ASSESSING THE IMPACT
OF
AUDITORY PERIPHERAL DISPLAYS FOR UAV OPERATORS

H.D. Graham
M.L. Cummings

MASSACHUSETTS INSTITUTE OF TECHNOLOGY*

PREPARED FOR CHARLES RIVER ANALYTICS, INC.

HAL2007-09

NOVEMBER 2007

Table of Contents

List of Figures

List of Tables

ABSTRACT

A future implementation of unmanned aerial vehicle (UAV) operations is having a single operator control multiple UAVs. The research presented here explores possible avenues of enhancing audio cues of UAV interfaces for this futuristic control of multiple UAVs by a single operator. This project specifically evaluates the value of continuous and discrete audio cues as indicators of course deviations or late arrivals to targets for UAV missions. It also looks at the value of the audio cues in single and multiple UAV scenarios.

To this end, an experiment was carried out on the Multiple Autonomous Unmanned Vehicle Experimental (MAUVE) test bed developed in the Humans and Automation Laboratory at the Massachusetts Institute of Technology with 44 military participants. Specifically, two continuous audio alerts were mapped to two human supervisory tasks within MAUVE. One of the continuous audio alerts, an oscillating course deviation alert was mapped to UAV course deviations which occurred over a continual scale. The other continuous audio alert tested was a modulated late arrival alert which alerted the operator when a UAV was going to be late to a target. In this case the continuous audio was mapped to a discrete event in that the UAV was either on time or late to a target. The audio was continuous in that it was continually on and alerting the participant to the current state of the UAV. It either was playing a tone indicating the UAV was on time to a target or playing a tone indicating the UAV was late to a target. These continuous alerts were tested against more traditional single beep alerts which acted as discrete alerts. The beeps were discrete in that when they were used for monitoring course deviations a single beep was played when the UAV got to specific threshold off of the course or for late arrivals a single beep was played when the UAV became late.

The results show that the use of the continuous audio alerts enhances a single operator's performance in monitoring single and multiple semi-autonomous vehicles. However, the results also emphasize the necessity to properly integrate the continuous audio with the other auditory alarms and visual representations in a display, as it is possible for discrete audio alerts to be lost in aural saliency of continuous audio, leaving operators reliant on the visual aspects of the display.

INTRODUCTION

In current military conflicts, the unmanned aerial vehicle (UAV) is one of the most highly sought after air assets (Whitney, 2007). UAVs provide real time information and support to battlefield operations remotely from either local or global control station platforms. Each of these operational environments presents fluid atmospheres of noise and activity.

Currently many of the medium and high altitude UAV operations require multiple operators to control a single UAV; the desire for the future is to invert the numbers and have a single operator controlling multiple UAVs (Culbertson, 2006). To achieve this, research is needed to understand how best to aid and integrate the human operator into a system of semi-autonomous UAVs. Due to potential information overload of operators in visually saturating domains such as UAV supervisory control, reducing the workload by offloading some cognitive processing to the other senses, primarily auditory and haptic, is critical. The point of this research is to explore possible auditory enhancements that can reduce operator workload and increase operator performance.

BACKGROUND

Improving UAV operator performance via audio enhancements has been an active area of research, particularly in the areas of spatial audio, which has been shown to aid UAV operators in target recognition, and 3D audio, which has been shown to enhance alert and warning detection (for a review of the literature, see (Nehme & Cummings, 2006)). One area in the audio domain that has received less attention is the use of continuous audio peripheral displays, particularly in complex, time pressured domains such as UAV operation in command and control settings.

Research in the medical domain indicates that continuous auditory displays (also known as sonifications) can aid anesthesiologists in monitoring patient status (Watson & Sanderson, 2004). Deviations from normal breathing patterns matched to musical notes allowed doctors to aurally detect a difference in patient status, while visually focusing on other tasks. Given that such an approach has worked in this medical monitoring domain, it is not clear whether continuous auditory displays can be applied to a command and control setting, particularly for the supervisory control of one or more unmanned vehicles by a single operator.

While continuous audio displays can allow operators to effectively divide attention so that disparate information can be taken in through different channels (such as visual and aural), the environment plays a major role. In the case of the anesthesia study, operating rooms are generally not as noisy as those settings typically found in command and control environments. For example, Shadow UAV operators must work in the back of a truck with significant noise due to nearby generators. In addition, unlike an operating room, UAV operators wear headsets so that they can communicate with both inter and intra team members. Also unlike the operating room, which may be connected to an intercom in another room, there is significant communication coming into UAV operators from the outside world. UAV operators must contend with an air traffic control-like environment, constantly scanning the audio landscape for possible incoming

communications, as well as listening to general radio chatter in order to gain situational awareness for the overall mission.

Given these significant UAV environmental constraints, which are tantamount to noisy working conditions, the design of an audio display, particularly a continuous one, is difficult. Moreover, a sonification should map to the task, i.e. perceptual changes should accurately match the change in the data (Barrass & Kramer, 1999). Thus, discrete events should map to discrete signals, and continuous events should map to a continuous auditory display. In addition, in high workload environments, whether sonifications can improve performance, or whether operators will ignore the auditory displays and rely primarily on the visual, is not clear. In order to investigate these issues, an experiment was conducted which is detailed in the remainder of the paper.

THE MAUVE SIMULATOR

The Massachusetts Institute of Technology's Humans and Automation Laboratory (HAL) developed the Multiple Autonomous Unmanned Vehicle Experimental (MAUVE) test bed, which is a two screen interface that allows the simulation of any number of unmanned vehicles conducting strike operations in hostile environments. MAUVE, for this experiment, allowed the operator to arm and fire on targets with various priorities while monitoring each UAV's flight of path, event timeline, and radio traffic.

The test bed provided a map for geo-spatial tracking of UAVs on their preset flight paths along with an interaction control panel for making UAV control inputs (Map Display, Figure 1). Control inputs included arming and firing a UAV to destroy targets, directing specific path changes, and acknowledging radio calls. The path changes included commanding a UAV to skip a target to get a UAV back on schedule. The operator could also reset the UAV's onboard navigation to get the vehicle back on course if it drifted.

A timeline for each UAV, a scheduling decision support tool for each UAV, and chat interfaces were also provided (Timeline Display, Figure 1). This display helped operators ensure UAVs would be on time to targets by monitoring when they would arrive at targets, and in what phase they were currently located. In particular, the timeline used the color-coding provided in Table 1 to represent phase of flight.
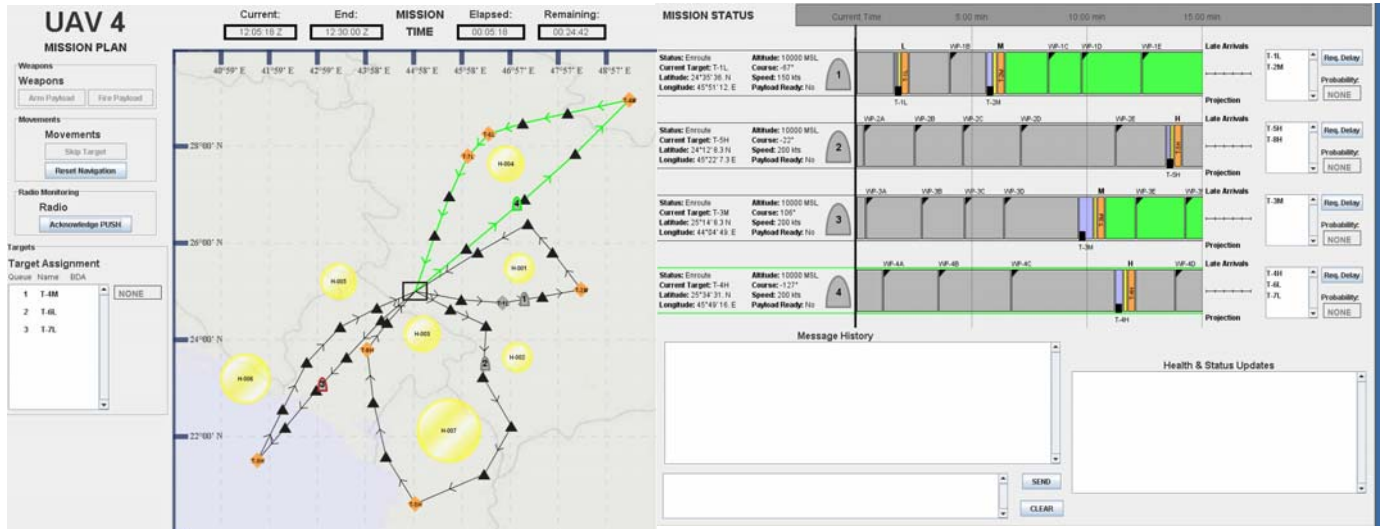
Map Display	**Timeline Display**

Figure 1: Multiple Autonomous Unmanned Vehicle Experimental (MAUVE) test bed.

Table 1: UAV Color-Coded Flight Phases

UAV Action	Color
Enroute	Gray
Loitering	Blue
Arming Payload	Yellow
Firing Payload	Orange
Return to Base	Green

The general task for MAUVE operators for this experiment was to monitor one or more UAVs on a preplanned mission involving several waypoints and targets. The primary task of the operator was to ensure weapons were dropped on the correct targets during the specified windows of opportunity. However, unexpected head winds or crosswinds could cause UAVs to slow their speed or drift off course, requiring the operator to take corrective action for projected target late arrivals or course deviations. Late arrivals were a higher priority than course deviations. In addressing late arrivals, the priority of the target (high, medium, or low) also had to be addressed, such that it was more important to hit a high priority target than a medium priority target, and so on. In addition, operators also had to ensure the UAVs would return to base by a specified time.

4

SONIFICATION DESIGN AND INTEGRATION

Four auditory alerts were created by Charles River Analytics, Inc. for incorporation into MAUVE, which represent experimental independent variables. These four auditory alerts were the 1) threshold course deviation alert, 2) the oscillating course deviation alert, 3) the threshold late arrival alert, and 4) the modulated late arrival alert. These four auditory alerts were designed to test the differences between discrete and continuous alarms for the two conditions of late arrivals and course deviations. The threshold alerts were designed to be discrete alarms; they emitted a single discrete audio alert when some precondition was met. The oscillating and modulated alerts were designed to be continuous alarms; they provided continual audio information about the state of a particular variable within MAUVE. Instead of indicating a clear distinct threshold, the continuous and modulating alerts gradually increased as the severity of either the late arrival or course deviation condition increased.

The threshold course deviation alert consisted of a single beep with a fundamental frequency of 1000 Hz. The beep lasted ~8 ms and played whenever a UAV moved completely away from its plotted course line.

The oscillating course deviation alert represented both the existence and severity of the UAV course deviations. It consisted of comb filters that were applied to a mix of pink noise and the ambient signal. Using pink noise means that filters equalized the energy of sound at each octave level, resulting in a constant level presentation of the audio. The mix ranged from 0.2 pink noise for low deviation to 0.9 pink noise for high deviation. The comb filters had a base delay of 0.2 ms, with a 50 percent each mix of the base signal and a feed forward delay. The delay values were then oscillated to create a periodic signal. Because this was a continuous audio scheme, it played continually to provide an auditory image of UAV path position. As a UAV drifted further off course, the frequency of oscillation of the comb filter delay decreased from 17 Hz to 4.5 Hz, and the depth of oscillation increased from 0.2 ms to 0.7 ms, thus changing on a continual scale.

The threshold late arrival alert consisted of a single beep with a fundamental frequency of 415 Hz. The beep lasted ~18 ms, and played whenever a UAV was projected to be late to any target. A UAV was determined late to a target when it had exceeded its window of opportunity to destroy the target and was not going to be able to get over the target in time to fire on the target during the set time window it was required to be at the target.

The modulated late arrival alert consisted of discrete harmonic signals continuously playing to indicate when a projected late arrival at a target. These sounds played continually until the operator addressed the issue with the corrective action of choosing to delay the UAV, skip a target, or the issue disappeared because the UAV automatically continued past the target when it was going to be late to the next target. The harmonic signals were composed of five Formant filters that were applied to a mix of pink noise and the ambient signal. During the condition of no late arrivals, a baseline audio condition was generated with two filters set to 261.6 Hz, two filters set at 329.4 Hz, and one filter set at 392.0 Hz (a major C triad with extra balance in the root and third). If it was late to a low priority target, a signal was generated with two filters set to 261.6 Hz, two filters set at 311.1 Hz, and one filter set at 392.0 Hz (a minor C triad with extra balance in the root and third). If it was late to a medium priority target, a signal was generated with three filters set to 293.6 Hz and two filters set at 415.3 Hz (a tritone interval with a D root). If the

UAV was predicted to be late to a high priority target, a signal was generated with three filters set to 369.9 Hz and two filters set at 523.25 Hz (a tritone interval with an F# root). As the priority increased, the pink noise mix also increased, from 0.25 for the baseline, to 0.7, 0.8, and 1.0 for the three priority levels.

All audio alerts were presented equally in both ears through the Sensimetrics HDiSP. Pictured in Figure 2, the HDiSP is an over-the-head headset with sealed around-the-ear ear cups. It provides ambient noise attenuation ranging from 35 to 40 db for the spectrum between 500 and 8 kHz. The headset has integrated digital signal processors, which produce the signals used in this experiment based on generative audio filters and processing of the ambient signal, which was received from microphones mounted on the headband.

Figure 2: Sensimetrics HDiSP.

EXPERIMENTAL PROTOCOL

Experiment Objective

The objective of this experiment was to determine whether sonifications maximize the information conveyed to UAV operators more efficiently than typical discrete alarms typically used in current ground control stations. In addition, the impact of continuous versus discrete alerting on operators attempting to control multiple UAVs was also a research question.

Research Hypothesis

There were three null hypotheses explored in this research:

1. For various combinations of discrete and continuous alerts, there was no difference in the operator's performance.
2. For operators controlling either a single or multiple UAVs, there was no difference in the operator's performance.

3. There was no interaction between the discrete and continuous alerts combinations and whether the operator was controlling one or more UAVs.

The alternative hypotheses for the first two null hypotheses, therefore, were that there is a difference, and for the third null hypothesis, that there was interaction.

Task

For this experiment, the operator was given the following as the primary objective for each mission:

Make sure the UAV(s) maximize the number of targets engaged as well as arrive back at the base safely.

Further, supervision of each of the UAV(s) for each mission was broken down into the following prioritized sub-tasks, from highest priority to lowest:

1. Return to base (RTB) within the time limit for the mission (this limit was clearly marked).
2. Comply with recovery rules for course deviations.
3. Comply with recovery rules for target late arrivals.
4. Destroy all targets before their time on target (TOT) window ended.
5. Avoid taking damage from enemies by avoiding all threat areas.
6. Acknowledge all "Push" radio calls.

Participants were trained to follow these priorities, and the objectives and priorities above were displayed to them throughout the experiment. Within the priorities, recovery rules of how to recover the UAV from a course deviation or late arrival were also displayed. See Appendix A for pictorial descriptions of a course deviation and late arrival. The instructions on how to comply with recovery rules for course deviations were to click "Reset Navigation." This theoretically reset the onboard navigation in the UAV and caused it to fly back to the planned flight path.

The instructions on how to comply with recovery rules for target late arrivals were:

- Low Priority – Click "Skip Target"
- Medium Priority – Click "Skip Target" _or_ Employ the decision support visualization (DSV).
- High Priority – Employ DSV before Requesting Delay or Clicking "Skip Target"

For low priority targets, clicking the "Skip Target" button would cause the UAV to proceed immediately to the next target. For medium and high priority targets, the option was given to use the DSV, which was a tool that illustrated the effect that delaying the time on target for the one late target had on the remaining targets for that UAV's flight plan. With the assistance of the information presented in the DSV, participants then decided whether they should click the "Skip

Target" button or "Request Delay" button. If they clicked "Request Delay," they then had to wait to see if a delay was granted, and if the delay was granted the UAV was then no longer late to the target and would be able to destroy it. For a high priority target the immediate response was to employ the DSV.

Independent Variables

The experiment was a 4x2 fixed factor repeated measures model, with two independent variables: the audio condition (a between-subjects treatment), and the number of vehicles under control (a repeated within-subjects factor). The four levels of the audio condition factors were combinations of the four auditory alerts described earlier. The audio conditions were: 1) the threshold audio condition for both the late arrivals and course deviations, 2) the continuous oscillating course deviation audio condition with threshold alerts for the late arrivals, 3) the continuous modulated late arrival audio condition with a threshold alert for course deviations, and 4) the combination audio condition which consisted of the oscillating course deviation alert and the modulated late arrival alert.

The second independent variable, the number of vehicles under control, had two levels: single UAV and multiple UAV. In the single level, the participant supervised only one UAV, while in the multiple factor level, the participant supervised four UAVs. The framework of this experiment is represented in Table 2.

Table 2: Experimental Design.

Scenario (Repeated)

		Single UAV	Multi UAV
Audio Scheme (Between)	Both Threshold	--	--
	Continuous Course Deviation, Threshold Late Arrival	--	--
	Continuous Late Arrival, Threshold Course Deviation	--	--
	Combination	--	--

Dependent Variables

Dependent variables included the reaction time to correct course deviations, the number of missed course deviations, the reaction time to correct projected late arrivals, the number of missed late arrivals, NASA TLX scores, and number of missed radio calls. These are discussed in detail below.

1. Course deviation reaction time

In each test scenario, participants were expected to respond to four triggered course deviations. A course deviation was defined as when a UAV no longer appeared to be following its planned flight path. A general rule for defining a course deviation was a UAV that no longer appeared over the flight path's line or a UAV completely left the flight path.

2. Course deviation errors of omission

Anytime the participant failed to respond to one of four triggered course deviations, an error of omission was counted.

3. Late arrival reaction time

Four late arrivals occurred in each test scenario. Late arrivals were caused when a UAV slowed down because of unforeseen headwinds and was no longer able to reach a target in time to complete the firing mission.

4. Late arrival errors of omission

Errors of omission for late arrivals occurred when the participant failed to respond to one of the four triggered late arrivals.

5. Radio calls missed (secondary workload assessment)

As a secondary workload task, the number of missed radio was an indication of the operator's level of mental workload. The count of missed radio calls measured spare mental capacity. Participants were instructed to monitor a recording of continual air traffic radio chatter for the word "Push." The word "Push" occurred 62 times in a 30 minute session, with an average of 27 seconds between each "Push" radio call. To acknowledge the radio call, participants clicked an "Acknowledge Push" button on the display.

6. NASA TLX score (subjective workload assessment)

The NASA Task Load Index gathered participants' subjective assessment on a scale of 1 to 20 of mental demand, physical demand, temporal demand, effort, performance, and frustration. A participant was then asked to rate these six dimensions against each other to determine their importance in the participant's workload (Hart & Staveland, 1988). Since there was no physical demand in this experiment, each participant was told to rate the physical demand so that it was zeroed out of the score. No difference between the single and multiple UAV scenario was expected because the scenarios were designed with equal workloads. Appendix B shows how the single and multiple UAV scenarios were designed with the same events. The only difference was that the multiple UAV scenario has the events divided over 4 UAVs while the single scenario has them all occurring with 1 UAV.

Apparatus

The experiment was administered in a testing room with a background ambient noise level of ~64dB on a C-scale. MAUVE was run on a multi-modal workstation (MMWS). Pictured in Figure 3, the MMWS is a four screen computer display.

Figure 3: The Multi-Modal Workstation (MMWS).

The three screens across the top were 21 inches and set to display at 1280 x 1024 pixel, 16 bit color resolution, and the 15 inch bottom center screen was set at 1024 x 768 pixels, 32 bit color resolution. The computer used to run the simulator was a Dell Optiplex GX280 with a Pentium 4 processor and an Appian Jeronimo Pro 4-Port graphics card. Participants controlled the simulator through a generic corded computer mouse. Throughout the experiment, the top left display showed the mission objectives and priorities, and the bottom center displayed the color-coding for MAUVE (Table 1). The top center display contained the left MAUVE map and interaction display, and the top right display included the right MAUVE timeline and decision support display (Figure 1). During the test sessions, the screen recordings for the left and right MAUVE displays were captured by Camtasia ® recording software.

Participants

Forty-four military personnel were paid $10 an hour to participate. The experiment took anywhere from 2.5 to 3 hours to complete. Participants' ages ranged from 20 years to 42 years, with an average age of 26 years. There were 3 Navy midshipmen, 1 Army reservist specialist, 1 Air Force staff sergeant, and 39 officers from the Army/Navy/Air Force. Overall, the personnel

tested had a combined experience of over 250 years of active duty military service, with each member having an average of 5.8 years of active duty service. Five pilots contributed as test participants, and most of the 28 junior officers tested will be future military pilots.

For the data analysis, 5 participants were dropped because of problematic data. The first two were omitted because of the failure of the test proctor to administer proper training. The third was dropped because of a 3 day interruption between the single and multiple UAV test sessions. The fourth participant was not used because his secondary data was an outlier for both the single and multiple UAV scenarios. The fifth participant dropped was an active duty Air Force lieutenant colonel with 4 years experience as a maintenance officer and an additional 4 years of flight line time as a pilot and flight test engineer. He reported having been diagnosed with tinnitus and loss of high frequency tone in one ear. During the audio training, he reported not being able to hear the differentiation between a low and high course deviation with the oscillating course deviation alert. His hearing test results indicated age induced hearing loss.

Procedure

Each participant experienced three separate consecutive phases in completing the experiment. Each had a 60 to 70 minute training session followed by a 70 to 80 minute test session, and then a 10 minute post-test survey. For the training session, each participant completed a participant consent form and demographics survey, received standardized training from a PowerPoint® tutorial, a hearing test, audio training for the audio condition they were presented, and a thirty minute practice test session with a trial run through of NASA TLX. If required, a participant could do an additional 10 minute practice session to ensure understanding of the proper MAUVE control actions. Copies of the participant consent form, demographics survey, tutorial, and post-test survey are in Appendices C,D, E, and F respectively.

The hearing test was designed to identify whether the participant suffered any hearing loss (either temporary or permanent). For the hearing test, participants listened to a CD of test tones from 40dBSPL down to 16dBSPL, with 5 frequencies tested (500 Hz, 1 kHz, 2 kHz, 4 kHz, and 8 kHz). The listener acknowledged how many beeps they heard for each ear and each frequency. The range tested ensured they could hear to within 3dB of the noise floor of the headset in the normal lab conditions. Based on the method of descending limits, the test was designed so that hearing 5 beeps or more at each frequency showed adequate hearing (Snodgrass, 1975; Wundt, 1902) If fewer beeps were heard at 8 kHz than at 1 kHz, participant exhibited high frequency loss (Cooper & Owen, 1976; Humes, Joellenbeck, & Durch, 2006). Age related permanent loss showed up most at high frequencies (i.e. 8 kHz congestion and other losses show up at lower frequencies and across range) (Walden, Prosek, & Worthington, 1975). The hearing test was generated by Dr. Thomas Edward von Wiegand of Sensimetrics Corporations, in conjunction with the HDiSP.

The audio training gave each participant exposure to all of the audio alerts they would hear. They were presented with their respective late arrival and course deviation alerts, as well as the radio chatter. After introducing each of these individually, they were played together to give the full effect. After a demonstration of the audio condition, participants were encouraged to play

with the audio demonstration software until they were comfortable with recognizing each of the sounds.

After the training was completed, each participant then completed both the single UAV scenario and the multiple UAV scenario. The experiment was counterbalanced; half of the participants completed the single UAV scenario first, and the other half finished the multiple UAV scenario first. Additionally, the participants were randomly assigned to the four audio schemes to maintain a balanced study. At the conclusion of every test session, MAUVE generated a data log, with all the previously discussed dependent variables. Additionally, each participant completed a NASA TLX following each test session. At the conclusion of the test session, each participant completed a post-test survey, was paid, and thanked for their time and involvement.

RESULTS

For statistical analysis, repeated measures ANOVAs were conducted for each dependent variable. A summary of the dependent variables with significant results is presented in Table 3. The audio scheme factor is a between-subjects factor for the four audio conditions and the scenario is a repeated within-subjects factor of single versus multi UAV levels. The current data analysis is based on 39 participants, which meets the 23 participant requirement of the *a priori* power analysis for a power of 0.80. Excluding the late arrival reaction times and error of omission counts, all of the remaining variables analyzed met normality and homogeneity assumptions. While $\alpha = 0.05$, p values between 0.05 and 0.10 are considered marginally significant. See Appendix G for further details on the statistical analysis tests.

Table 3: Summary of experiment statistical results.

Dependent Variable	Audio Scheme	Scenario	Interaction
Course Deviation Reaction Times	$F_{(3,35)}=2.878$ p=.05	$F_{(1,35)}=3.215$ p=.08	NS
Late Arrival Reaction Times	$F_{(3,35)}=3.345$ p=.03	$F_{(1,35)}=20.737$ p<.001	NS

1. Course deviation reaction time

Figure 4 shows the means for the course deviation reaction times across the four audio conditions (combination, continuous late arrival, continuous course deviation, and threshold). The omnibus test shows there is a statistically significant difference in the performance of participants based on the audio scheme ($F_{(3,35)}=2.878$, p=.05), and a marginally significant difference due to the scenario ($F_{(1,35)}=3.215$, p=.08). A Tukey post hoc comparison shows that the primary points of difference are for the combination to threshold audio conditions for the single UAV scenario (p=0.02). There is not a significant difference between any of the other points.
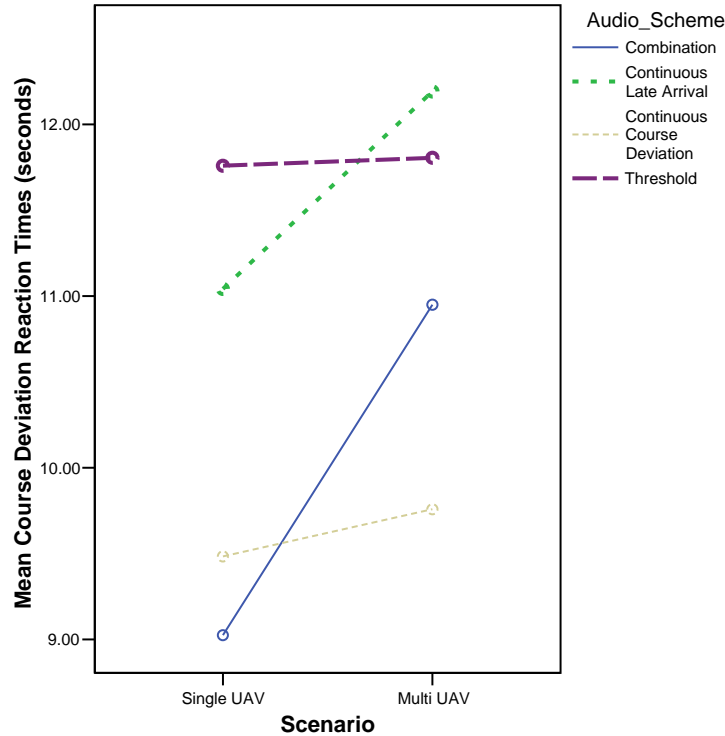
Figure 4: The treatment means plots for the course deviations reaction times.

In Figure 4, there is an apparent division in the data based on which audio alert is present for the course deviation. As part of the post hoc analysis based on the data clustering in Figure 4, the model was reconfigured as a 2x2 fixed factor repeated measures model with audio scheme and scenario as the independent variables, respectively (Table 4). Instead of the four audio conditions, the factor levels were collapsed into two levels: those exposed to the oscillating course deviation alert and those exposed to the threshold course deviation alert. The audio scheme remained still a between-subjects factor while the scenario remained the same as a repeated within-subjects factor.

Table 4: Post Hoc Audio Alert Experimental Design

		Scenario (Repeated)	
		Single UAV	Multiple UAV
Audio Scheme (Between)	Oscillating Course Deviation Alert	--	--
	Threshold Course Deviation Alert	--	--

The combination and continuous course deviation audio conditions use the oscillating course deviation alert condition while the threshold course deviation alert condition is used in the continuous late arrival and threshold audio conditions. As seen in Figure 5, the analysis confirms that the participants with the continuous oscillating alert condition performed significantly better than those with the threshold course deviation alert condition ($F_{(1, 37)} = 8.874$, p=.01).
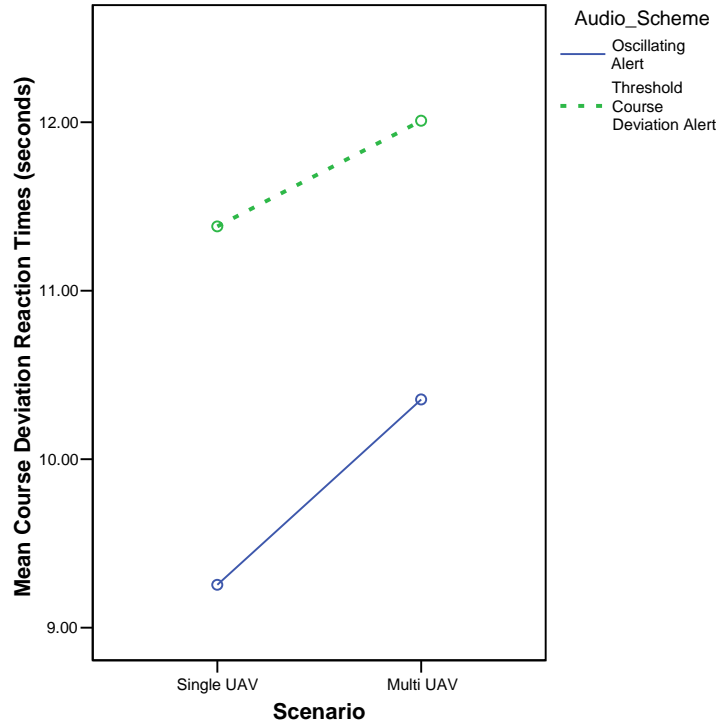
Figure 5: The treatment means plots for the two alert conditions for the course deviations reaction times.

2. Course deviation errors of omission

Out of the 156 course deviations presented to the 39 participants, 9 were missed (5.8%). The data is not normal and the non-parametric tests (Wilcoxon-Signed Rank, Kruskal-Wallis) show no significant differences.

3. Late arrival reaction time

The original late arrival reaction times are not normal, thus a natural log transformation was performed. The omnibus results show a significant difference in the performance of participants between the four audio conditions ($F(3,35)=3.345$, $p=.03$) and a significant difference due to the scenario ($F(1,35)=20.737$, $p<.001$).

As seen in Figure 6, there is an apparent split in the data. A post hoc Tukey comparison reveals there is significant and marginally significant difference between the data for the continuous course deviation audio condition and the other three audio conditions: threshold audio condition ($p=.05$), continuous late arrival audio condition ($p=.08$), and combination audio condition ($p=.07$). Point-to-point comparisons reveal the difference in the audio scheme is not significant for any of the single UAV comparisons, but for the multiple UAV audio comparisons, there is a significant difference between the combination and continuous course deviation audio conditions ($p=.02$), the continuous course deviation and continuous late arrival audio conditions ($p=.06$), and the continuous course deviation and threshold audio conditions ($p=.07$). In addition, there is a significant difference in performance from the single to the multiple UAV scenarios for the combination audio condition ($p=.04$) and the continuous course deviation audio condition

14

(p=.03). These post hoc comparisons confirm the apparent differences within the treatment means plot in Figure 6.
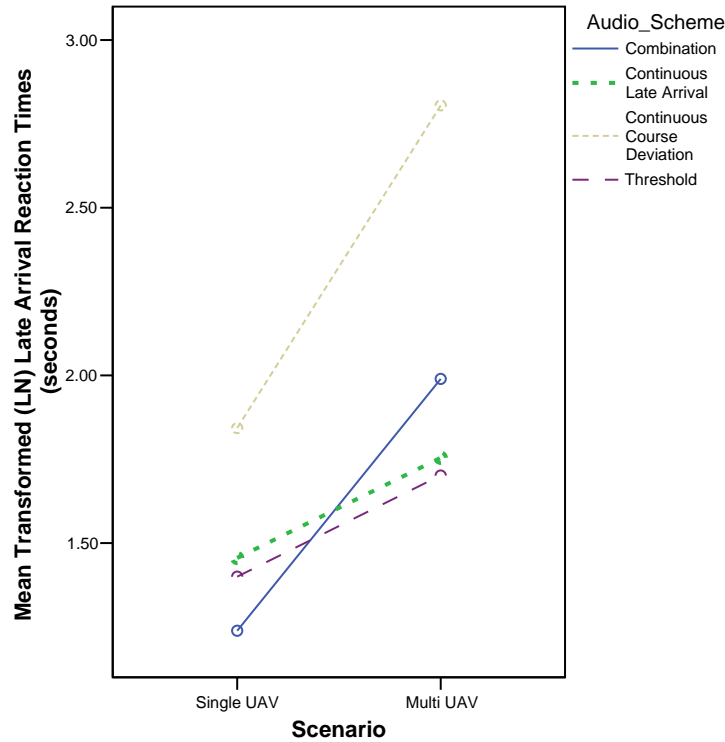

Figure 6: The treatment means plots for the transformed (LN) late arrival reaction times.

In fact, when the combination, continuous late arrival, and threshold audio conditions are combined as one condition and compared to the continuous course deviation audio condition the results show a significant difference based on audio condition ($F(1,37) = 10.528$, $p<.001$) and scenario ($F(1,37) = 23.730$, $p<.001$).

4. Late arrival errors of omission

All 39 participants responded to all of the 4 late arrivals presented to each of them, thus there were no late arrival errors of omission.

5. Radio calls missed (secondary workload assessment)

For the number of missed radio calls, there is no significant difference between the audio conditions or across the scenarios. According to this secondary workload measurement, the participants appear to be equally work saturated regardless of their assigned audio condition, or whether they supervised a single or multiple UAVs. The box plots in Figure 7 show larger variations of performance in the multi UAV scenario as compared to the single scenario for all audio conditions except the threshold audio condition.
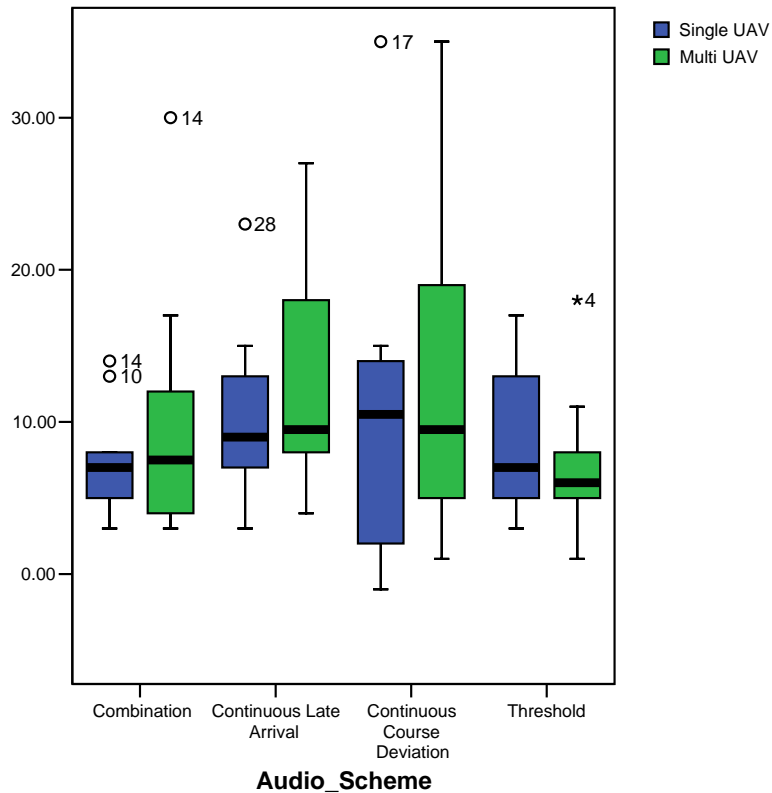
Figure 7: The box plots of the number of missed radio calls

6. NASA TLX score (subjective workload assessment)

Similar to the secondary workload measurement, the subjective measurement of the NASA TLX scores are not significantly different for the audio condition or scenario. Because these are subjective scores, this tool is best for comparing within-subjects across the scenarios, so this result is not surprising.

DISCUSSION

The results of this experiment show that sonifications for two different events, the combination audio scheme, promoted the best performance for the two reaction time dependent variables. This combination scheme consisted of oscillating course deviations alerts and modulated late arrival alerts, and as will be discussed below, their integration was just as critical as their existence for the individual alerts.

The post hoc analysis shows that the oscillating course deviation alert is beneficial in controlling both a single and multiple UAVs. Overall, participants were quicker to respond to course deviations with the oscillating course deviation alert than with the threshold course deviation alert. This was true when a continuous alert was used either just for the course deviation scenario as well as when combined with the late arrival continuous alert. One reason for the

16

longer course deviation reaction times for the threshold alert may be that participants who have visually identified the course deviation may wait for the threshold alert as confirmation. With the oscillating course deviation, the auditory confirmation comes sooner as they integrate the continuous audio with the visual display. One participant reported doing this. On the whole, most participants reported relying on the audio as their primary alert. They did not even visually monitor for a course deviation until they were alerted by the audio. The better performers, therefore, were the participants with the oscillating course deviation alert because they were forced to integrate the audio feedback continually and, as a result, were more actively visually monitoring UAV progression on or off course.

The results for the late arrival reaction do not show the same trends. Whether participants had a continuous modulated alert for late arrival events across both the single and multiple UAV conditions did not make a difference in late arrival reaction times. However, those participants with the continuous course deviation audio condition and the threshold alert for late arrivals performed significantly worse *in terms of late arrival reaction time* across both single or multiple UAVs scenarios. This is notable because when presented with threshold alerts for both the course deviation and late arrival events, participants responded more quickly to late arrival alerts than when the continuous course deviation alert was used.

This conflicting result could be indicative of an aural saliency problem. It appears the oscillating course deviation audio alert may saturate the audio channel to a point that the late arrival's threshold beep alerts are not heard, resulting in longer late arrival reaction times. Interestingly, the combination audio condition with continuous alerts for both the late arrival and course deviation conditions produced statistically better reaction times than when just the course deviation continuous alert was used. The integration scheme of the two continuous alerts could explain this result. The continuous modulated late arrival alert continues to sound until the participant addresses it; the participant cannot easily miss this alert, even in the presence of the oscillating course deviation alert. In contrast, the late arrival threshold alert sounds once and is easily missed in the presence of the continuous course deviation alert. This was noted by many participants in the post experiment survey, where 16 participants asked for alerts with "more than just a simple beep" but instead wished for a more repetitive alert that will "last" so they will have to address it.

In terms of error rates in responding to both course deviation and late arrival events, only a very small percentage of course deviation alerts were missed, and all of the late arrivals were responded to by the participants. Both events were also represented visually which demonstrates the value of coding on multiple channels. One difference in the visual alert representation that could cause this slight difference in error rates is that the late arrival visual alert is easily seen by the physical displacement and color change of the target icon. In contrast, whether a UAV is truly off course and not correcting is a judgment call until the UAV is clearly off the path. Appendix A contains images of the late arrival and course deviation representations. In the simulation, if a participant did not recognize and correct an off course error in 30 seconds, the UAV automatically resumed its correct course.

In addition to the audio condition independent variable, the impact of workload was examined. Objectively, the secondary task results show that workload appears to be the same between the four audio conditions and the two scenarios. This workload equality is not a surprise, because as

discussed earlier and as shown in Appendix B, the two scenarios were designed to have comparable workloads, with the multi UAV scenario dividing the monitoring tasks over four vehicles instead of just one. The subjective ratings, the NASA TLX data, confirm that the participants felt the two scenarios were comparable in workload.

However, while participants observed that audio alerts were "very helpful" when focusing on other tasks, it was "harder to comprehend with four UAVs than one." The results generally support these comments. Across both the course deviation and late arrival reaction times, participants reacted more quickly to the single UAV scenario than the multiple UAV case. This result is not surprising since this trend is seen throughout air traffic controller studies, and is a known effect of increased cognitive complexity (Cummings & Tsonis, 2005). As the number of aircraft that require supervisory control increases (akin to the number of UAVs under control), reaction to the various tasks degrades because the user's attention is divided across multiple entities.

Comments from the participants were generally positive about having auditory cues complement the visual display. They reported that the audio "helped draw my attention to the matter so I could make a decision." However, not all subjective comments were positive. Many participants, including actual Air Force pilots, reported that they do not want to listen to continuous "noise for hours at a time day after day." A few participants commented that the perceived noise made them feel "overwhelmed" and was "fatiguing." A few suggested "dampening or eliminating" the "constant background noise when" the UAVs operate within normal parameters. It is important to note that these scenarios were 30 minute test scenarios, and actual UAV operations occur over much longer periods. Integrating this audio alert into operations must be balanced against what operators will subjectively tolerate. It is possible that continuous audio should only be used over limited time spans, during high risk and high workload periods of the mission. These are areas of possible future research.

CONCLUSION

This research demonstrates that sonifications can help single operators supervise multiple semi-autonomous vehicles. However, auditory cues must be integrated in a fashion that the operators will be able to tolerate subjectively, and these signals must also be integrated with other audio cues to ensure balanced aural saliency. This research also showed that while workload is an important factor in overall performance, the sonifications were equally effective for participants controlling one or four UAVs.

Pontillo and Brian Malley, undergraduate assistants, who contributed by helping to set up the experiment, run participants, and process data.

REFERENCES

Barrass, S., & Kramer, G. (1999). Using sonification. *Multimedia Systems, 7*, 23-31.

Cooper, J. C., & Owen, J. H. (1976). Audiologic profile of noise-induced hearing loss. *Archives of Otolaryngology, 102*(3), 148-150.

Culbertson, E. (2006). COMUSAFE: unmanned aircraft key to future decision superiority. from http://www.af.mil/news/story.asp?storyID=123029520

Cummings, M. L., & Tsonis, C. G. (2005). *Deconstructing Complexity in Air Traffic Control.* Paper presented at the HFES Annual Conference, Orlando, FL.

Hart, S., & Staveland, L. (1988). Development of the NASA-TLX: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (1 ed., pp. 139-183). Amsterdam: North Holland.

Humes, L. E., Joellenbeck, L. M., & Durch, J. S. (Eds.). (2006). *Noise and Military Service Implications for Hearing Loss and Tinnitus*. Washington, D.C.: Institute of Medicine of the National Academies.

Nehme, C. E., & Cummings, M. L. (2006). *Audio Decision Support for Supervisory Control of Unmanned Vehicles* (No. HAL2006-06). Cambridge, MA: Humans and Automation Laboratory.

Snodgrass, J. G. (1975). Psychophysics. In B. Scharf (Ed.), *Experimental Sensory Psychology*.

Walden, B. E., Prosek, R. A., & Worthington, D. W. (1975). *The Prevalence of Hearing Loss Within Selected U.S. Army Branches*. Washington, D.C.: Walter Reed Army Medical Center.

Watson, M., & Sanderson, P. (2004). Sonification Supports Eyes-Free Respiratory Monitoring and Task Time-Sharing. *Human Factors, 46*(3), 497-517.

Whitney, R. (2007). Air Force stands up first unmanned aircraft systems wing. from http://www.spacewar.com/reports/Air_Force_Stands_Up_First_Unmanned_Aircraft_Systems_Wing_999.html

Wundt, W. M. (1902). *Principles of Physiological Psychology*. London: Swan Sonnenschein & Co. Lim.

APPENDIX A COURSE DEVIATION AND LATE ARRIVAL IMAGES

COURSE DEVIATION

The green line in the images below indicates the desired flight path. Anytime the UAV has completely departed from the green course line, it is considered to have deviated. During the simulation, the UAV constantly fluctuates back and forth to a small degree on the flight path, but participants are trained to respond when the UAV actually is no longer touching the course line. Below are images illustrating UAV 3 on course, almost off course, and off course.

On Course	**Almost Off Course**	**Off Course**

LATE ARRIVAL

The image below shows the visual representation for a late arrival. A black rectangle on the timeline shows when the UAV is going to arrive at the labeled target, and changes from black to red when the UAV is late. This can also be seen when the black rectangle moves past the orange window of opportunity. On the timeline below, the red rectangle labeled "T-1L" indicates that the UAV is going to be late to target T-1L. In contrast, the black rectangle labeled "T-2M" indicates that the UAV is going to be on time to target T-2M.
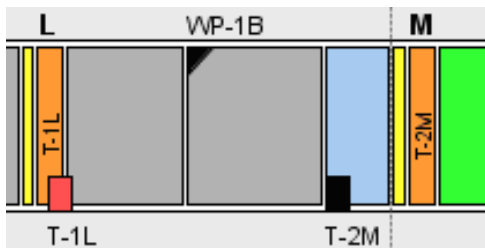
APPENDIX B SCENARIO EVENTS

These slides show the major events in the single UAV and multiple UAV 30 minute test scenarios.
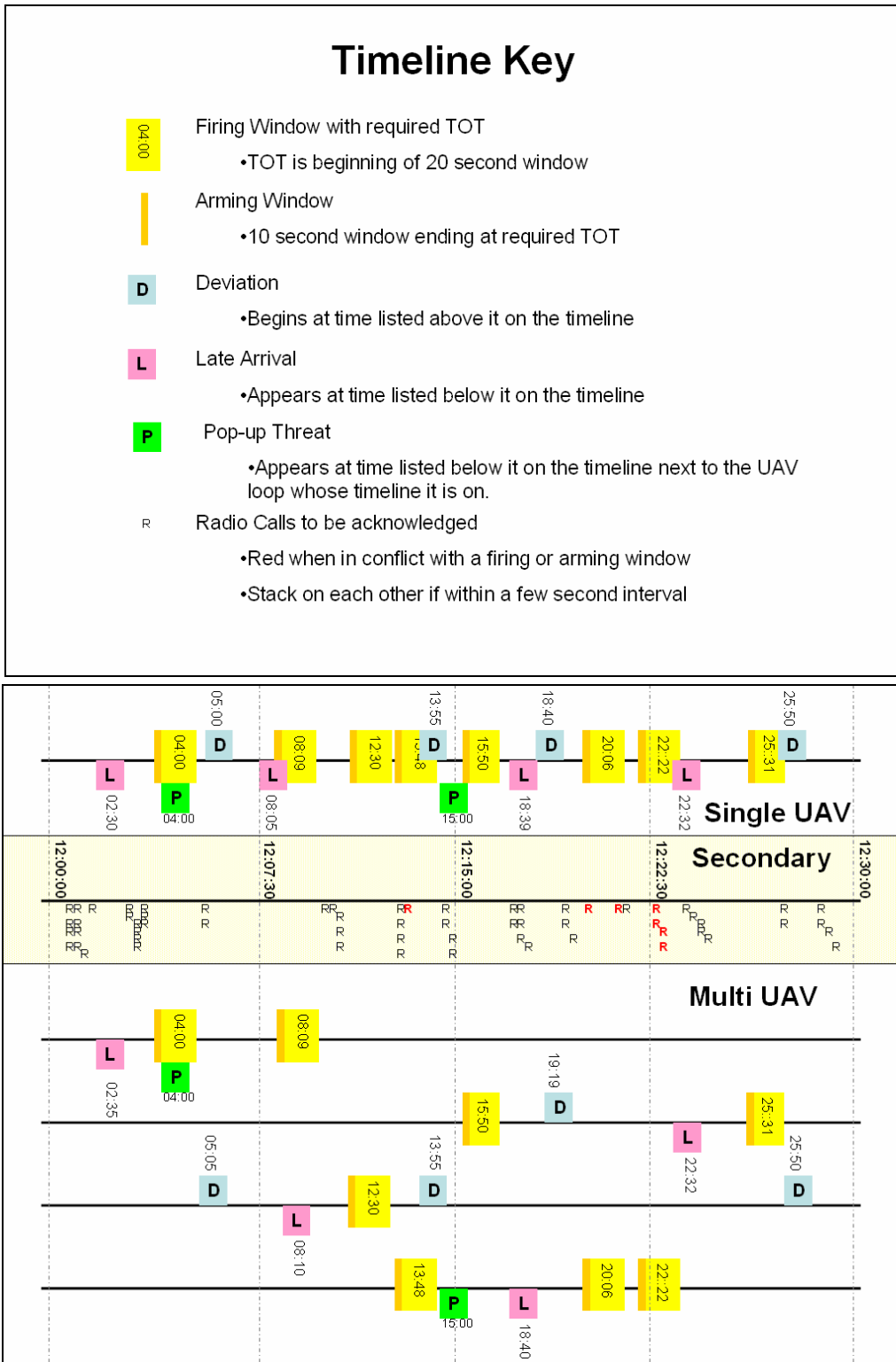
The table below shows side by side images of the timeline display and map display for the single UAV and multi UAV scenarios. The displays show that in both scenarios the operator has the same number of tasks to complete. The only difference between the two scenarios is that for the multi scenario the tasks are divided over four UAVs instead of just one UAV.

	Single UAV Scenario	Multi UAV Scenario
Timeline Display		
Map Display		
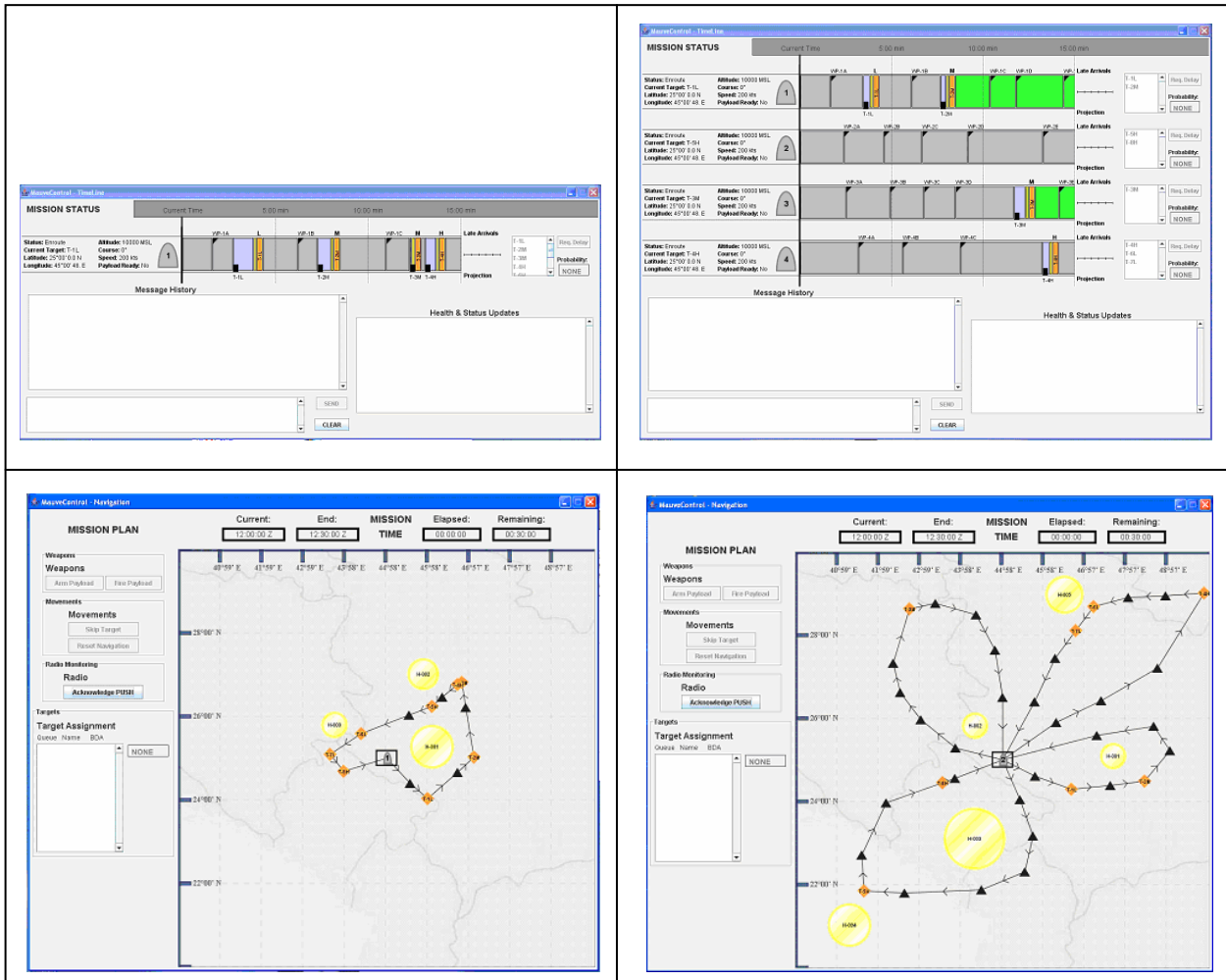
APPENDIX C PARTICIPANT CONSENT FORM
CONSENT TO PARTICIPATE IN
NON-BIOMEDICAL RESEARCH

Developing Decision Support for Supervisory Control of Multiple Unmanned Vehicles

You are asked to participate in a research study conducted by Professor Mary Cummings Ph.D, from the Aeronautics and Astronautics Department at the Massachusetts Institute of Technology (M.I.T.). You were selected as a possible participant in this study because the expected population this research will influence is expected to contain men and women between the ages of 18 and 50 with an interest in using computers. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

- **PARTICIPATION AND WITHDRAWAL**

Your participation in this study is completely voluntary and you are free to choose whether to be in it or not. If you choose to be in this study, you may subsequently withdraw from it at any time without penalty or consequences of any kind. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

- **PURPOSE OF THE STUDY**

The study is designed to evaluate how decision support tools or recommendations, both audio and visual, assist an operator supervising multiple simultaneous dynamic tasks, and how decision support assistance and effectiveness changes as workload increases. In measuring the effectiveness of decision support, an operator's performance and situation awareness are used as metrics. Situation awareness is generally defined as the perception of the elements in the environment, the comprehension of the current situation, and the projection of future status of the related system.

- **PROCEDURES**

If you volunteer to participate in this study, we would ask you to do the following things:

- Attend a training and practice session to learn a video game-like software program that will have you supervising and interacting with multiple unmanned aerial vehicles (estimated time 0.75 hours).
- Practice on the program will be performed until an adequate level of performance is achieved, which will be determined by your demonstrating basic proficiency in monitoring the vehicles, redirecting them as necessary, executing commands such as firing and arming of payload at appropriate times, using decision support visualizations and/or recommendations to mitigate timeline problems, and responding to radio calls by clicking an acknowledge button on the software interface (estimated time 0.75 hours).
- Execute two thirty minute trials consisting of the same tasks as above (1 hour)
- Attend a debriefing to determine your subjective responses and opinion of the software (10 minutes).
- Testing will take place in MIT building 37, room 301.

- Total time: 2-3 hours, depending on skill level.

- **POTENTIAL RISKS AND DISCOMFORTS**

There are no anticipated physical or psychological risks in this study.

- **POTENTIAL BENEFITS**

While there is no immediate foreseeable benefit to you as a participant in this study, your efforts will provide critical insight into the human cognitive capabilities and limitations for people who are expected to supervise multiple complex tasks at once, and how decision support tools can support their task management.

- **PAYMENT FOR PARTICIPATION**

You will be paid $10/hr to participate in this study which will be paid upon completion of your debrief. Should you elect to withdraw in the middle of the study, you will be compensated for the hours you spent in the study.

- **CONFIDENTIALITY**

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. You will be assigned a subject number which will be used on all related documents to include databases, summaries of results, etc. Only one master list of subject names and numbers will exist that will remain only in the custody of Professor Cummings.

- **IDENTIFICATION OF INVESTIGATORS**

If you have any questions or concerns about the research, please feel free to contact the Principal Investigator, Mary L. Cummings, at (617) 252-1512, e-mail, missyc@mit.edu, and her address is 77 Massachusetts Avenue, Room 33-305, Cambridge, MA 02139. The student investigators are Hudson D. Graham (719-238-1713, email: hgraham@mit.edu), and Amy Brzezinski (617-276-6708, amybrz@MIT.EDU).

- **EMERGENCY CARE AND COMPENSATION FOR INJURY**

"In the unlikely event of physical injury resulting from participation in this research you may receive medical treatment from the M.I.T. Medical Department, including emergency treatment and follow-up care as needed. Your insurance carrier may be billed for the cost of such treatment. M.I.T. does not provide any other form of compensation for injury. Moreover, in either providing or making such medical care available it does not imply the injury is the fault of the investigator. Further information may be obtained by calling the MIT Insurance and Legal Affairs Office at 1-617-253-2822."

- **RIGHTS OF RESEARCH SUBJECTS**

You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you feel you have been treated unfairly, or you have questions regarding your rights as a research subject, you may contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T., Room E25-143b, 77 Massachusetts Ave, Cambridge, MA 02139, phone 1-617-253-6787.

SIGNATURE OF RESEARCH SUBJECT OR LEGAL REPRESENTATIVE

I understand the procedures described above. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

_____
Name of Subject

_____   _____
Signature of Subject   Date

SIGNATURE OF INVESTIGATOR

In my judgment the subject is voluntarily and knowingly giving informed consent and possesses the legal capacity to give informed consent to participate in this research study.

_____   _____
Signature of Investigator   Date

APPENDIX D DEMOGRAPHICS SURVEY
MAUVE-MITUS Demographic Survey

1. **Age:** _____

2. **Gender:** □ Male   □ Female

3. **Occupation:** _____

   If student:
   a. *Class Standing:* □ Undergraduate □ Graduate
   b. *Major:* _____

   If currently or formerly part of any country's armed forces:
   a. *Country/State:* _____
   b. *Status:* □ Active Duty □ Reserve □ Retired
   c. *Service:* □ Army □ Navy □ Air Force □ Other _____
   d. *Rank:* _____
   e. *Years of Service:* _____
   f. *Did you ever serve in high noise environments?* □ Yes □ No
      If yes*, please explain what the duties were, how long the shifts were, and how many times you served these shifts?*
      _____
      _____

4. **Do you have experience with remotely piloted vehicles (land, sea, air)?**
   □ Yes
   □ No

   If yes:
   a. *Vehicle type(s)/class(es):* _____
   b. *Number of hours:* _____

5. **Do you have experience with radios such as those used for communication in flying?**
   □ Yes
   □ No

   If yes:
   Please explain: _____
   _____

6. **Have you been to a music concert in the last Month?**
   □ Yes
   □ No

   If yes:
   a. *Concert type:*_____
   b. *When:* _____
7. **Do you have any hearing loss?**
   □ Yes
   □ No

If yes:
   Please explain: _____
   _____

8. **How often do you play video games?**
   □ Never
   □ Less than 1 hour per week
   □ Between 1 and 4 hours per week
   □ Between 1 and 2 hours per day
   □ More than 2 hours per day

9. **Are you color blind?**
   □ Yes
   □ No

   If yes:

   Which type of color blindness (if known)_____

Spring 2007

Multi-Aerial Unmanned Vehicle Experiment (MAUVE) TUTORIAL

Hudson Graham – MIT Humans and Automation Lab

Introduction

Welcome!

This tutorial is designed to give you some background on the Multi-Aerial Unmanned Vehicle Experiment (MAUVE) interface before you arrive on testing day. Please take the time to look over the following slides and come prepared with questions. Before testing you will be thoroughly trained on the actual interface, but being exposed to it beforehand will be invaluable in speeding up this process.

Thank you in advance for your participation!

Experiment Overview

In this experiment, you are an unmanned aerial vehicle (UAV) operator that is responsible for supervising 1 to 4 UAVs collectively tasked with destroying a set of time-sensitive targets in a suppression of enemy air defenses mission. The area contains enemy threats capable of firing on your UAVs.

The UAVs are highly autonomous, and therefore only require high level mission execution from you. The UAVs launch with a pre-determined mission plan, so initial target assignments and routes have already been completed for you. Your job will be to monitor their progress, re-plan aspects of the mission in reaction to unexpected events, and in some cases manually execute mission critical actions such as arming and firing of payloads.

The interface we have developed for this experiment is called the Multi-Aerial Unmanned Vehicle Experiment (MAUVE) and will be referred to by this name from here out.

Objectives

Your primary objective in this mission is:

To make sure the UAV(s) maximize the number of targets engaged as well as arriving back to the base safely.

Supervision of the UAVs can be broken down into the following prioritized sub-tasks, from highest priority to lowest:

1. Return to base (RTB) within the time limit for the mission (this limit will be clearly marked).
2. Comply with recovery rules for course deviations.
3. Comply with recovery rules for target late arrivals.
4. Destroy all targets before their time on target (TOT) window ends.
5. Avoid taking damage from enemies by avoiding all threat areas.
6. Acknowledge all "Push" radio calls.

These sets of objectives will often conflict with one another. In these cases, you must perform the actions that have the highest priority first.

Your performance will be judged based on how well you follow the above priorities.

Audio Alerts

To help you meet your objectives you will receive auditory signals for both course deviations and late arrivals. Both are induced by unanticipated high winds along the planned flight path.

Course deviations are when a UAV is blown off of the planned path. It is significantly deviated when you visually see that the UAV has left the course line. Deviations may occur over targets as well.

Late arrivals are when the UAV has hit stronger than anticipated head winds and slows down. As a result it will now be late to the next target.

Your test proctor will provide further training as to what these auditory signals sound like during the test day training.

Other auditory sounds to be familiar with are: (Note all three are the same because all are related to a new message in your message box.)

For new messages in your message boxes

For pop-up threats

For when your UAV is being fired upon while flying through a threat area.

Color Coding

Throughout the displays you're about to see, the following color coding is used to indicate each of the 5 possible actions a UAV can perform in MAUVE:

UAV Action	Color
Enroute	Gray
Loitering	Blue
Arming Payload	Yellow
Firing Payload	Orange
Return to Base	Green

Displays – Overview

During the experiment, you will see two side-by-side displays that contain the following major elements:

- **Left Display**
  - **Mission Time**
  - **Map Display**
  - **Mission Execution**

- **Right Display**
  - **Unmanned Aerial Vehicle (UAV) Status**
  - **Decision Support**
  - **Chat Box**
  - **UAV Health & Status Updates**

The following slides will show these displays in detail and explain how to use them properly.

Left Display – Overview

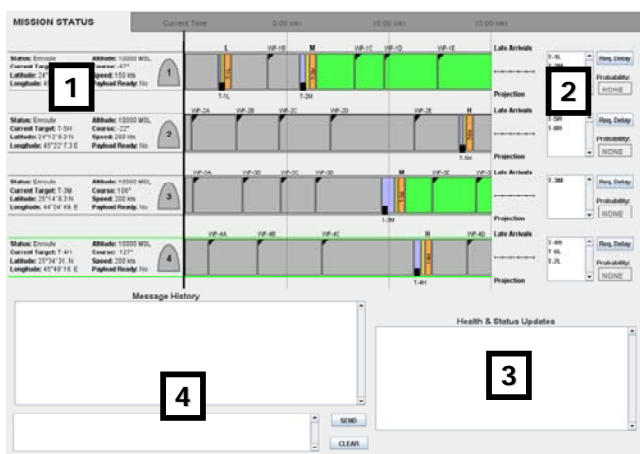The three major screen elements on the left display are:

1
Mission Execution

2
Mission Time

3
Map Display
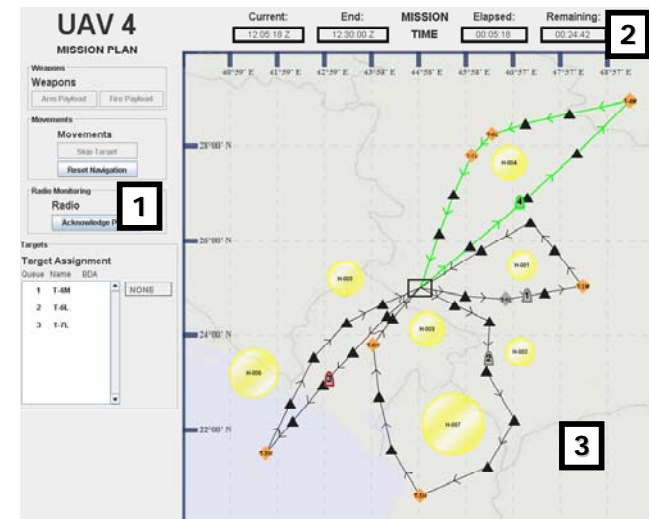
Right Display – Overview

The four major screen elements on the right display are:

1
UAV Status

2
Decision Support

3
UAV Health & Status Updates

4
Chat Box

Left Display – Detail

The following slides detail all of the elements contained on the left display, in this order:

- **Map Display**
- **Mission Execution**
- **Mission Time**

Map Display – Detail – 1

Key Map Display Elements

1. Active Target
2. Waypoint
3. Base
4. UAV
5. Threat Area

Mission Plans

- The solid black lines indicate each UAV's current mission plan
- The currently selected mission plan is highlighted green
- One UAV will always be highlighted with a red border. This UAV has the greatest course deviation. It corresponds to the course deviation auditory signal you will be receiving.
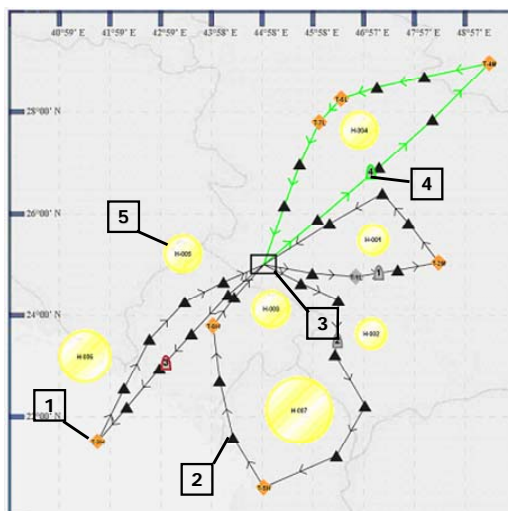
Map Display – Detail – 2

Naming Conventions

- *UAVs*
  - Numbered 1-4
- *Targets*
  - T-XXP where XX = target number and P = priority
  - Priority may be High (H), Medium (M), or Low (L)
  - Examples:
    - T-1H – Target 1 a high priority target
    - T-12M – Target 12 a medium priority target
    - T-23L – Target 23 a low priority target
- *Waypoints (WP)*
  - WP-XY where X = UAV# the waypoint is associated with and Y = waypoint letter
  - Examples: WP-1A, WP-2C
- *Threats/Hazards*
  - H-XXX where XXX = threat number
  - Example: H-001, H-012

Mission Time – Detail

The mission time display element shows the following:

1. Absolute Time
   - These clocks are on the left side of the "Mission Time" title
   - Two formats
     - Current Time: Current mission time
     - End Time: End time of the current mission
2. Relative Time
   - These clocks are on the right side of the "Mission Time" title
   - Two formats
     - Elapsed Time: How long has elapsed since the start of the scenario
     - Remaining Time: How long remains until the mission is over

Example Mission Time Display Element

Current:	End:	**MISSION TIME**	Elapsed:	Remaining:
12:02:43 Z	12:30:00 Z		00:02:43	00:27:17

Mission Execution – Detail – 1

Each UAV has its own mission execution bar.

To bring it up on the left display, click anywhere on the desired UAV's status window on the right display OR on the UAV icon itself on the left display.

Light green highlighting around the UAV's status bar and its current mission plan on the map display tell you which UAV/route is currently selected

In the display below UAV 4 is highlighted so this is the mission execution bar on the left side of the left display.
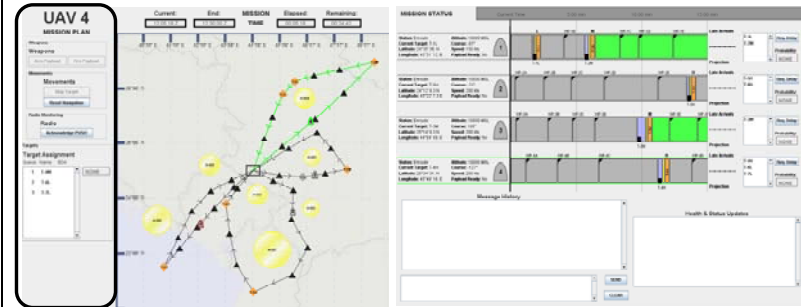
Mission Execution – Detail – 2

Mission Execution Functions

1. Arm Payload
   - This button is only enabled if the UAV is selected while directly on top of a target, and within the arming or firing windows
2. Fire Payload
   - This button is only enabled if the UAV is selected while directly on top of a target, armed, and within the firing window for that particular target
3. Skip Target
   - This button is used if you decide to skip a target because you are going to be late to it. It causes the UAV to skip the next target/waypoint and move to the next waypoint/target

---
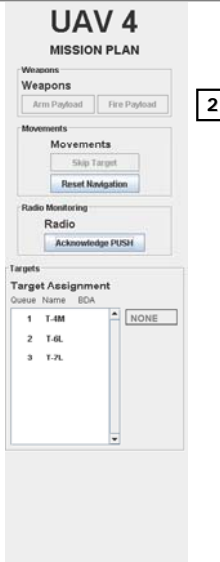
Mission Execution – Detail – 3

Mission Execution Functions

4. Reset Navigation
   - Click anytime you have a significant course deviation and want to return to the planned course; it will return the UAV to the next plotted waypoint/target.
   - Causes the UAV to be inactive while resetting; you may not be able to arm or fire with that UAV when its navigation system is resetting.
5. Radio Monitoring
   - Radio chat from the Boston ground control will be playing and you must click the "Acknowledge PUSH" each time "push" is called on the radio chat. *Hint: Push calls usually come in pairs (but not always); the tower and an aircrew or vice versa.*
6. Target Assignment Queue
   - You will not be using this portion of the display.

31
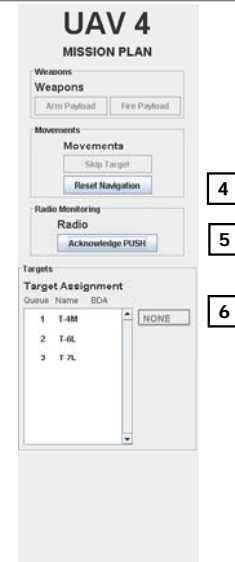
Right Display – Detail

The following slides detail all of the elements contained on the right display, in this order:

- **UAV Status**
- **Health & Status Updates**
- **Decision Support**
- **Chat Box**

A Reminder of How it All Fits Together – Right Display

The four major screen elements on the right display are:

1
UAV Status

2
Decision Support

3
UAV Health & Status Updates

4
Chat Box
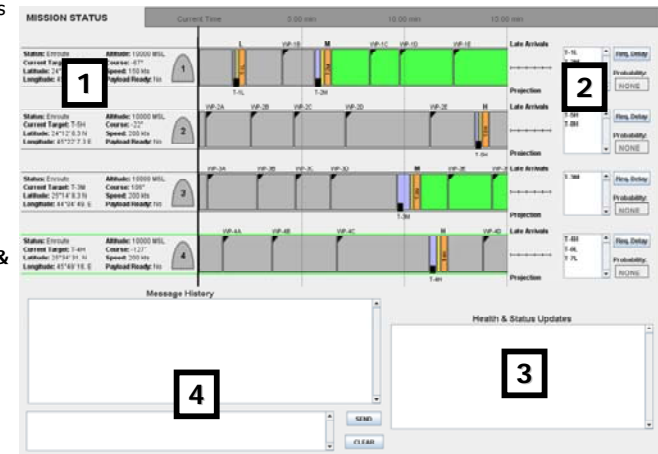
UAV Status – Detail – 1

The UAV status display shows the following real-time information for each UAV:

- **Status / Current Action**
  - This is written out as well as represented by the color of the UAV icon to the right
  - For example: a blue UAV would be loitering. This means it has arrived over the target but is there before the arming and firing window so is in a holding pattern waiting for these windows.
- **Current Target Name**
- **Position in Latitude & Longitude**
  - You will not need this in the scenario.
  - Given in degrees, minutes, and seconds
- **Altitude**
  - This is a static number that is not used in the scenario in any way

(continued on next slide)

Example UAV Status Display Element

UAV Status – Detail – 2

The UAV status display shows the following real-time information for each UAV:

- **Course**
  - You will not need this during the scenario.
  - 0˚ indicates due north; increases in a clockwise manner
- **Speed**
  - The UAVs are set to travel at a constant speed of 200 kts.
  - If there is significant head wind the UAV may slow down. This can precipitate late target arrivals
- **Payload Ready**
  - This reflects whether the UAV has a payload ready for the current target
  - Will say "Yes" if the UAV is armed, "No" if not

Example UAV Status Display Element

UAV Health & Status Updates – Detail

The Health & Status Updates box contains messages from specific UAVs intended to inform the operator. Messages are color coded as follows:

- **Red** = UAV Health messages
  - UAV is under fire from a threat
  - Again, a standard audio alert will play when you receive red messages.
- Bold Black = UAV Status messages, action required
  - UAV is available to arm or fire
- Black = UAV Status messages, no action required
  - UAV has completed arming or firing

Example Health & Status Updates Window
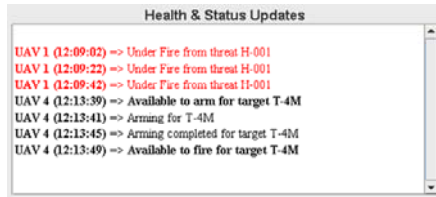
Decision Support – Remember the Color Coding

Color coding is an important element of the decision support, so take a look at it again!

UAV Action	Color
Enroute	Gray
Loitering	Blue
Arming Payload	Yellow
Firing Payload	Orange
Return to Base	Green

Decision Support – Detail – 1

The active level of decision support contains a visual representation of what and when targets are approaching through a relative timeline and projective decision support display for each UAV.

- The arming and firing windows cannot by changed solely at the will of the the operator. i.e. Operators may request time on target (TOT) delay requests, but must get approval before the arming and firing window will be moved back (if approved).

Example Active Decision Support Window
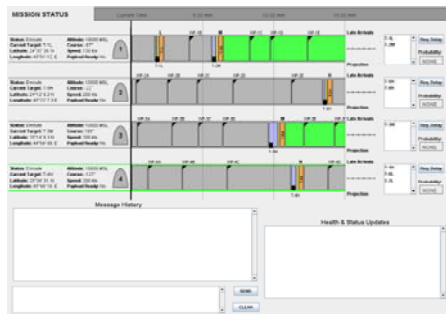
Decision Support – Detail – 2

Arming and firing elements are color coded in the same way as corresponding UAV actions. For each target the following information is represented:

1. Arming Window = Yellow
   - 10 seconds long and takes approximately 3-7 seconds to arm.
   - Payload for the relevant target may be armed, but not fired during this time
   - Always occurs immediately before the firing window
2. TOT/Firing Window = Orange
   - 20 seconds long and takes approximately 3-7 seconds to fire.
   - Payload must be fired at the relevant target during this time
   - In addition to the arming window, a payload may also be armed during this time
   - Target name is printed vertically in the center of the window, priority is printed above the window.

Decision Support – Detail – 3

Mission planning information reflects when UAVs will reach important points in the scenario, such as:

1. **Waypoints/Loiterpoints/Base = Black Triangles**
   - **Names are printed above the relevant UAV's timeline**
2. **UAV Arrival at Targets = Black Rectangles**
   - **Names are printed below the relevant UAV's timeline**
   - **Note that each target name will appear twice on the timeline, once for when the UAV will arrive at that target and once at the center of that target's firing window**
3. **Late UAV Arrival = Red Rectangles**
   - **Black Rectangle turns red and moves past the target to when the UAV will arrive.**
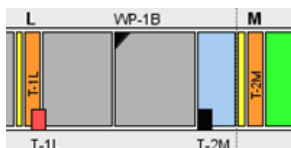
Decision Support – Detail – 4

The active level of decision support aids the user by identifying possible late target arrivals and areas of potential high workload on the timeline.

- **A late target arrival is defined as when a UAV will arrive to a target after its scheduled time on target**
- **Corrective actions for late arrivals are based on the priority of the target you are projected to be late to. For:**
  - **Low priority targets, skip them by clicking "Skip Target" on the navigation screen.**
  - **Medium priority targets, either skip them or use the decision support visualization (DSV) to possibly request a delay. Remember your priorities of wanting to hit all the targets. *See the next four slides for an explanation of how to use the DSV (shown below).***
  - **High priority targets, use the DSV and then decide whether to request a TOT delay or to skip the target by clicking "Skip Target"**

Note: The corrective actions above should only be taken when a late arrival is projected by the red rectangle on the screen and your audio.
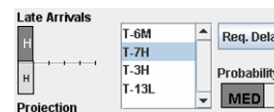
Decision Support – Detail – 5

The decision support visualization (DSV) helps the user manage the schedule by showing timeline issues and projecting "what if" conditions of the effects on the timeline based upon user decisions.

- **Each UAV's DSV is uses emergent features to show problems that currently exist or that may exist if a TOT delay is given.**
- **No issues (late target arrivals) are indicated by no rectangles being displayed.**
  - **The picture below shows the DSV for when there is no late arrivals. It depicts this by having no rectangles above the line on the left side of the display in the "Late Arrival" section.**

Note: The DSV will be inactive except for when you are going to be late to a medium or high priority target.

Decision Support – Detail – 6

- **Late arrivals are represented by a rectangle occurring on the DSV above the line on the left side of the display in the "Late Arrivals" section. It will also be highlighted yellow as it is below.**
  - **The target's priority is indicated within the rectangle and by the rectangle's size. The higher the priority of the target that the UAV will be late for the taller the rectangle will be.**

34
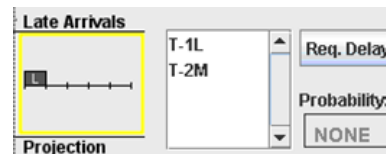
Decision Support – Detail – 7

- Below the center line is for the "what if" condition; after the user selects a target that they might request a TOT delay for it will show the projected future late arrivals for that UAV below the centerline on the left side of the display in the "Projection" section.

  - The example below shows that if a TOT delay request is granted for target T-16H, the UAV will then be late to a Low Priority target.

  - The Probability in the bottom right of this display shows the likelihood of a TOT delay being granted. The further in advance a delay is requested the higher the likelihood of it being granted. Do not request a delay again if your first request for a delay on that target is denied.
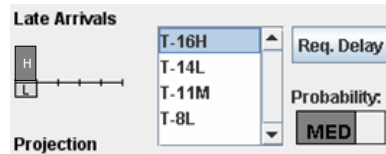
Decision Support – Detail – 8

Each UAV possesses a DSV display to help the user understand the potential effects of decisions

- A list of all the mission targets on the timeline appears to the right of each UAV's DSV display.

- In using the display to the right for the top UAV:
  - The user is considering requesting a delay for T-7H, a high priority target. However, it shows they will now be late to another high priority target even with this delay. This is where you, as the user, will have to make a value call. Do you request a delay or skip the target by clicking "Skip Target?" Most of the time you will not want to request a delay if you know it is going to create another delay.

- In using the display to the right for the second UAV:
  - The user is considering requesting a delay for T-16H, a high priority target. This will result in the UAV being late to a low priority target. So in this case you, as the user, would want to request this delay because it means you can hit the high priority target for the trade-off of now missing a low priority target.
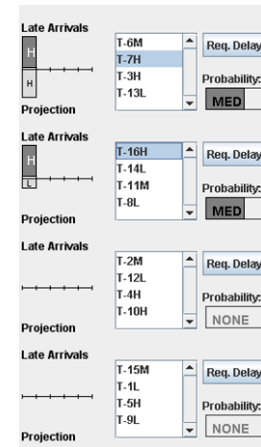
Chat Box – Detail

The chat box contains a time history of all *human* interactions. These interactions are color coded as follows:

- **Red** = Intelligence updates
  - Again, a standard audio alert will play when you receive red messages.
- **Black** = Message to/from Base, no response required
  - Messages that inform you, but do not require a specific response

The chat box is purely informational. It will provide you with updates, but you will not input anything in the chat box.

Example Message History Window
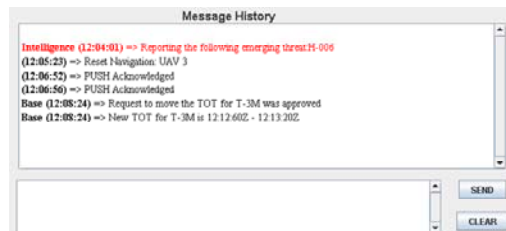
Conclusion

You are now ready to proceed to hands-on training with the MAUVE interface. Remember to bring any questions you have to the experimenter on testing day!

Predator B

Global Hawk

APPENDIX F POST-EXPERIMENT SURVEY

1. How did the audio cues help or hinder you in managing *late arrivals*?

_____

_____

_____

_____

_____

2. How did the audio cues help or hinder you in managing *course deviations*?

_____

_____

_____

_____

_____

3. How would you change the audio cues for the *late arrivals* to have them better aid you during your mission?

_____

_____

_____

_____

_____

4. How would you change the audio cues for the *course deviations* to have them better aid you during your mission?

_____

_____

_____

_____

_____

5. In general how would you change all the audio to help you with the control during your mission?

_____

_____

_____

_____

_____

_____

6. Please express any other comments you may like to share:

_____

_____

_____

_____

_____

APPENDIX G GLM ANALYSIS: SPSS OUTPUT

Course Deviation Reaction Time (for 4 audio conditions)

Tests of Normality

	Audio_Scheme	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	Df	Sig.
D_Sing	Combination	.175	10	.200(*)	.949	10	.653
	Continuous Late Arrival	.192	10	.200(*)	.912	10	.296
	Continuous Course Deviation	.180	10	.200(*)	.893	10	.181
	Threshold	.113	9	.200(*)	.978	9	.952
D_Multi	Combination	.193	10	.200(*)	.927	10	.415
	Continuous Late Arrival	.192	10	.200(*)	.918	10	.339
	Continuous Course Deviation	.267	10	.042	.843	10	.048
	Threshold	.165	9	.200(*)	.964	9	.836

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

Levene's Test of Equality of Error Variances(a)

	F	df1	df2	Sig.
D_Sing	.721	3	35	.546
D_Multi	.310	3	35	.818

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Audio_Scheme Within Subjects Design: scenario

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	scenario	Type III Sum of Squares	df	Mean Square	F	Sig.
scenario	Linear	14.029	1	14.029	3.215	.082
scenario * Audio_Scheme	Linear	10.780	3	3.593	.824	.490
Error(scenario)	Linear	152.711	35	4.363		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	8998.235	1	8998.235	1089.907	.000
Audio_Scheme	71.289	3	23.763	2.878	.050
Error	288.959	35	8.256		

Tukey Comparison

Multiple Comparisons

Measure: MEASURE_1
Tukey HSD

(I) Audio_Scheme	(J) Audio_Scheme	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Combination	Continuous Late Arrival	-1.6292	.90862	.294	-4.0796	.8213
	Continuous Course Deviation	.3667	.90862	.977	-2.0838	2.8171
	Threshold	-1.7949	.93352	.237	-4.3125	.7227
Continuous Late Arrival	Combination	1.6292	.90862	.294	-.8213	4.0796
	Continuous Course Deviation	1.9958	.90862	.144	-.4546	4.4463
	Threshold	-.1657	.93352	.998	-2.6834	2.3519
Continuous Course Deviation	Combination	-.3667	.90862	.977	-2.8171	2.0838
	Continuous Late Arrival	-1.9958	.90862	.144	-4.4463	.4546
	Threshold	-2.1616	.93352	.114	-4.6792	.3560
Threshold	Combination	1.7949	.93352	.237	-.7227	4.3125
	Continuous Late Arrival	.1657	.93352	.998	-2.3519	2.6834
	Continuous Course Deviation	2.1616	.93352	.114	-.3560	4.6792

Based on observed means.

			D	Min	Max	q*	P
Sing	D1	Comb-Mod	-2.02	-4.48	0.45	-3.55	
	D2	Comb-Osc	-0.46	-2.92	2.00	-0.81	
	D3	Comb-Thresh	-2.73	-5.20	-0.27	-4.81	0.02
	D4	Mod-Osc	1.56	-0.90	4.02	2.74	
	D5	Mod-Thresh	-0.72	-3.18	1.75	-1.26	
	D6	Osc-Thresh	-2.28	-4.74	0.19	-4.00	
Multi	D7	Comb-Mod	-1.24	-3.70	1.22	-2.18	
	D8	Comb-Osc	1.19	-1.27	3.65	2.09	
	D9	Comb-Thresh	-0.86	-3.32	1.61	-1.50	
	D10	Mod-Osc	2.43	-0.03	4.90	4.28	
	D11	Mod-Thresh	0.39	-2.08	2.85	0.68	
	D12	Osc-Thresh	-2.05	-4.51	0.42	-3.60	

Course Deviation Reaction Time (for 2 audio alerts)

Tests of Normality

	Audio_Scheme _Osc	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
D_Sing	Oscillating Alert	.145	20	.200(*)	.933	20	.179
	Beep Alert	.130	19	.200(*)	.971	19	.799
D_Multi	Oscillating Alert	.180	20	.089	.940	20	.237
	Beep Alert	.107	19	.200(*)	.967	19	.718

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

Levene's Test of Equality of Error Variances(a)

	F	df1	df2	Sig.
D_Sing	.328	1	37	.570
D_Multi	.167	1	37	.685

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Audio_Scheme_Osc Within Subjects Design: scenario

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	scenario	Type III Sum of Squares	Df	Mean Square	F	Sig.
Scenario	Linear	14.534	1	14.534	3.311	.077
scenario * Audio_Scheme _Osc	Linear	1.089	1	1.089	.248	.621
Error(scenario)	Linear	162.403	37	4.389		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Intercept	9007.397	1	9007.397	1146.991	.000
Audio_Scheme_Osc	69.684	1	69.684	8.874	.005
Error	290.564	37	7.853		

Transformed (Natural Log) Late Arrival Reaction Time

Tests of Normality

	Audio_Scheme	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
LN_L_Sing	Combination	.162	10	.200(*)	.958	10	.768
	Continuous Late Arrival	.205	10	.200(*)	.943	10	.583
	Continuous Course Deviation	.246	10	.088	.839	10	.043
	Threshold	.180	9	.200(*)	.915	9	.349
LN_L_Multi	Combination	.156	10	.200(*)	.962	10	.812
	Continuous Late Arrival	.215	10	.200(*)	.903	10	.234
	Continuous Course Deviation	.242	10	.098	.915	10	.319
	Threshold	.170	9	.200(*)	.977	9	.944

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

Levene's Test of Equality of Error Variances(a)

	F	df1	Df2	Sig.
LN_L_Sing	.651	3	35	.587
LN_L_Multi	1.583	3	35	.211

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Audio_Scheme Within Subjects Design: scenario

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	scenario	Type III Sum of Squares	Df	Mean Square	F	Sig.
Scenario	Linear	6.513	1	6.513	20.737	.000
scenario * Audio_Scheme	Linear	1.619	3	.540	1.719	.181
Error(scenario)	Linear	10.992	35	.314		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	244.837	1	244.837	305.851	.000
Audio_Scheme	8.028	3	2.676	3.343	.030
Error	28.018	35	.801		

Tukey Comparison

Multiple Comparisons

Measure: MEASURE_1
Tukey HSD

(I) Audio_Scheme	(J) Audio_Scheme	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Combination	Continuous Late Arrival	.0092	.28293	1.000	-.7539	.7722
	Continuous Course Deviation	-.7099	.28293	.076	-1.4729	.0532
	Threshold	.0633	.29069	.996	-.7206	.8473
Continuous Late Arrival	Combination	-.0092	.28293	1.000	-.7722	.7539
	Continuous Course Deviation	-.7191	.28293	.071	-1.4821	.0440
	Threshold	.0542	.29069	.998	-.7298	.8381
Continuous Course Deviation	Combination	.7099	.28293	.076	-.0532	1.4729
	Continuous Late Arrival	.7191	.28293	.071	-.0440	1.4821
	Threshold	.7732	.29069	.054	-.0107	1.5572
Threshold	Combination	-.0633	.29069	.996	-.8473	.7206
	Continuous Late Arrival	-.0542	.29069	.998	-.8381	.7298
	Continuous Course Deviation	-.7732	.29069	.054	-1.5572	.0107

Based on observed means.

			D	Min	Max	q*	P
Sing	D1	Comb-Mod	-0.22	-0.95	0.52	-1.28	
	D2	Comb-Osc	-0.60	-1.34	0.13	-3.57	
	D3	Comb-Thresh	-0.16	-0.89	0.57	-0.95	
	D4	Mod-Osc	-0.39	-1.12	0.34	-2.29	
	D5	Mod-Thresh	0.06	-0.68	0.79	0.33	
	D6	Osc-Thresh	0.44	-0.29	1.18	2.62	
Multi	D7	Comb-Mod	0.24	-0.50	0.97	1.39	
	D8	Comb-Osc	-0.82	-1.55	-0.08	-4.82	0.02
	D9	Comb-Thresh	0.29	-0.44	1.02	1.70	
	D10	Mod-Osc	-1.05	-1.78	-0.32	-6.21	0.06
	D11	Mod-Thresh	0.05	-0.68	0.78	0.31	
	D12	Osc-Thresh	1.10	0.37	1.84	6.53	0.07

	Sing-Multi	Min	Max	Q*	P
Comb	-0.75088	-1.48	-0.02	-4.44	0.04
Mod	-0.29899	-1.03	0.43	-1.77	
Osc	-0.96236	-1.69	-0.23	-5.69	0.03
Thresh	-0.30183	-1.03	0.43	-1.79	

Transformed (Natural Log) Late Arrival Reaction Time (with Combination/Continuous Late Arrival/Threshold Combined against Continuous Course Deviation)

Tests of Normality

	Late_Comparison	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
LN_L_Sing	Continuous Course Deviation	.246	10	.088	.839	10	.043
	Comb_ContLate_Thresh	.090	29	.200(*)	.959	29	.316
LN_L_Multi	Continuous Course Deviation	.242	10	.098	.915	10	.319
	Comb_ContLate_Thresh	.149	29	.101	.961	29	.341

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

Levene's Test of Equality of Error Variances(a)

	F	df1	Df2	Sig.
LN_L_Sing	1.055	1	37	.311
LN_L_Multi	3.216	1	37	.081

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Late_Comparison Within Subjects Design: scenario

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	scenario	Type III Sum of Squares	Df	Mean Square	F	Sig.
scenario	Linear	7.476	1	7.476	23.730	.000
scenario * Late_Comparison	Linear	.954	1	.954	3.029	.090
Error(scenario)	Linear	11.657	37	.315		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Intercept	228.000	1	228.000	300.624	.000
Late_Comparison	7.984	1	7.984	10.528	.002
Error	28.062	37	.758		

NASA TLX Scores

Tests of Normality

	Audio_Scheme	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	Df	Sig.
TLX_Sing	Combination	.136	10	.200(*)	.969	10	.882
	Modulated	.175	10	.200(*)	.952	10	.688
	Oscillating	.209	10	.200(*)	.886	10	.153
	Threshold	.231	9	.181	.939	9	.574
TLX_Multi	Combination	.256	10	.063	.846	10	.051
	Modulated	.146	10	.200(*)	.963	10	.818
	Oscillating	.151	10	.200(*)	.922	10	.376
	Threshold	.129	9	.200(*)	.957	9	.769

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

Levene's Test of Equality of Error Variances(a)

	F	df1	df2	Sig.
TLX_Sing	2.424	3	35	.082
TLX_Multi	1.038	3	35	.388

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Audio_Scheme Within Subjects Design: scenario

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Scenario	Type III Sum of Squares	Df	Mean Square	F	Sig.
scenario	Linear	3.364	1	3.364	.058	.811
scenario * Audio_Scheme	Linear	215.197	3	71.732	1.234	.312
Error(scenario)	Linear	2035.166	35	58.148		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	96127.322	1	96127.322	188.194	.000
Audio_Scheme	1244.548	3	414.849	.812	.496
Error	17877.551	35	510.787		
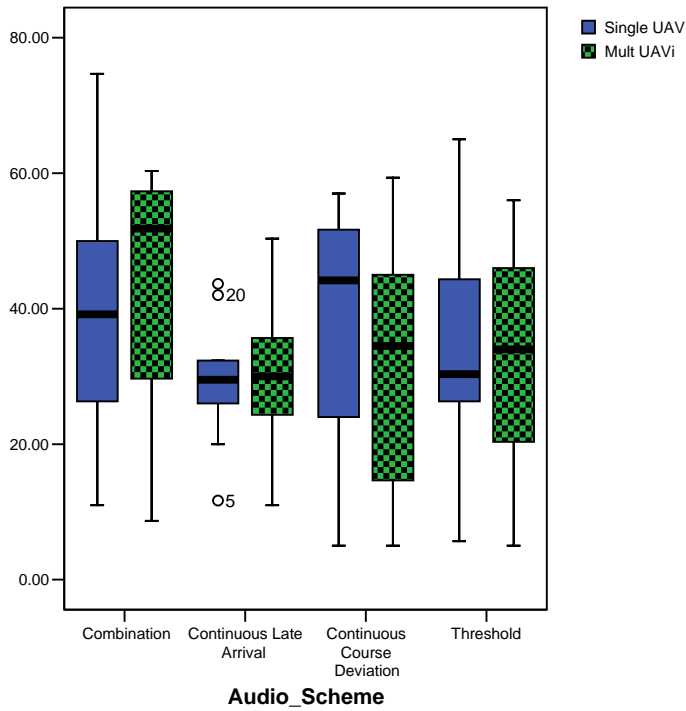
Audio_Scheme

Missed Radio Calls

Tests of Normality

	Audio_Scheme	Kolmogorov-Smirnov(a)			Shapiro-Wilk		
		Statistic	Df	Sig.	Statistic	df	Sig.
Missed_Sing	Combination	.234	10	.127	.893	10	.184
	Continuous Late Arrival	.165	10	.200(*)	.919	10	.351
	Continuous Course Deviation	.245	10	.089	.865	10	.087
	Threshold	.221	9	.200(*)	.910	9	.315
Missed_Multi	Combination	.214	10	.200(*)	.814	10	.022
	Continuous Late Arrival	.249	10	.079	.894	10	.190
	Continuous Course Deviation	.177	10	.200(*)	.909	10	.273
	Threshold	.217	9	.200(*)	.914	9	.349

* This is a lower bound of the true significance.
a Lilliefors Significance Correction

Levene's Test of Equality of Error Variances(a)

	F	df1	df2	Sig.
Missed_Sing	2.150	3	35	.111
Missed_Multi	1.684	3	35	.188

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Audio_Scheme Within Subjects Design: scenario

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	scenario	Type III Sum of Squares	df	Mean Square	F	Sig.
scenario	Linear	34.307	1	34.307	1.411	.243
scenario * Audio_Scheme	Linear	65.774	3	21.925	.902	.450
Error(scenario)	Linear	850.944	35	24.313		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	7966.465	1	7966.465	92.043	.000
Audio_Scheme	210.033	3	70.011	.809	.498
Error	3029.300	35	86.551		