APPENDIX L – LESSONS LEARNED DOCUMENT

ASSURE A46
Validation of Visual Operation Standards for Small Uncrewed Aircraft Systems

Task 5: Lessons Learned

18 September 2023

v 1.1

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS

ADS-B	Automatic Dependent Surveillance-Broadcast
ASSURE	Alliance for System Safety of UAS Through Research Excellence
EP	External Pilot
EVLOS	Extended Visual Line of Sight
GPS	Global Positioning System
KSU	Kansas State University
NMSU	New Mexico State University
PIC	Pilot in Command
RPIC	Remote Pilot in Command
UA	Uncrewed Aircraft
UAS	Uncrewed Aircraft System
VLOS	Visual Line of Sight
VO	Visual Observer
WSU	Wichita State University

1 INTRODUCTION

The Alliance for System Safety of UAS through Research Excellence (ASSURE) A46 research team investigated the effectiveness of Visual Observer (VO) performance for Extended Visual Line of Sight (EVLOS) operations. For Task 4, the Kansas State University (KSU) team conducted flight test experiments to quantify the effectiveness of VO performance, identify potential visual detection limitations, and assess VO decision-making limitations when directing Uncrewed Aircraft (UA) avoidance maneuvers. Assessments of VO performance and limitations through A46 Task 4 were part of an effort to inform VO safety training for Visual Line of Sight (VLOS) and EVLOS operations for Uncrewed Aircraft Systems (UAS). The KSU research team conducted flight tests in Gypsum, Kansas, approximately 15 miles from the KSU-Salina campus. Flight tests occurred over eight non-consecutive days, beginning on November 1, 2022, and ending on August 11, 2023. This study included a total of 19 participants.

Prior to KSU's flight tests, New Mexico State University (NMSU) conducted three days of initial flight testing to validate the research methodology and experimental plan. Initial flight tests at NMSU provided the A46 research team with valuable insight and allowed the NMSU team to generate recommendations to improve the experiment design. These recommendations are articulated in the latter sections of this document.

The following sections describe the research team's approach to measuring dependent variables throughout testing and capturing lessons learned from the NMSU and KSU-led flight test campaigns. Lessons learned from the flight test campaigns inform actions for future methodological design. This report also captures impressions and perceptions from the research team about the research methodology and experiment design based upon lessons learned.

2 RECRUITMENT, SCHEDULING, AND SET-UP

2.1 Participant Recruitment

KSU performed multiple rounds of participant recruitment. Recruitment methods included campus-wide emails, classroom visits, word-of-mouth advertising, and posters on campus in common areas. Low participation was a recurrent problem, which made multiple rounds of recruitment necessary. As a component of the second round of recruitment, the team received approval from the sponsor to compensate participants. Compensating participants marginally increased interest, yet low participation was evident throughout this study. The third and final round of recruitment advertised a reduction in participant time commitment, reducing test blocks from five hours to four. The initial five-hour time block accounted for VO training, briefings, flight test runs, and breaks, including incidental delays. While many individuals voiced interest in participating, they indicated that even dedicating four hours of their day was not feasible.

2.2 Scheduling

The weather in central Kansas created scheduling complexities for flight testing. Originally, testing was planned for the summer of 2022 but was rescheduled to the fall of 2022 due to significant weather delays. Fall of 2022 proved an inopportune time, as many days experienced below-freezing temperatures, which made conditions unsafe for participants. Only two flight test campaigns were conducted during the fall of 2022. Due to inclement weather conditions in the fall of 2022, an additional round of testing was scheduled for the spring of 2023. The research team anticipated spring would bring warmer temperatures and safer conditions for participants. However, spring conditions were still not conducive for flight tests. Rain delayed testing, and high winds made it impossible to persuade participants they were participating in a real flight operation. Consequently, only two flight test campaigns were completed during the spring of 2023. The team conducted a third and final round of testing in the summer of 2023, adding an additional four days of testing.

2.3 Test Set-up

The day prior to the testing, the research team prepared all documents required for the day of testing. These documents included the air boss record chart, perceptual testing documents, the informed consent form, and other administrative and operational records. The "station bins," as seen in Figure 1, contained all the necessary evaluation tools and were set out to ensure all equipment was accounted for prior to departing for the test site. The station bins corresponded with the numbered station where the participant (VO), the Remote Pilot In Command (RPIC), and an assigned researcher were positioned on the day of testing. Two to three stations were in the field for each day of testing. The number of stations was dependent upon the number of participants. As previously mentioned, the station bins (Figure 1) contained all the necessary testing equipment, such as tables, chairs, radios, and documentation for the RPIC and researchers. The test station bins were organized by station number prior to testing. The day prior to testing, the lead researcher emailed the participants to reconfirm availability and provide directions to the testing location. The emails included instructions to bring food/water and to wear proper attire for environmental conditions.

Figure 1. Staging Station Bins for Flight Testing.

During each day of testing, the research team met 1½ hours prior to the scheduled experiment start time to brief, load the truck and trailer, and travel to the test site. Participants were asked to meet the team at the testing location at a predetermined start time. This arrangement allowed the research team time to set up stations and equipment and enabled participants to leave when testing was complete. The KSU research team used the following lessons learned from NMSU's initial round of testing and incorporated them into their procedures and experiment setup:

1. Generators used in the field should be placed far away from the stations to prevent interference with participants' hearing during testing.
2. Take photos of all areas for test setup reference – staged with team personnel, not with participants, as seen in Figure 2.
3. Have a dedicated station for onsite light and sound meters.
4. Provide a compass so participants have a reference for cardinal directions.

Figure 2. Station Reference Set-up with Team Personnel.

The KSU team modified the original setup procedures, incorporating lessons learned by NMSU. The generators used at each station were located approximately 20 feet from the participants. Wind direction was also considered when setting up the generators. Generators were placed downwind to reduce ambient noise. The team captured a reference photo of the testing area during an initial dry run. The KSU team also placed ground marking flags to indicate the setup location of each station. These flags helped ensure test station locations remained consistent between testing days. The team recorded the Global Positioning System (GPS) coordinates for each test station prior to each round of flight testing. GPS coordinates served as a backup to the ground marking flags, as activity in the test field could cause marker flags to shift or be crushed/removed by farm equipment.

Researchers placed weather stations with a single light and sound meter between the three stations. The location of each weather station ensured light and sound meters remained undisturbed during testing. During the initial dry run, the team identified that when sound meters were placed at the VO stations, they captured unwanted noise – e.g., typing, conversations, and other undesirable artifacts. Similarly, if the light meters were on the station tables, shadows from the crew, participants, and the shade canopies negatively impacted the ambient light readings. Locating the light and sound meters between test stations offered more accurate data collection with less interference.

Further, each table at each station was marked with a printed compass rose indicating the cardinal directions. Tables at each station were set up with participants facing north to reduce confusion. Additionally, each station included a computer for the RPIC, data collection forms, GoPro cameras, and a VO script for the participants to reference during testing. The VO script served as a reminder of key information a VO should call out – i.e., identify intruder aircraft, estimate intruder altitude and direction, suggested maneuver, etc.

3 TESTING COMPONENTS

Flight tests began with a 30-minute VO training PowerPoint, which provided an overview of the prerequisite requirements to serve as a VO. This training session also provided a description of the experiment tasks. The research team determined that 30 minutes for this training would suffice to accomplish the participant training requirements. The VO training also included practice scenarios to familiarize participants with the language outlined in the training. Originally, the experiment plan indicated that VO training would be conducted in the field. However, NMSU found outdoor training was ineffective and not conducive as a training environment. With recommendations from NMSU, KSU transitioned to indoor training within a mobile command trailer to reduce outside noise, reduce glare on computer screens, and to provide a more comfortable environment for training participants. Additional suggestions provided by NMSU are as follows:

1. NMSU suggested providing more examples for review. The training provided five slides with interactive examples to practice the concepts associated with VO tasking.
2. Providing participants a real-life visual examples of a 500-foot distance provided context for approximating visual measurements. The VO trainer used a water tower as an example.

3.1 Perceptual Testing

Following the VO training, participants underwent a series of perceptual tests, including a Snellen vision test, a color deficiency test, and an auditory test. NMSU determined these tests were best conducted indoors to limit the potential effects of varying outdoor environmental conditions. Due to limited space, KSU could not perform all perceptual tests indoors. The color deficiency and hearing tests were conducted inside the KSU mobile command trailer, where adequate lighting was available. The Snellen eye test was the only test conducted in the outdoor environment as space inside the mobile command trailer was limited, and participants were required to maintain a 20-foot distance from the test chart. The 20-foot distance for the Snellen eye test did cause issues for some individuals in the outdoor environment, particularly on bright/sunny days. In one instance, test administrators had to relocate the Snellen eye chart from the East side of the trailer to the West, as the sun impeded participants' ability to read the chart. Future testing would benefit from a dedicated indoor location where test administrators could run all participants through the testing with consistent environmental conditions.

3.2 Communication

NMSU established communication procedures that were used throughout this study. These communication procedures incorporated three sets of radios, each with a defined function. The first radio set connected all ground personnel – i.e., RPICs and the air boss. The second radio set connected the air boss to the pilot flying the intruder aircraft. Communications to the intruder aircraft were isolated such that VO teams could not hear radio calls that may give away the aircraft's position. KSU added a third set of radios for the RPIC to call out maneuvers to a fictional External Pilot (EP). The KSU team added the third radio set and the fictional EP to ensure the simulated encounters appeared realistic to participants. All participants were informed the EP was located approximately one mile from the VO stations.

Before the flights began, KSU research staff informed participants the RPIC would call out suggested maneuvers to the EP during each encounter, but EPs would be too busy maneuvering the aircraft to respond. For future test flights, an additional team member acting as the EP may help convince the participants they are participating in a real flight operation.

A minor change in the VO test location identifier was made between the NSMU test flights and KSU's flights. For radio callouts and data entry, NMSU used the identifiers "North," "South," and "West." To reduce potential confusion on radio callouts, KSU opted to identify the VO locations as "Station 1," "Station 2," and "Station 3." Both sets of identifiers worked for the specific universities, and no confusion resulted from this difference in nomenclature. Either option would work well for future testing, but identifiers such as those used by KSU would limit radio communication errors or confusion.

3.3 Flight Test Execution

During flight tests, the KSU team modified the test cards as required due to changes in wind speed and direction. Wind conditions varied on each day of testing, which impacted the intruder aircraft's performance. Thus, it was necessary for the KSU team to incrementally shorten the inbound/outbound legs of the intruder aircraft's flight path from five miles to three miles to maintain consistent timing. Figure 3 provides an example of the intruder aircraft's flight path. These modifications allowed the pilot of the intruder aircraft to make consistent turns around a known point within the desired operating area. During this time, the air boss, researchers, and the intruder aircraft pilot were the only individuals cognizant of the intruder aircraft's altered flight path. The overall time between runs and the ability to keep participants engaged also influenced the decision to shorten the turn-around time. The air boss and researchers met multiple times throughout each test day to discuss the state of testing. The first day of testing was slower than anticipated, averaging ten to fifteen minutes between runs, leading to fewer data points and participant boredom. This influenced the decision to gradually shorten the inbound/outbound leg from the originally planned five-mile distance to three miles.

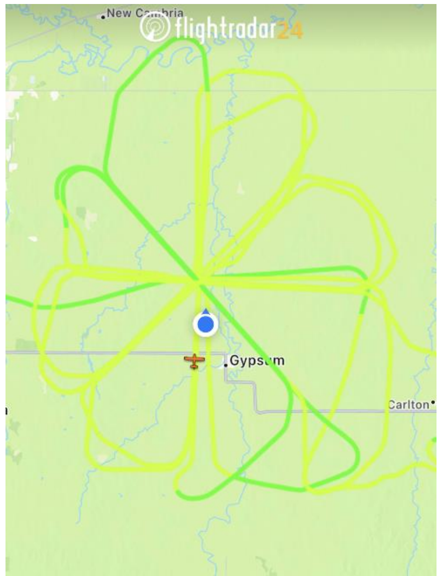

Figure 3. Example Intruder Aircraft Flightpath – Turns and Inbound/Outbound Legs.

During the inbound/outbound leg, the researchers were not actively engaged with participants. Instead, the lead researcher on site worked with the air boss to determine if the intruder aircraft was visible with the reduced turn-around distances. The air boss and the lead researcher on site agreed that three miles was an ideal turn-around distance to increase the efficiency of test runs, maintain participant engagement, and ensure participants could not see the intruder aircraft. Figure 4 illustrates the view of the intruder aircraft at one of the VO stations as it passed overhead.

Figure 4. Intruder Aircraft Flying Over the Testing Area.

The KSU team utilized various electronic devices during testing, including computers, radios, and meters, to capture ambient light and sound. All these devices had some form of internal clock. Prior to deployment for testing, the team synchronized the internal clocks for all computers, radios, and data collection devices. The air boss called the experiment stations for a time check every 15 minutes as a secondary measure of ensuring time synchronization. Researchers at each station used these 15-minute callouts to ensure that all computers and test equipment at each station displayed the correct time. If the time was different, the RPIC or researcher noted the discrepancy on their data collection sheets. No time-synching issues were found during flight testing. Future iterations of flight testing would benefit from utilizing similar practices for data consistency.

Two noteworthy deviations from the original experiment plan were removing the post-run questionnaire and RPIC overrides to the VO's suggested maneuvers. During the initial dry run, RPIC found that the post-run questionnaire was too cumbersome to complete as they reset the simulation and recorded data for each run. Similarly, RPICs found that overriding the VO's suggested maneuver while recording data was impractical. A separate person solely responsible for the time sheets/data recording would be beneficial for future testing. This would allow the RPICs to focus exclusively on maneuvering the simulated UA and monitoring the status of the simulation without the responsibility of recording data.

3.4 Intruder Aircraft/Pilot Perspective

The intruder aircraft consisted of a single-engine fixed-wing airplane piloted by a single Pilot in Command (PIC). The aircraft's autopilot ensured each approach track was correctly aligned and that each track was consistent. However, ground speeds were harder to maintain on test days with windy or gusty conditions, and inbound/outbound tracks were adjusted for more consistent runs. When required, the intruder pilot used the aircraft's flaps to maintain a more consistent ground speed, but flying at lower speeds at low altitudes posed its own risks. The intruder aircraft utilized navigation and strobe lights on initial tracks. This increased the visibility of the aircraft and made it easier for participants to track the aircraft while on the wagon wheel tracks. For subsequent tracks, the strobe lights were turned off so participants could not track the intruder aircraft between runs.

Automatic Dependent Surveillance-Broadcast (ADS-B) onboard the intruder aircraft provided situational awareness of other non-participating traffic in the area. However, the workload onboard the intruder aircraft was moderately high, and monitoring all local traffic while flying low-level approaches became burdensome for the pilot. Therefore, the air boss also used an ADS-B receiver in the UAS Mobile Operations Command trailer to monitor ADS-B and inform the intruder aircraft pilot about local traffic. Non-participating air traffic did affect the intruder aircraft's approaches. If non-participating air traffic created a hazard or negatively affected the experiment, the pilot of the intruder aircraft established a loiter outside of the VO's field of view and waited until the non-participating traffic passed. When the operating area was clear, the intruder pilot coordinated with the air boss to resume the experiment.

The KSU team collected data from the intruder aircraft utilizing the onboard Garmin G1000 avionics suite and a backup GPS tracker mounted to the dash. One flight resulted in the G1000 not collecting data due to the full SD card. To rectify this, the PIC would clear the onboard G1000 SD card before the intruder aircraft took off from the Salina Regional Airport (KSLN).

3.5 Data Analysis

Wichita State University (WSU) analyzed the data collected during flight tests performed by KSU and NMSU. The data collection format used by both KSU and NMSU was consistent for their individual tests, with minor variations between each university. Due to the minor variations in the formats between KSU and NMSU data sets, WSU used different scripts to extract the data collected by each university. Using different scripts for data extraction did not affect the outcomes or conclusions reached from the experiment.

NMSU utilized multiple GPS receivers to ensure redundancy for recording the intruder aircraft flight parameters. This data included baseline GPS data points such as latitude, longitude, and altitude. A portable Qstarz GPS unit recorded the intruder aircraft's heading. Researchers retrieved the aircraft's internal log after each flight. During the data analysis, WSU observed that the Qstarz GPS data recording did not cover all the flights, and therefore, the intruder heading data was unavailable for several encounters. Future testing would benefit from ensuring the flight recording systems record all the basic flight parameters, including the heading, for the entire duration of flight tests.

NMSU relied on audio recordings of the VO responses to determine the VO callouts for intruder detection and suggested maneuvers. However, some of the VO responses could not be understood, as moderate wind gusts affected the clarity of audio recordings. The time-stamp callouts were clearly recorded and essential to perform the time-sync between the intruder aircraft, UAS flight data logs, and the VO responses.

KSU implemented the lessons learned from NMSU during their flight tests. The KSU team manually recorded the VO callout time stamps in a spreadsheet instead of relying solely on the audio recording of the VO callouts. This significantly reduced the data analysis processing time. Overall, the data collection methodology and formats were consistent and allowed for a quick turnaround time for data analysis.

3.6 Post Testing

After the participants completed the flight tests, they were escorted back to the mobile command trailer for a debrief while the research team went through the experiment closedown procedures. The participants were informed during the debrief that flight tests used a simulated UA. Participants were asked if they were surprised by this information, and approximately half were unsurprised. Many stated they had been looking for any sign of the UA where the research team suggested that it would be but could not locate the UA in the airspace. Others stated they knew it was simulated because no one responded to the RPIC maneuver callouts on the radio.

To rectify this for subsequent testing, the air boss went into more detail the flight operation, specifying that participants would not see the UA because of the tree line and the UA's cruising altitude. The air boss also explained why participants would not hear the EP respond to the RPIC. This was not discussed on the first day of testing. Fewer participants said they knew it was simulated after making those changes to the VO training. Future testing would benefit from having a designated EP to mimic real-world operations more closely. This was suggested by participants and was captured in Section 3.2 of this report (Communication).

4 CONCLUSIONS AND RECOMMENDATIONS

ASSURE A46 flight testing investigated critical aspects of VO performance during EVLOS operations. Based on the research team's experience, the following seven items should be considered for follow on work to ensure consistency and efficiency during future flight test operations:

> For test setup, schedule a dry run a few days before flight testing to ensure all those involved understand their assignments and the flow of the days.

During the dry run, the team should take pictures of each test station to ensure consistency during set up on the day of flight testing.
Place equipment in bins or containers for each station to keep supplies organized during set up and tear down.
Perform all perceptual testing inside a well-lit room with minimal noise. While this was the initial plan, the perceptual testing was moved outside due to limitations onsite at the test location. Performing such tests inside will reduce potential errors in perceptual data collection and reduce strain on the participants.
Providing a visual example of a 500 ft distance during VO testing, such as a tower, pole, or building, could increase participants' understanding and improve their performance when perceiving objects at a distance. This could be especially helpful to novice individuals and/or those with little or no VO experience.
The KSU team believes EP responses via radio upon suggestion and completion of UA maneuvers would improve the realism of the experiment and aid in convincing participants that they are making decisions for real aircraft.
Conducting this research in a simulated environment would mitigate the need for a 5-hour block of time from participants and staff. A simulation would also permit a team to conduct more runs, allowing a team to study additional variables in a quantitative manner. The ability to analyze more variables in a more controlled environment would contribute to a more robust understanding of the performance effectiveness of VOs in different conditions.