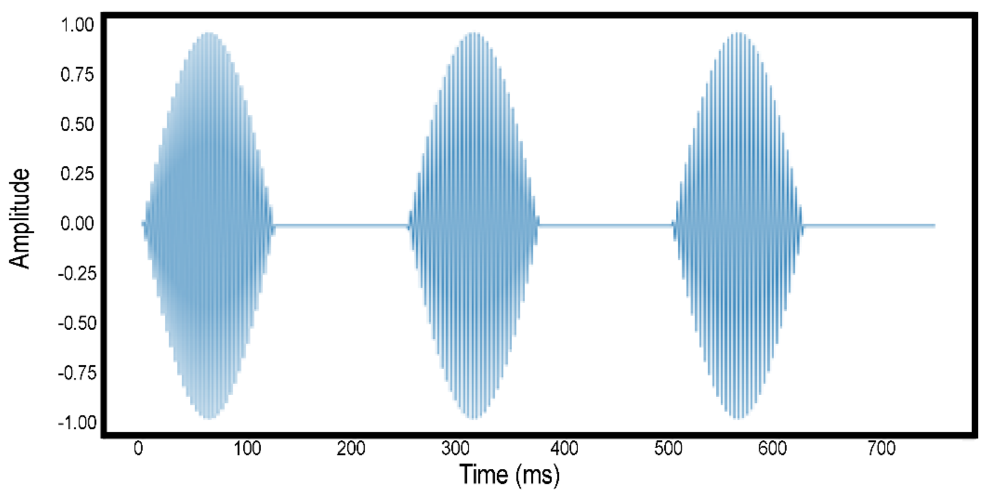
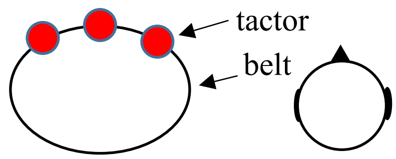
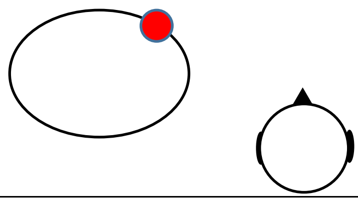
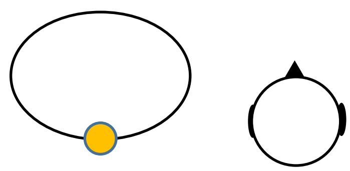


Fig. 6 Caution tactile waveform: 125-ms pulse-on-time and 125-ms ISI with the same signal duration of 750 ms

Table 2 Examples of cues and their implementations

Exemplar cues	Tactile signal	Auditory signal
Vehicle mobility	Warning tactile waveform 	Warning tone + TTS: "Warning: teleop required" Note: both tactile and auditory signals are persisted until addressed
Threat: 2 o'clock 15 m	Warning tactile waveform 	Warning tone + TTS: "Warning: threat, 2 o'clock 15 meters" Note: both tactile and auditory signals persist until addressed
Threat: 6 o'clock 30 m	Caution tactile waveform 	Caution tone + TTS: "Caution: threat, 6 o'clock 30 meters" Note: neither tactile nor auditory signal persists
A new BSO added	None	Chirp tone alert/notification

2.3.3.3 Auditory Signals

We implemented auditory cues by combining a tone followed by speech delivered through a gaming headset (shown in Fig. 7). The tones (warning, caution, and notification) were modeled from warning signals found in aircraft cockpit (see Figs. 8, 9, and 10 for their waveforms). The speech signals were generated dynamically with Microsoft text-to-speech (TTS) engine implemented with Python TTS pyttsx3 library. See Table 2 for details.

Fig. 7 HyperX CloudX Gaming Headset with integrated microphone was used to deliver auditory cues. In later studies, this function also permitted communication among crew members.

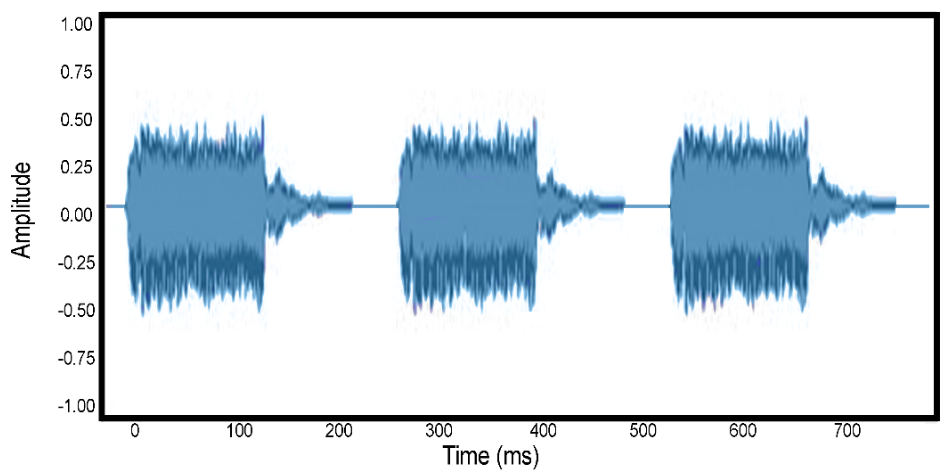


Fig. 8 Warning auditory waveform

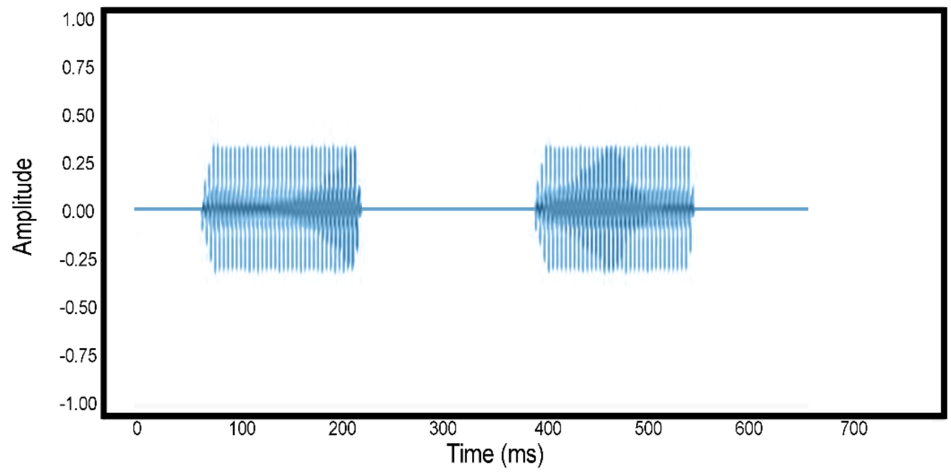


Fig. 9 Caution auditory waveform

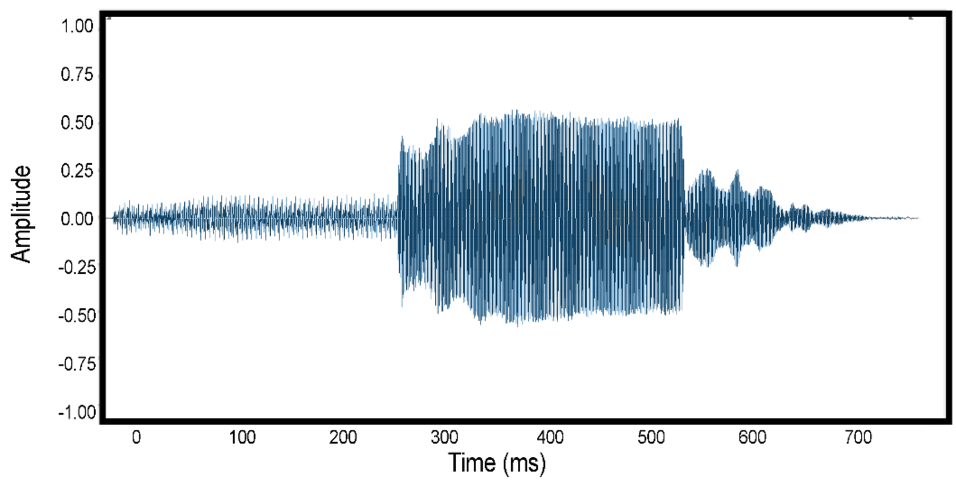


Fig. 10 Notification auditory waveform

3. Methods

3.1 Participants

Participants (*n* = 8) for the present study comprised volunteers from the CCDC Army Research Laboratory civilian workforce.

3.2 Screening for Participant Safety

In order ensure participant safety, we administered two types of tests designed to test susceptibility to motion sickness and simulator sickness experienced during the study. At the beginning of the experiment, participants completed the Motion Sickness Susceptibility Questionnaire-short (MSSQ-short; Golding 2006) to judge participants' trait propensity to develop motion sickness during real or simulated travel in moving vehicles. Based on prior studies generating normalized scores on this questionnaire, we selected the 75th percentile as a cutoff for participation, corresponding to a score of 19 or higher. If a participant's score exceeded 19, they would be dismissed from the study. During the study, we administered the Simulator Sickness Questionnaire (SSQ; Kennedy et al. 1993) before and after each experimental session to determine whether the participants were currently experiencing simulator sickness. Prior research indicates that participants can complete experiments with scores on this questionnaire exceeding 29 (Balk et al. 2013); however, for the purpose of the current study we set the cutoff at a more conservative score of 20. Any score above 20 on the SSQ would prompt the experimenter to discuss possible withdrawal from the study and implement mitigation techniques according to the INFORMS laboratory standard operating procedure, which included leading the participant to a dark room and providing them water and snacks. At no point during the study was any participant required to terminate their participation as a result of simulator sickness, or for any other reason, nor did any participants report experiencing simulator sickness symptoms.

3.3 Materials and Equipment

3.3.1 Laboratory Hardware

The experiment took place in the INFORMS laboratory using one of the crew station setups. The INFORMS laboratory crew stations each contain a steering yoke and pedals, as well as three touchscreen monitors necessary for controlling a manned or unmanned NGCV. In addition, each crew station is outfitted with headphones to allow the crew to communicate with one another and, in the current study, present auditory cues. For the purpose of this test, the crew station was also outfitted with a tactile belt and related hardware to present the tactile cues as part of the MMCS. A detailed description of the laboratory and sample image of the crew station layout can be found in the first technical note in this series (ARL-TN-1003 [Perelman et al. 2020a]). For the present study, participants executed the required tasks using a single INFORMS crew station, which consisted of three touchscreen monitors displaying the WMI, a steering yoke and pedals for manually teleoperating the RCV, a communication headset with microphone, and a tactile belt in the appropriate conditions (Section 2.3.3). Two types of WMI were employed during the course of this study. The Baseline WMI was presented as received in November 2018 (i.e., without the MMCS). The Project 2–developed Transparent WMI consisted of the Baseline WMI plus the MMCS running in parallel.

3.3.2 The Simulated Environment

The experiment employed a simulated environment developed in the Autonomous Navigation Virtual Environment Laboratory, or ANVEL, (see ARL-TN-1003 and ARL-TN-1004 for more information). The simulated environment was designed to facilitate a route reconnaissance scenario in which participants could navigate among points in the environment, while scanning for targets and avoiding threats. The environment contained a mixture of rural, suburban, and urban terrain and was intended to represent a possible future operational environment. Importantly, two versions of the environment were created. One version featured only static terrain and buildings and was used to create the overhead map available to participants via the WMI during the study. The other version featured the same terrain and buildings plus dynamic terrain objects (such as roadblocks and debris), and it also included human character models that were used for the experimental tasks; this version of the environment was visible to participants through the RCV's camera sensors. The use of these two environments allowed us to present a map to participants that could be used for planning. Participants were told that the map was based on a priori information collected by an ISR asset while not giving full information about a changing, dynamic battle space or disclosing the locations of information required to complete the experimental tasks.

3.3.2.1 Human Character Models

Human character models were placed in groups around the environment for use in experimental tasks. These character models were generally depicted as insurgent fighters, armed with either small caliber weapons (e.g., assault rifles, sniper rifles, and light machine guns) or antitank weapons (i.e., rocket-propelled grenade [RPG]-7 launchers). These character models were placed in groups around the environment to serve as targets for experimental tasks. As these character models were designated as neutral for the purpose of the scenario, they did not trigger alerts because they were not populated on the WMI COP as threats.

3.3.2.2 Simulated Dynamic Threats

To introduce elements of dynamism into the experiment, predetermined locations in the environment were designated to contain different types of threats that would spawn. These threats were selected to map onto the types of features expected during future operations and included kinetic threats (e.g., mines or improvised explosive devices [IEDs]) and electromagnetic (EM) spectrum threats (e.g., signal jammers); examples are shown in Fig. 11. The kinetic and EM threats were shown as unknown (suspected) or hostile MIL-STD-2525D (2014) BSOs, respectively, on the WMI COP. Threats were set to populate dynamically on the WMI COP when the vehicle passed to within 40 m of each threat's predetermined spawn location and thereby serve as stimuli during the experiment.

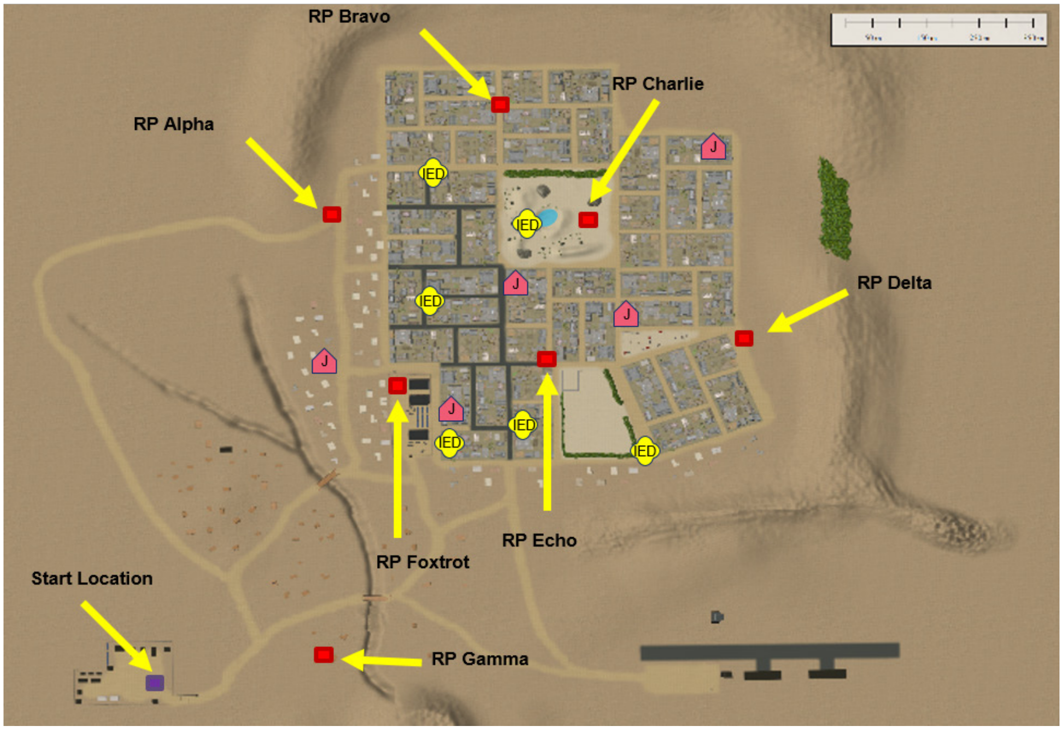


Fig. 11 Top-down view of the experiment map used as the simulated environment for the study. Yellow arrows indicate different rally points along the driving route. Potential kinetic threats are marked as "IED" and the EM threats are marked as "J" to indicate signal jammers.

3.4 Experiment Design

The experiment followed a within-subjects design in which participants completed the scenario designed to compare two levels of one independent variable: the WMI as delivered November 2018 (described here as Baseline) and the MMCS-equipped WMI (referred to herein as Transparent WMI) systems. The two WMIs were visually identical, but the Transparent WMI also presented tactile and auditory cues via the MMCS. To account for order effects, since participants would not be reasonably expected to train to asymptotic performance in the various WMI functions due to time constraints, the presentation order of each of the WMI conditions (i.e., Baseline vs. Transparent WMI) was counterbalanced. In this way, four participants completed the scenario with the baseline system first, and four participants completed the scenario with the Transparent WMI first. The dependent variables in this study were broadly aimed at operationalizing crew members' understanding of unmanned vehicles actions, intentions, goals, and general reasoning processes (see Hypotheses, Section 3.6). For the purpose of this experiment, these were defined as crew members' *inferred* SA (measured behaviorally, and recognizing caveats therein as discussed by Endsley [1990, 1995]) with respect to the RCV's current mobility status and position relative to environmental features during mission execution. We expect that the MMCS will improve crew members' ability to maintain SA over the vehicle's current mobility status and intervene when necessary, as well as to maintain local SA around the vehicle and react more quickly to dynamic threats in the environment.

3.5 Experimental Tasks

At a high level, the experimental task was a route reconnaissance task in which participants were asked to navigate a semi-autonomous RCV through a series of nine rally points, in sequence, while marking targets (Section 3.5.1), ensuring that the autonomous mobility kept the vehicle moving (Section 3.5.2), and avoiding dynamic threats (Section 3.5.3). Successfully accomplishing this task required the crew member to maintain SA of the sensor screens and the WMI COP (essentially a map plus icons indicating the locations of friendly and enemy forces). Participants were asked to use the autonomous mobility functionality of the vehicle as much as possible to mitigate their workload. The subtasks involved in this scenario are described in more detail in Sections 3.5.1–3.5.3.

3.5.1 Target Marking (Distractor) Task

The environment was populated with groups of static human character models, skinned as insurgent fighters, to represent targets. These targets could be seen through the vehicle's sensors as it moved through the environment. Participants were asked to mark these targets with a neutral BSO according to the following classification scheme: groups of one to three fighters were considered a team, while four or more fighters constituted a *squad*. Furthermore, groups of fighters should be classified as infantry, unless one or more of them was carrying an RPG launcher, in which case they could potentially pose a threat to the vehicle and should be marked as antiarmor. The marking task required participants to use neutral BSOs to differentiate them from dynamic threats on the WMI COP (see Section 3.5.3). We do not have specific hypotheses regarding this task, as it was simply used to increase participants' workload and distract them from the mobility and threat SA tasks.

3.5.2 Mobility Task

Participants were asked to use the route planning functionality of the WMI and execute as much of the vehicle's route using autonomous mobility as possible. However, the autonomy frequently had trouble successfully negotiating the narrow streets and alleyways of the environment. Participants were expected to intervene (i.e., take over control via manual teleoperation) if the vehicle appeared to be experiencing a mobility challenge. Mobility challenges were specifically operationalized as conditions where the vehicle's mobility mode was set to autonomous navigation (i.e., "plan following") mode, and it was moving below a certain speed threshold (1 m/s) for at least 5 s. The first two conditions (mobility mode and speed) were used to define *stalls* in mobility, whereas the persistence of these conditions for at least 5 s indicated a more serious challenge to mobility.

3.5.3 Threat SA Task

In addition to the target marking task, participants were tasked with avoiding areas of the environment that contained potential threats. As the vehicle traversed the environment, threats would spawn, visible as MIL-STD-2525D (2014) BSOs on the WMI COP (see Section 3.3.2.2). Participants were instructed that each threat had an area of effect radius of 40 m, and that they should avoid passing within 40 m of each threat.

3.6 Hypotheses

The MMCS-equipped Transparent WMI was tested against the baseline system according to the Decision Point 1 (DP1) goal constructs (crew's comprehension of vehicle actions, intentions, goals, and general reasoning; see ARL-TN-1003). The following hypotheses served to guide the test.

3.6.1 Mobility Task

During each scenario, participants were asked to ensure that the vehicle's autonomous mobility was functioning properly as it executed waypoint navigation. Participants' ability to detect any perturbations to normal function of the autonomous mobility, and intervene to mitigate them, maps onto their understanding of the vehicle's actions (i.e., what it is currently doing) relative to its goals (i.e., what it is trying to achieve). We operationalize these perturbations in two ways: mobility *stalls*, defined as a state in which the vehicle is moving at a speed of less than 1 m/s, and more serious mobility *challenges*, where a stall lasts longer than 5 s. We expect that the MMCS will improve participants' ability to respond to these mobility perturbations.

H1) Participants will react more quickly to stalls in mobility using the MMCS-equipped Transparent WMI than the baseline system.

H2) Participants will react more quickly to major mobility challenges using the MMCS-equipped Transparent WMI than the baseline system.

3.6.2 Threat SA Task

In addition to ensuring autonomous mobility, the participant was also tasked with maintaining SA on the WMI COP to avoid dynamic threats that populated in the environment as the vehicle moved between rally points. The ability to maintain this SA of the vehicle's position relative to threats represents comprehension of the vehicle's actions relative to threats in the environment. We expected that the auditory and tactile cues provided by the MMCS would improve the crew's ability to respond to the dynamic threats, stopping earlier to avoid entering threat zones, and initiating responses farther away from the threat.

H3) Participants will respond to dynamic threats more quickly using the MMCS-equipped Transparent WMI than the Baseline WMI.

H4) Participants will respond to dynamic threats at a greater distance using the MMCS-equipped Transparent WMI than the Baseline WMI.

3.7 Statistical Tests

Hypotheses H1–H4 were tested using standard null hypothesis significance tests, and effect size is reported as Cohen's *d*. Furthermore, to address the goals of DP1, effects are furthermore reported in terms of percent change between the Transparent and Baseline WMI systems. Statistical tests were conducted using the R Statistical Computing Language (R Core Team 2014).

3.8 Procedure

Participants arrived at the INFORMS laboratory and provided verbal consent to participate in the study. After consenting, participants were given a brief tutorial during which they learned the functions of the vehicle and WMI necessary to complete the scenario. The tutorial included targeting training on the following:

- Loading/saving, planning, editing, and executing waypoint plans using autonomous mobility
- Teleoperation
- Placing/editing BSOs to mark targets

After demonstrating the ability to execute the tasks (Section 3.5), the scenario was described to the participants in detail. Each block of the experiment (i.e., completing the scenario using the Baseline and Transparent WMIs, respectively) took roughly 1.5 h for a total experiment duration of roughly 3 h. Before the experiment, participants completed the MSSQ-short as a screening questionnaire to assess propensity for motion sickness. Before and after each experimental session, participants also completed the SSQ to ensure that they were not experiencing adverse motion- or simulator-related symptoms as a result of participation in the experiment.

4. Results and Discussion

4.1 Autonomous Mobility Response

Our first two hypotheses pertained to autonomous mobility and crew members' ability to ensure the vehicle maintained course by intervening with teleoperation or replanning when necessary. We operationalized perturbations to normal autonomous mobility in terms of both minor perturbations (stalls) and major perturbations (challenges) to characterize participants' performance using each WMI (i.e., the Baseline WMI, and the MMCS-equipped Transparent WMI) at multiple levels of severity. These stalls and challenges were generated via analysis of the vehicle's mobility data; the number of these perturbations generated by participants in each experimental condition is available in Table 3.

Table 3 Mobility perturbations by type occurring in each WMI condition

Perturbation	Baseline	Transparent
Mobility stalls	288	340
Mobility challenges	119	166

We do not offer any hypotheses regarding the frequency of mobility challenges. Judging by Table 3, participants experienced more mobility perturbations using the Transparent WMI, though the reasons for this are not immediately clear. The purpose of the system was to increase participants' awareness of these mobility perturbations and to reduce the time required for participants to respond to them. Therefore, we do not offer a priori hypotheses or potential explanations as to why participants experienced greater or fewer perturbations in each of the conditions and instead focus on distributions of the durations of these events. The statistical tests employed herein are robust to violations of homogeneity of variance, which we would expect from different numbers of observations in each condition.

To test H1, mobility stall durations were pooled for each of the participants in each WMI condition, yielding eight observations (one per participant) for each WMI. A t-test did not find a significant difference between the duration of mobility stalls experienced using the Baseline WMI (M = 7727.78 ms, SD = 9373.93 ms) and MMCS-equipped Transparent WMI (M = 6372.86 ms, SD = 7759.46 ms) despite the latter durations being 11.69% shorter (see Fig. 12). In terms of minor mobility perturbations, these results suggest that the MMCS did not substantially improve the crew's performance to ensure normal autonomous mobility relative to baseline.

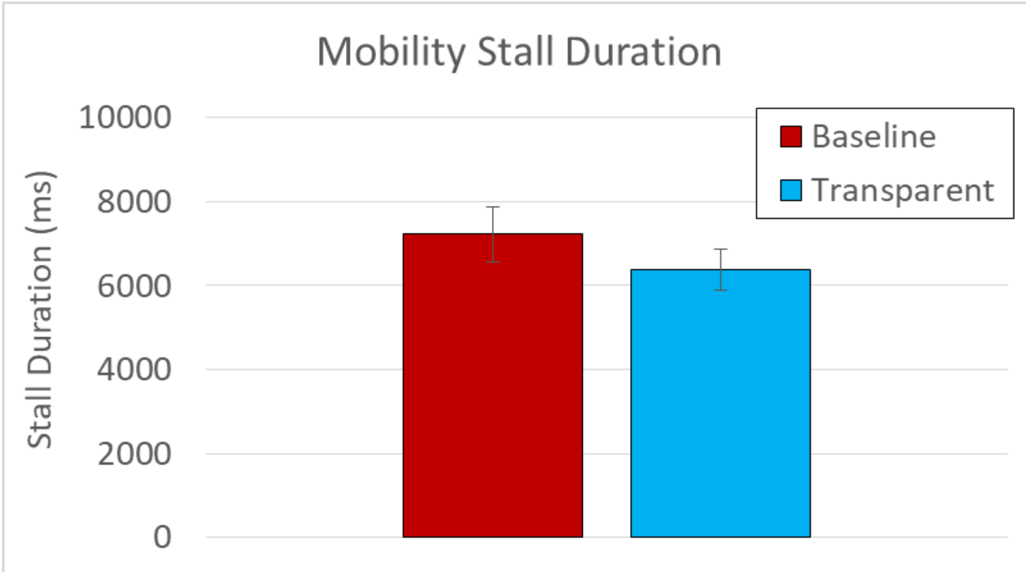


Fig. 12 Mean mobility stall duration in each experimental condition. Error bars indicate standard error.

H2 asked a similar question, but isolated more severe mobility perturbations (challenges; stalls lasting longer than 5 s). A t-test found a significant effect of WMI condition, $t(245.66) = 2.78$, $p = 0.006$, $d = 0.389$ such that participants took longer to respond to mobility challenges using the baseline system (M = 18129.93 ms, SD = 12346.77 ms) versus the Transparent WMI (M = 13725.57 ms, SD = 10279.19 ms). This equates to a roughly 4.5 s difference in average duration between the two systems and a 32.09% improvement for the MMCS over baseline (see Fig. 13).

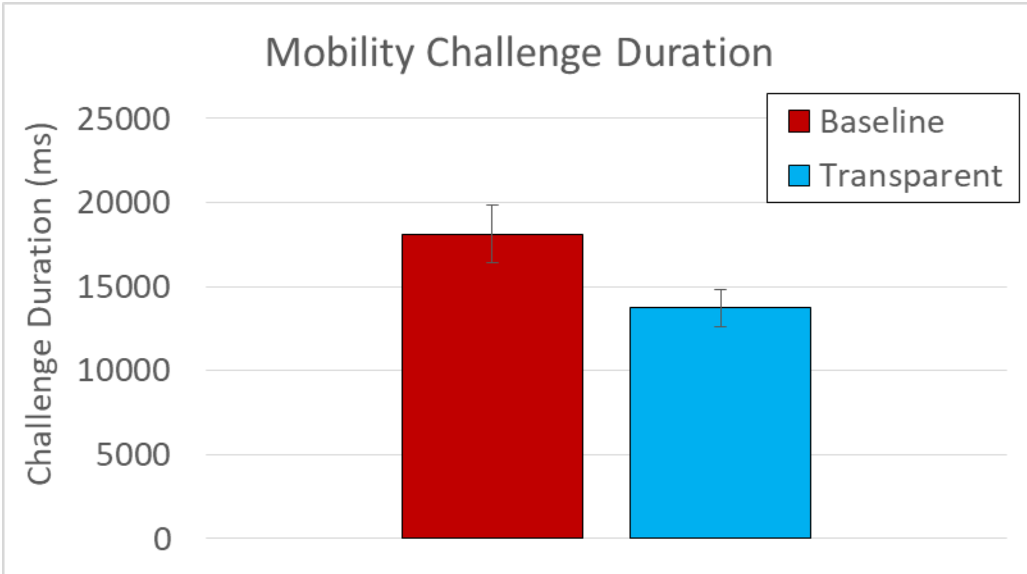


Fig. 13 Mean mobility challenge duration in each experimental condition. Error bars indicate standard error.

These results support H2 in that the MMCS-equipped Transparent WMI improves crew members' ability to respond to severe mobility challenges. Juxtaposed with the results reported previously for H1, one interpretation of these results is that the MMCS helps crew members maintain SA over the vehicle's autonomy performance and mitigate perturbations to normal autonomous mobility. This effect was more pronounced when examining the more severe perturbations in isolation. The results described in this section demonstrate that the MMCS-equipped Transparent WMI offers advantages to the crew in terms of their ability to execute mobility using a remotely operated autonomous vehicle.

4.2 Dynamic Threat Response

Our third hypothesis, H3, was that the MMCS-equipped Transparent WMI would improve participants' awareness of dynamic threats, reducing the time required to take them to take corrective action (response time). To test this, we isolated the data logs at time t, when the threat BSO deployed on the map, through time t + 60 s, giving an upper bound of 1 min to respond. This interval was extremely generous given the data and was intended to be highly inclusive, differentiating cases with responses from those in which the participant failed to respond at all. Further filtering was required to remove cases where the vehicle was not in the appropriate mobility mode to measure a response, or moved to within the radius of a threat and began another task, such as replanning their route. The environment contained 25 threats, meaning 25 possible observations per participant. Table 4 shows the threat interactions available for analysis in each condition. Since the data-logging software in the WMI does not accurately record braking and other responses that would be appropriate during teleoperation, we further down-selected observations to isolate only those occurring when the vehicle was in autonomous navigation mode (Table 4, row 3). In some cases, participants failed to respond altogether (Table 4, row 4). The remaining number of observations used for the analysis is shown in Table 4, row 5.

Table 4 Threats available for analysis in each experimental condition. Row 1 shows the total threats per condition. Row 2 contains the number of observations generated in each condition, which equates to the number of threats revealed across all trials in those conditions. Data are presented as total number, followed by the percentage of the total in parenthesis.

Data available	Baseline	Transparent
Total threats across all participants	200 (100%)	200 (100%)
Threats revealed across all participants	120 (60%)	178 (89%)
Threats revealed during autonomous navigation	102 (51%)	139 (69.5%)
No participant response to threat	39 (19.5%)	11 (5.5%)
Usable response time/distance observations	63 (31.5%)	128 (64%)

Of particular note here is that, while participants revealed (i.e., triggered by passing within the threat's radius; see second row) more threats when completing the scenario using the MMCS-equipped Transparent WMI, their response rate was substantially higher (row 4). Examining the trials during which participants encountered a threat while allowing the autonomy to drive, participants failed to respond in 39 cases using the Baseline WMI system indicating a failure rate of 38.24%; using the MMCS-equipped Transparent WMI, participants failed to respond in only 11 cases, indicating a failure rate of 7.91%. These percentages represent the situations in which participants should have responded, but did not.

To directly address H3, we compared participants' response times in both conditions, finding a significant effect of WMI condition, $t(91.45) = 3.65$, $p < 0.001$, $d = 0.744$. Participants were able to respond faster on average using the MMCS-equipped Transparent WMI ($M = 3095.12$ ms, $SD = 1348.64$ ms) than the baseline system ($M = 4267.88$ ms, $SD = 1803.20$ ms), which equated to a 25.21% performance increase (see Fig. 14).

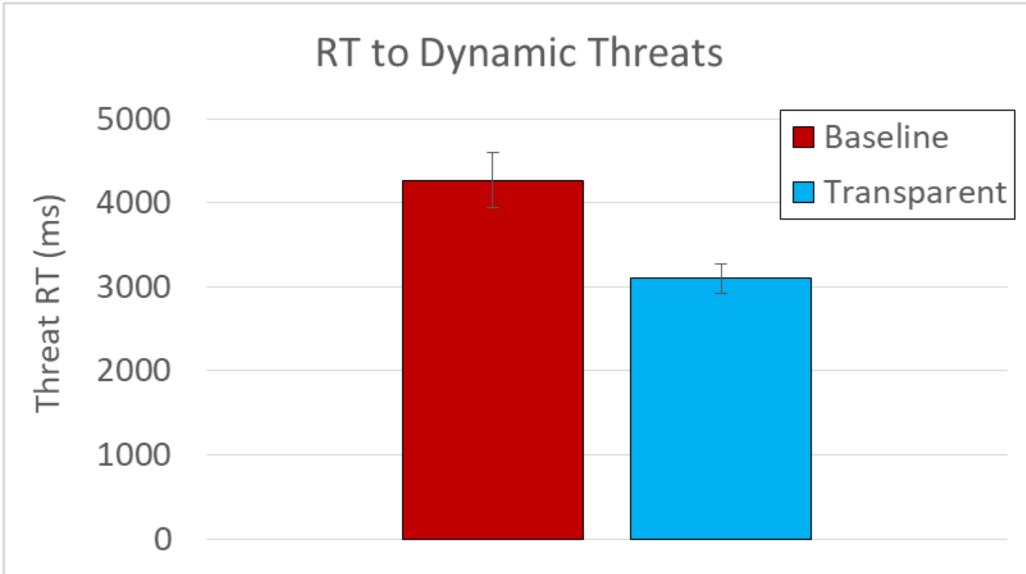


Fig. 14 Participants' mean response times to dynamic threats using the baseline and MMCS-equipped Transparent WMIs. Error bars indicate standard error.

These results support H3 and indicate that the MMCS greatly improves crew members' ability to respond to threats as they populate dynamically on the WMI COP.

The fourth, and related, hypothesis H4 was that the MMCS would improve crew members' ability to stop the vehicle as it was entering dangerous locations in the environment. Using the same trials described in Table 4, with response distance to the threat as the dependent variable, a t-test found a significant effect, $t(122.13) = 2.35$, $p = 0.020$, $d = 0.466$. Participants were able to initiate a response far from dynamic threats when aided by the MMCS ($M = 29.00$ m, $SD = 6.55$ m) than using the Baseline WMI ($M = 25.96$, $SD = 6.54$), corresponding to an 11.74% improvement (see Fig. 15).

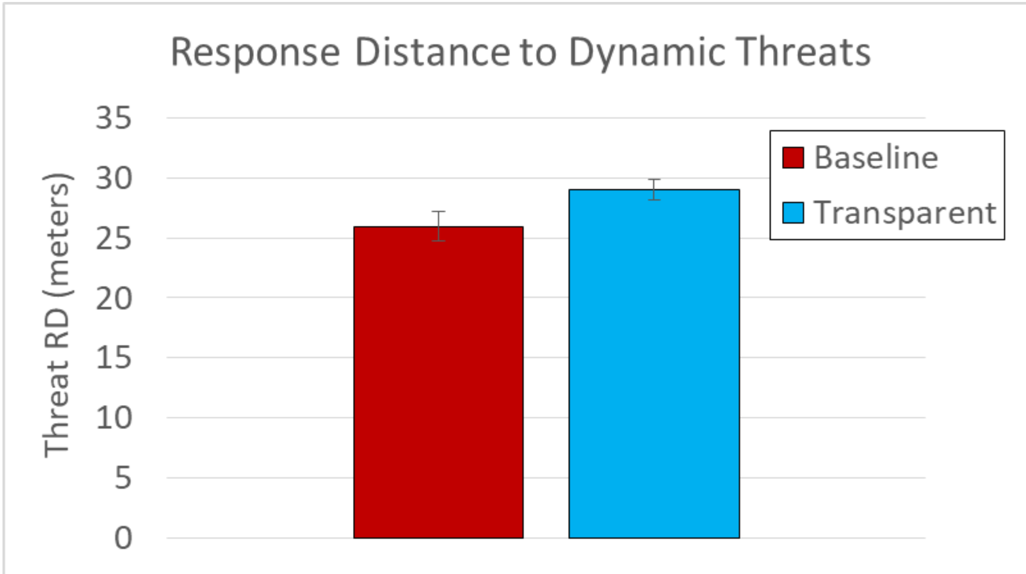


Fig. 15 Participants' mean response distances to dynamic threats in each experimental condition. Error bars indicate standard error.

These results substantiate H4, and indicate that the MMCS improves crew members' ability to engage response farther away from dynamic threats. The results of the analyses in this section suggest that the MMCS improves crew members' ability to access information contained within the WMI COP and respond to it effectively and quickly, in concordance with H3 and H4.

4.3 General Comments and Limitations

First, we must note that the MMCS parameters are customizable, and that the values selected for use in this test were found to be appropriate to active the cues through anecdotal testing. We would expect that an operational system will undergo further refinement and include functionality for the vehicle commander to customize the MMCS criteria in accordance with Mission, Enemy, Terrain, Troops available, Time, and Civilian considerations. Second, we acknowledge that while behavioral measures can be used to infer SA, they do not directly measure it (Endsley 1990, 1995). We opted for a behavioral measure as such metrics are more congruent with the long-term goal of the research program to measure and predict team performance in real time, using methods that could extend to operational settings. Administration of a subjective SA probe, as is frequently done in experimental settings (such as those used in Situation Awareness Global Assessment Technique, known as SAGAT [Endsley 1988]), would be too obtrusive for this application. Third, while the present study appropriately characterizes the performance of an individual, NGCV RCV and Manned Combat Vehicle teams are expected to work in dyads and it might be possible to mitigate some of the burden associated with maintaining SA on the WMI COP while completing other tasks through allocation between the crew members. Future work will need to address these same research questions with at least a dyad, and preferably a section-sized element to determine the effectiveness in a team environment. Finally, some effects observed herein are still not understood; for example, was there a systematic reason that participants revealed more threats when using the MMCS-equipped WMI?

5. Conclusions

With the integration of remotely operated autonomous systems into operations, the US Army faces novel challenges with respect to Soldiers' ability to understand the actions, intentions, goals, and general reasoning of those systems. In summary, the results of this experiment demonstrate the utility of multimodal cueing for improving crew member performance and ability to understand the behavior of such systems during armored vehicle operations. While the Baseline WMI provides NGCV crew members with the ability to monitor and maintain security of their vehicles using the visual modality, supplementing this sensory information with tactile and auditory stimuli can improve the saliency of operationally relevant information, improving crew members' ability to respond to events in dynamic scenarios.

6. References

Baldwin CL, Eisert JL, Garcia A, Lewise BA, Pratt SM, Gonzalez CA. Multimodal urgency coding: auditory, visual, and tactile parameters and their impact on perceived urgency. Work. 2012;41:3586–3591.

Balk SA, Bertola MA, Inman VW. Simulator sickness questionnaire: twenty years later. Proceedings of the Seventh International Driving Symposium on Human Factors in Driving Assessment Training and Vehicle Design 2013; 2013 June 17–20; Bolton Landing, NY. Iowa City (IA): The University of Iowa; c2013. p. 257–263.

Brewster S, Brown LM. Tactons: structured tactile messages for non-visual information display. AUIC 2004. Proceedings of the Australasian User Interface Conference 2004; 2004 Jan 18–22; Dunedin, New Zealand. Darlinghurst (Australia): Australian Computer Society c2004. p. 15–23.

Brown LM, Brewster SA, Purchase HC. A first investigation into the effectiveness of tactons. Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems; 2005 18–20 Mar; Pisa, Italy. p. 167–176.

Calhoun GL, Ruff HA, Draper MH, Guilfoos BJ. Tactile and aural alerts in high auditory load UAV control environments. Proceedings of the Human Factors and Ergonomic Society 49th Annual Meeting; 2005 Sep 1; Orlando, FL. Thousand Oaks (CA): SAGE Publications; c2005. p. 145–149.

Edworthy J, Hellier E, Hards R. The semantic associations of acoustic parameters commonly used in the design of auditory information and warning signals. Ergonomics. 1995;2341–2361.

Edworthy J, Loxley S, Dennis I. Improving auditory warning design: relationship between warning sound parameters and perceived urgency. Human Factors. 1991;33(2):205–231. doi:https://doi.org/10.1177/001872089103300206.

Elliott LR, Coovert MD, Redden ES. Overview of meta-analyses investigating vibrotactile versus visual display options. Proceedings of the 13th International Conference on Human-Computer Interaction; 2009 July 19–24; San Diego, CA. Berlin Heidelberg (Germany): Springer; c2009. p. 435–443.

Elliott LR, Duistermaat M, Redden ES, Van Erp J. Multimodal guidance for land navigation. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2007. Report No.: ARL-TR-4295.

Elliott LR, van Erp J, Redden ES, Duistermaat M. Field-based validation of a tactile navigation device. IEEE Transactions on Haptics. 2010;3(2):78–87.

Endsley MR. Situation awareness global assessment technique (SAGAT). In Proceedings of the National Aerospace and Electronics Conference (NAECON); 1988; 789–795. New York: IEEE.

Endsley MR. Predictive utility of an objective measure of situation awareness. Proceedings of the Human Factors Society 34th Annual Meeting; 1990; Santa Monica, CA. Human Factors Society. p. 41–45.

Endsley, M.R. Toward a theory of situation awareness in dynamic systems. Human Factors. 1995;37(1):32–64.

Ferris TK, Sarter NB. Cross-modal links among vision, audition, and touch in complex environments. Human Factors. 2008;50(1):17–26.

Fitch GM, Hankey JM, Kleiner BM, Dingus TA. Driver comprehension of multiple haptic seat alerts intended for use in an integrated collision avoidance system. Transportation Research Part F: Traffic Psychology and Behaviour. 2011;14(4):278–290.

Fitch GM, Kiefer RJ, Hankey JM, Kleiner BM. Toward developing an approach for alerting drivers to the direction of a crash threat. Human Factors. 2007;49(4):710–720.

Forster B, Cavina-Pratesi C, Aglioti SM, Berlucchi G. Redundant target effect and intersensory facilitation from visual-tactile interaction in simple reaction time. Experimental Brain Research. 2002;143(4):480–87.

Golding JF. Motion sickness susceptibility. Autonomic Neuroscience. 2006;129(1–2):67–76.

Graham R. Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. Ergonomics. 1999;42(9):1233–1248.

Haas EC, Casali JG. Perceived urgency and response time to multi-tone and frequency-modulated warning signals in broadband noise. Ergonomics. 1995;38:2313–2326.

Haas EC, van Erp JB. Multimodal warnings to enhance risk communication and safety. Safety Science. 2014;29–35.

Hartnett GA, Elliott LR, Baraniecki L, Skinner A, Riddle K, Pettitt R. Wearable concepts for gesture based robot control, direction localization, and communication to a tactile belt display. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2018. Report No.: ARL-TR-8597.

Jones LA, Sarter NB. Tactile displays: guidance for their design and application. Human Factors. 2008;50:90–111.

Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. The International Journal of Aviation Psychology. 1993;3(3):203–220.

Laroche C, Quoc HT, Hetu R, McDuff S. Detectsound: a computerized model for predicting the detectability of warning signals in noisy workplaces. Applied Acoustics. 1991;32(3):193–214.

MIL-STD-1472G(1). Design criteria standard: human engineering. Washington (DC): Department of Defense Standards; 2019 [accessed 2019 Mar 31]. https://assist.daps.dla.mil/.

MIL-STD-2525D. Department of Defense interface standard: joint symbology standard. Fort Meade (MD): Defense Information Systems Agency; 2014 June 10.

Murata A, Kuroda T, Karwowski W. Effects of auditory and tactile warning on response to visual hazards under a noisy environment. Applied Ergonomics. 2017;60:58–67.

Nondahl DM, Shi X, Cruickshanks KJ, Dalton DS, Tweed TS, Wiley TL, Carmichael LL. Notched audiograms and noise exposure history in older adults. Ear and Hearing. 2009;30(6):696–703.

Partan SR, Marler P. Issues in the classification of multimodal communication signals. The American Naturalist. 2005;166(2):231–245.

Patterson RD. Guidelines for auditory warning systems on civil aircraft, civil aviation authority paper 82017. Cheltanham (UK): Civil Aviation Authority, Airworthiness Division; 1982.

Perelman BS, Lakhmani S, Wright JL. Human-autonomy teaming essential research program project 2: transparent multimodal crew interface designs. Technical note 1: general overview. Aberdeen Proving Ground (MD): CCDC Army Research Laboratory; 2020a Jan. Report No.: ARL-TN-1003.

Perelman BS, Wright JL, Lieberman GA, Lakhmani S. Human-autonomy teaming essential research program project 2: transparent multimodal crew interface designs. Technical note 2: transparency in mobility planning. Aberdeen

Proving Ground (MD): CCDC Army Research Laboratory; 2020b Jan. Report No.: ARL-TN-1004.

Prewett MS, Elliott LR, Walvoord AG, Coovert MD. A meta-analysis of vibrotactile and visual information displays for improving task performance. IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews. 2012;42:123–132.

R Core Team. R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2014. ISBN: 3-900051-07-0. http://www.R-project.org/.

Sarter NB. Multimodal information presentation: design guidance and research challenges. International Journal of Industrial Ergonomics. 2006;439–445.

Sarter NB. Multimodal displays: conceptual basis, design guidance, and research needs. In: Lee JD, Kirlik A, editors. The Oxford Handbook of Cognitive Engineering. Oxford (UK): Oxford University Press; 2013. p. 556–565.

Scharine AA, McBeath MK. Natural regularity of correlated acoustic frequency and intensity in music and speech: auditory scene analysis mechanisms account for integrality of pitch and loudness. Auditory Perception and Cognition. 2019;1(3–4):205–228.

Stephan KL, Smith SE, Martin RL, Parker SP, McAnally KI. Learning and retention of associations between auditory icons and denotative referents: implications for the design of auditory warnings. Human Factors. 2006;48(2):288–299.

van Erp JBF, van Veen HAHC. Vibrotactile in-vehicle navigation system. Transportation Research Part F. 2004;7:247–56.

van Erp JB. Tactile displays for navigation and orientation: perception and behaviour. Utrecht (Netherlands): Utrecht University; 2007.

van Erp JB, Toet A, Janssen JB. Uni-, bi- and tri-modal warning signals: Effects of temporal parameters and sensory modality on perceived urgency. Safety Science. 2015;1–8.

Vause N, Letowski T, Ferguson L, Mermagen T. Auditory issues in handheld land mine detectors. In: Dubey AC, Harvey JF, Broach JT, Dugan RE, editors. Pro. SPIE 3710, Detection and Remediation Technologies for Mines and Minelike Targets IV; 1999 Apr 5–9; Orlando, FL. p. 778–786. Bellingham (WA): Society of Photo-Optical Instrumentation Engineers; c1999. doi: 10.1117/12.357099.

Appendix. Multimodal Cueing System Algorithm Description

The multimodal cueing system (MMCS) algorithms consists of multiple scripts executing different functions to provide auditory and tactile cues based on real-time information updated from the Warfighter Machine Interface (WMI). Here, we provide high-level descriptions of each script needed to implement MMCS.

BSO tracker: tracks Battle Space Object (BSO) in the environment, updates the BSO list by adding new BSO and remove one from the list when BSO is no longer in the environment.

- **listen_for_bsos**: listens for BSOs on the current Lab Streaming Layer (LSL) stream, and adds the BSO metadata to the shared BSO list.
- **check_threats**: listens to the mobility stream and tracks the given vehicle's position (inherited from the mobility predicate) to determine if the vehicle is near a BSO in the shared list.
- **update_alerts**: resets the BSO values, allowing for appropriate BSOs to be cued at timed interval.

Vehicle tracker: tracks vehicle position through LSL samples, creates an lsl_queue to manage mobility samples.

- **mobility_predicate**: a predicate for a specific vehicle used to retrieve mobility stream.
- **belt_handle**: connects the belt to the PC, write and read data for tactile cues.
- **cue_manager**: resolves outgoing cues and responsible for cues generated by vehicle handler.

Speak: a small script that runs the text-to-speech engine as a subprocess. Given a line of text and an audio file as two arguments, the script plays the audio file using playsound and outputs the text as audio using pyttsx3.

Cue: represents a cue, and any visual, auditory, or tactile data associated with a given cue. If a text-to-speech engine is given, sends an audio cue to the speakers. If a belt device is given, sends a tactile cue to the belt. If a beep sound is given with an audio cue, sends a beep alongside the audio cue once cued and records to the provided LSL stream.

Lsl queue: modified queue implementation to handle mobility samples through a specified buffer size

Cue manager: runs in the foreground and manages the BSO/vehicle trackers, handles and sends cue based on priority.

Bibliography

Abouchacra KS, Breitenbach J, Mermagen T, Letowski T. Binaural helmet: improving speech recognition in noise. Human Factors. 2001;43(4):584 0150-594.

Bolia RS, D'Angelo WR, McKinley RL. Aurally aided visual search in three-dimensional space. Human Factors. 1999;41:664–669.

Brungart DS. Speech-based distance cueing in virtual auditory displays. Proceedings of the IEA 2000/HFES Congress. 2000;44(22)714–717.

Chen JYC, Terrence PI. Effects of tactile cueing on concurrent performance of military and robotics tasks in a simulated multitasking environment. Ergonomics. 2008;51(8):1137–1152.

Chen JYC, Terrence PI. Effects of imperfect automation and individual differences on concurrent performance of military and robotics tasks in a simulated multitasking environment. Ergonomics. 2009;52(8):907–920.

Coovert MD, Gray AA, Elliott LR, Redden ES. Development of a framework for multimodal research: creation of a bibliographic database. Tampa (FL): University of South Florida Tampa; 2007.

Elliott LR, Mortimer B, Hartnett-Pomranky G, Zets G, Mort G. Augmenting Soldier situation awareness and navigation through tactile cueing. In: International Conference on Human Interface and the Management of Information; 2015 Aug 2–7; Los Angeles, CA. Cham (Switzerland): Springer, c2015. p. 345–353.

Elliott LR, Mortimer BJ, Pomranky-Harnett RA, Rapozo F, Pettitt RA, Wooldridge RE, Mort GR. Tactile cues: taction characteristics, salience, ease of learning, and recall. Aberdeen Proving Ground (MD): CCDC Army Research Laboratory; 2019 Feb. Report No.: ARL-TR-8650.

Elliott LR, Mortimer BJ, Pettitt RA, Wooldridge RE. Assessment of wearable tactile system: perception, learning, and recall. Augmented cognition: users and contexts. In: Lecture Notes in Computer Science, vol. 10916. Las Vegas (NV): Springer; 2018. p. 67–77.

Elliott L, Mortimer B, Hartnett-Pomranky G, Zets G, Mort G. Augmenting Soldier situation awareness and navigation through tactile cueing. Proceedings of the 17th International Conference on Human–Computer Interaction; 2015; Los Angeles, CA. p. 345–353.

Elliott L, Redden E. Reducing workload: a multisensory approach. In: Martin J, Allender L, Savage-Knepshield P, Lockett J, editors. Designing Soldier systems; current issues in human factors. Farnham, Surrey, UK: Ashgate;. 2009. Chapter 4. p. 69–96.

Elliott L, Schmeisser E, Redden E. Development of tactile and haptic systems for US infantry navigation and communication. Proceedings of the 14th International Conference of Human–Computer Interaction; 2011; Orlando, FL. p. 399–407.

Elliott L, Skinner A, Pettitt R, Vice J, Walker A. Utilizing glove-based gestures and a tactile vest display for covert communications and robot control. Aberdeen Proving Ground, (MD): Army Research Laboratory (US); 2014 June. Report No.: ARL-TR-6971.

Ferguson L, Mermagen T, Vause N, Letowski T. Auditory signals for enhanced operator performance with hand-held mine detectors. In: Dubey AC, Harvey JF, Broach JT, Dugan RE, editors. Proc. SPIE 4038, Detection and Remediation Technologies for Mines and Minelike Targets V; 2000 Apr 24–28; Orlando, FL. Bellingham (WA): Society of Photo-Optical Instrumentation Engineers; c2000. p. 130–135. doi.org/10.1117/12.396239.

Haas E. Can 3-D auditory warnings enhance helicopter cockpit safety? Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting; 1998 Oct 5–9; Chicago, IL. Thousand Oaks (CA): SAGE Journals. 1998;42:1117–1121.

Haas EC, Pillalamarri RS, Stachowiak CC, Lattin MA. Audio cues to assist visual search in robotic system operator control unit displays. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2005 Dec. Report No. ARL-TR-3632.

Haas E, Edworthy J. An introduction to auditory warnings and alarms. In: Wogalter M, editor, Handbook of Warnings. Mahwah (NJ): Lawrence Erlbaum and Associates; 2006. p. 189–198.

Haas E, Gainer C, Wightman D, Couch M, Shilling R. Enhancing system safety with 3-D audio displays. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Thousand Oaks (CA): SAGE Journals. 1997;41:868–872.

Haas E, Hill S, Stachowiak CF. Designing and evaluating a multimodal interface for Soldier-swarm interaction. Proceedings of the Human Factors and

Ergonomics Society Annual Meeting. Thousand Oaks (CA): SAGE Journals. 2009;53:259–263.

Haas E, Pillalamarri K, Stachowiak C, Fields M. Multimodal controls for soldier/swarm interaction. RO-MAN, 2011 IEEE; 2011. p. 223–228.

Krausman A, White T. Detection and localization of vibrotactile signals in moving vehicles. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2008 May. Report No.: ARL-TR-4463.

Krausman A, Pettitt R, Elliott L. Effects of redundant alerts on platoon leader performance and decision making. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2007 Jan. Report No.: ARL-TR-3999.

Letowski TR, Letowski, ST. Auditory spatial perception: auditory localization. Aberdeen Proving Ground, MD: Army Research Laboratory (US); 2012. Report No.: ARL-TR-6016.

Letowski T, Karsh R, Vause N, Shilling R, Ballas J, Brungart D, McKinley R. Human factors military lexicon: auditory displays. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2001 Sep. Report No.: ARL-TR-2526.

Letowski T, Schmeisser E, Russo D, Rash C. Helmet-mounted displays: sensation, perception and cognition issues. Fort Rucker (AL): Army Aeromedical Research Laboratory (US); 2009.

MIL-STD-1474E. Design criteria standard noise limits. Washington (DC): Department of Defense Standards; 2015 Apr 15 [accessed 2019 Mar 31] https://assist.daps.dla.mil/.

Patterson R. Auditory warning systems for high-workload environments. Ergonomics International 85. Abingdon (UK): Taylor and Francis; 1985:163–166.

Perrott DR, Cisneros J, McKinley RL, D'Angelo WR. Aurally aided detection and identification of visual targets. Proceedings of the Human Factors Society 39th Annual Meeting; 1995. p. 104–108.

Perrott DR, Cisneros J, McKinley RL, D'Angelo WR. Aurally aided visual search under virtual and free-field listening conditions. Human Factors. 1996:38(4):702–715.

Pettitt RA, Redden ES, Carstens CB. Comparison of army hand and arm signals to a covert tactile communication system in a dynamic environment. Aberdeen

Proving Ground (MD): Army Research Laboratory (US); 2006 Aug. Report No.: ARL-TR-3838.

Shilling RD, Shinn-Cunningham B. Virtual auditory displays. In: Stanney K, editor. Handbook of virtual environments. Mahwah (NJ): Erlbaum; 2002. p. 65-92.

Simpson BD, Brungart DS, Gilkey RH, McKinley RL. Spatial audio displays for improving safety and enhancing situation awareness in general aviation environments. Meeting Proceedings RTO-MP-HFM-123; 2005. Neuilly-sur-Seine (France): RTO. Paper 26. p. 26-1–26-16. http://www.rto.nato.int/abstracts.aps.

Simpson BD, Brungart DS, Gilkey RH, Cowgill JL, Dallman RC, Green RF, Youngblood KL, Moore TJ. 3-D audio cueing for target identification in a simulated flight task. Human Factors and Ergonomics Society 48th Annual Meeting; 2004 Sep 20–24; New Orleans, LA. p. 1836–1840.

Vause N, Abouchacra K, Letowski T, Resta E. Spatial (3-D) audio displays improve speech communication in the command and control vehicle (C2V). Proceedings of the 22nd Army Science Conference; 2000 Dec 11–13; Baltimore, MD. p. 834–841.

White T, Krausman A, Haas E. Tactile displays in Army operational environments. In: Martin J, Allender L, Savage-Knepshield P, Lockett J, editors. Designing Soldier systems; current issues in human factors. Farnham, Surrey (UK): Ashgate; 2012. p. 97–116.

List of Symbols, Abbreviations, and Acronyms

ANVEL	Autonomous Navigation Virtual Environment Laboratory
ARL	Army Research Laboratory
AZF	arena zone fragment
BSO	Battle Space Object
CCDC	US Army Combat Capabilities Development Command
COP	Common Operating Picture
DP1	Decision Point 1
EAI	Engineering Acoustics, Inc.
EM	electromagnetic
ERP	Essential Research Program
H1	Hypothesis 1
H2	Hypothesis 2
H3	Hypothesis 3
H4	Hypothesis 4
HAT	Human-Autonomy Teaming
IED	improvised explosive device
INFORMS	Information for Mixed Squads
ISI	inter stimulus interval
ISR	intelligence, surveillance, and reconnaissance
LSL	Lab Streaming Layer
METT-TC	Mission, Enemy, Terrain, Troops available, Time, and Civilian considerations
MMCS	multimodal cueing system
MSSQ	Motion Sickness Susceptibility Questionnaire
NGCV	Next Generation Combat Vehicle
RCV	robotic combat vehicle

RPG	rocket-propelled grenade
SA	situation awareness
SSQ	Simulator Sickness Questionnaire
TCU	tactile control unit
TTS	text-to-speech
USB	Universal Serial Bus
WMI	Warfighter Machine Interface

1 (PDF) DEFENSE TECHNICAL
INFORMATION CTR
DTIC OCA

1 (PDF) CCDC ARL
FCDD RLD CL
TECH LIB

1 (PDF) CCDC ARL
FCDD RLH B
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL
35898-7290

1 (PDF) CCDC ARL
FCDD HSI
J THOMAS
6662 GUNNER CIRCLE
ABERDEEN PROVING
GROUND MD
21005-5201

1 (PDF) USAF 711 HPW
711 HPW/RH K GEISS
2698 G ST BLDG 190
WRIGHT PATTERSON AFB OH
45433-7604

1 (PDF) USN ONR
ONR CODE 341 J TANGNEY
875 N RANDOLPH STREET
BLDG 87
ARLINGTON VA 22203-1986

1 (PDF) USA NSRDEC
RDNS D D TAMILIO
10 GENERAL GREENE AVE
NATICK MA 01760-2642

1 (PDF) OSD OUSD ATL
HPT&B B PETRO
4800 MARK CENTER DRIVE
SUITE 17E08
ALEXANDRIA VA 22350

ABERDEEN PROVING GROUND

13 (PDF) CCDC ARL
FCDD DAH C
L GARRETT
FCDD DAH W
F MORELLI
FCDD RLH
J LANE
Y CHEN
P FRANASZCZUK
K MCDOWELL
K OIE
A MARATHE
FCDD RLH BD
D HEADLEY
FCDD RLH F
J GASTON
FCDD RLH FA
A DECOSTANZA
FCDD RLH FC
D CHHAN
FCDD RLH P
A EVANS