DOT/FAA/TC-19/25

Federal Aviation Administration
William J. Hughes Technical Center
Aviation Research Division
Atlantic City International Airport
New Jersey 08405

Validation of Unmanned Aircraft Systems Contingency Procedures and Requirements En Route Human-in-the-Loop Simulation Technical Report

Lacey Thompson, ANG-C35
Randy Sollenberger, PhD ANG-E25
Amy Alexander, PhD MIT Lincoln Laboratory
Alex Konkel, PhD DSoft Technology, Engineering, and Analysis, Inc.
Eamon Caddigan, PhD DSoft Technology, Engineering, and Analysis, Inc.
John Bradley, ANG-E62
Steve Cullen, A3 Technology, Inc.

December 2020

Final report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/TC-19/25	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle Validation of Unmanned Aircraft Systems Contingency Procedures and Requirements En Route Human-in-the-Loop Simulation Technical Report	5. Report Date December 2020
	6. Performing Organization Code
7. Author(s) Lacey Thompson, ANG-C35 Randy Sollenberger PhD, ANG-E25 Amy Alexander PhD, MIT Lincoln Laboratory Alex Konkel PhD, DSoft Technology, Engineering, and Analysis, Inc. Eamon Caddigan PhD, DSoft Technology, Engineering, and Analysis, Inc. John Bradley, ANG-E62 Steve Cullen, A3 Technology, Inc.	8. Performing Organization Report No.
9. Performing Organization Name and Address Federal Aviation Administration UAS Engineering Branch William J. Hughes Technical Center Atlantic City International Airport, NJ 08405	10. Work Unit No. (TRAIS)
	11. Contract or Grant No.
12. Sponsoring Agency Name and Address Federal Aviation Administration NextGen New Entrants Division and NextGen Human Factors Division 800 Independence Avenue, S.W. Washington, DC 20591	13. Type of Report and Period Covered Technical Report
	14. Sponsoring Agency Code ANG-C1, ANG-C2

15. Supplementary Notes

16. Abstract

Objective: We investigated the time needed for air traffic control specialists (ATCS) to respond to Unmanned Aircraft System (UAS) lost link events in the en route environment. **Background:** When a UAS loses link, ATCS must react to an uncontrolled aircraft in the airspace. The time needed to do so has not been studied. **Method:** We conducted a human-in-the-loop (HITL) simulation to examine the human factors impact of lost link timing as well as related information delivery methods. **Results**: Participants preferred to receive information quickly via automation and to have more coordination time. **Conclusion:** We suggest that lost link information be given to ATCS within 2 minutes and lost link routes not involve a turn for at least 4 minutes after route delivery, particularly if automation is not assisting them.

17. Key Words Air Traffic Control, Automation, Contingency, En Route, Integration, Lost Link, Unmanned Aircraft Systems	18. Distribution Statement This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages	22. Price

ACKNOWLEDGEMENTS

This research was sponsored by the FAA's NextGen New Entrants Division (ANG-C2) and the Human Factors Division (ANG-C1), so first and foremost we would like to thank our sponsors Sherri Magyarits, Sherry Chappell (who has since retired), and Bill Kaliardos. We would also like to extend our gratitude to Greg Feldman and Dan Gutwein of Cavan Solutions, who were instrumental during the development phase of the research.

This research was a collaborative effort accomplished by a cross-cutting team of research analysts, human factors researchers, air traffic control subject matter experts, computer programmers/developers, and research support staff at the William J. Hughes Technical Center. We also extend our thanks to the NextGen Integration and Evaluation Capability (NIEC) staff, the Target Generation Facility (TGF) team, and the Distributed Environment for Simulation, Evaluation, and Experimentation (DESIREE) team for the dedicated support during scenario development and conduct of the simulation.

To those not mentioned who supported this research effort, we thank you.

Mike Beauvais, FAA, AJV-83, Air Traffic Control Subject Matter Expert
Brian Richardson, First American Systems and Services, AJV-83, Air Traffic Control Subject Matter Expert
Diane Flynn, Human Solutions Inc., AJV-83, Air Traffic Control Subject Matter Expert
Kevin Aurandt, Human Solutions Inc., AJV-84, Air Traffic Control Subject Matter Expert
Gerald Van Hook, Human Solutions Inc., AJT-2, Air Traffic Control Subject Matter Expert
Joey Ross, FAA, AJV-724, Air Traffic Control Subject Matter Expert
Matt Kukorlo, FAA, ANG-E141, Air Traffic Control Subject Matter Expert
Kim Bender, CSSI, ANG-C2, Human Factors Analyst
Dan Wegmann, GDIT, ANG-E142, DESIREE Developer/NIEC Support
Kevin Sounthavong, GDIT, ANG-E142, DESIREE Developer/NIEC Support
Kevin Krout, GDIT, ANG-E142, DESIREE Developer/NIEC Support
Chris Continisio, GDIT, DESIREE Developer/NIEC Support
Kevin Desmond, GDIT, ANG-E142, DESIREE Developer/NIEC Support
Gordon Bond, GDIT, ANG-E142, NIEC Support
George (Lee) Starks, GDIT, ANG-E142, NIEC Support
Cindy Salas, GDIT, ANG-E142, NIEC Technical Task Lead
Robin Peterson-Brown, FAA, ANG-E142, NIEC Lead
Dan Fumosa, FAA, ANG-E142, Concepts Integration Section Manager
Gary Mueller, FAA, ANG-E1, Laboratory Integration Lead
Laura Hamann, FAA, ANG-E25, Human Factors Branch DESIREE Developer
Terry Corley, Advanced Sciences & Tech, LLC, ANG-E153, Comm Support
Mike Galusha, FAA, ANG-E153, Comm Support
Timothy Swantek, ATEC, ANG-E161, TGF Developer/Sim Pilot Lab Support
Mai Truong, ATEC, ANG-E161, TGF Developer/Sim Pilot Lab Support
Richard Smail, FAA, ANG-E161, Target Generation Facility (TGF) Section Manager
Samantha Fullerton, JVN, ANG-E161, TGF Developer/Sim Pilot Lab Support
Bruce Slack, FAA, ANG-E161, Simulation Pilot
Bob Engiles, FAA, ANG-E161, Simulation Pilot

Tru Hall, FAA, ANG-E161, Simulation Pilot Coordinator
Tom Vispisiano, HiTec, Simulation Pilot
Tom Weston, HiTec, Simulation Pilot
Glenn Thompson, HiTec, Simulation Pilot
Frank Johnson, HiTec, Simulation Pilot
Eric Olson, HiTec, Simulation Pilot
Larry Pecan, HiTec, Simulation Pilot
Walt Ellis, HiTec, Simulation Pilot
Andre Lonchambon, HiTec, Simulation Pilot
Craig Johnston, HiTec, Simulation Pilot
Guy Bachi, HiTec, Simulation Pilot
Ron Strac, HiTec, Simulation Pilot
Max Hutchins, HiTec, Simulation Pilot
Barry Mastangelo, HiTec, Simulation Pilot
Charlie Haubrich, HiTec, Simulation Pilot
Kim Mortenson, HiTec, Simulation Pilot Manager
Mary Rozier-Wilkes, FAA, ANG-E161, Simulation Pilot Coordinator/Scheduler
Steve Weidner, NATCA, National UAS Representative
Juan Keller, A-3 Solutions Inc., ANG-C35, LimeSurvey Questionnaire Support

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Background

The goal of the Unmanned Aircraft Systems (UAS) community is to support advanced integration of UAS into the NAS, leading to "file and fly" capabilities that enable routine UAS operations in non-segregated civil airspace. Without an onboard pilot, there is a significant reliance on the Command and Control data link and a greater emphasis on the loss of functionality associated with off-nominal (contingency) events such as loss of link.

When a lost link does occur, air traffic control specialists (ATCS) must react by coordinating with other aircraft, since the UAS can no longer be controlled, and with ATCS in adjacent sectors. The time needed to do so, and the human factors impact on ATCS, is currently unknown.

Objectives

This research effort examined time lags for providing UAS lost link intent and maneuver information to air traffic control specialists (ATCS) via three different delivery mechanisms. It also assessed the impact/risks (e.g., workload, situation awareness) of the contingency procedures on en route ATCS. The purpose of the research was to determine acceptable time lags for delivery of lost link information.

Methods

This study assessed two timing parameters, T1 and T2. T1 was the lag between when the lost link UAS squawked 7400 and when the UAS intent was provided to the participant. T2 was the lag between when UAS intent information was provided and the aircraft's first maneuver off the original flight plan. We evaluated two T1 times, 1 and 4 minutes, and two T2 times, 2 and 4 minutes.

The T-times were evaluated with three different lost link intent delivery mechanisms: manual, automaton, and voice.

- The manual delivery method consisted of a manager (one of the researchers) providing the participant with a written copy of the UAS lost link routing.
- The automated delivery method consisted of the UAS lost link routing being pushed to the participant similarly to the current Airborne Reroute (ABRR) functionality of ERAM.
- The voice delivery method consisted of the UAS pilot providing the participant with the lost link routing at the expiration of the T1 time.

Objective and subjective data were collected to evaluate the impact of T-times and delivery mechanisms on en route airspace operations. Dependent measures collected during the simulation included safety, efficiency, communication, situation awareness measures, as well as controller workload.

Conclusions and Recommendations

Participants indicated that they could wait up to 4 minutes to receive lost link procedures (T1 time) although they preferred the shorter time (1 minute). Participants also expressed a need for at least 4 minutes between receiving the lost link procedure and the beginning of lost link maneuvering (T2 time), but even longer would be better. Notably, many lost link route amendments were not entered before the UAS began maneuvering when T2 was short (2 minutes). Participants had a strong preference for the automated (ABRR-like) mechanism compared to the other delivery methods. The manual method was also acceptable.

We recommend regulations that ensure quick delivery of lost link route information to ATCS along with a sufficient amount of time before a lost link UAS would turn off its current route. These times would likely need to be longer if lost link information is not provided via automation.

1. INTRODUCTION

This document, Validation of Unmanned Aircraft Systems (UAS)[1] Contingency Procedures and Requirements En Route Human-in-the-Loop (HITL) Simulation Technical Report, describes one of many activities in a portfolio that supports requirements for investigating issues pertaining to the integration of UAS into the National Airspace System (NAS).

Within the aviation community, interest in using UAS for a broad range of purposes is increasing, making UAS access to the NAS a priority. The Federal Aviation Administration (FAA) reached a significant milestone with the implementation of Part 107, which permits small UAS (less than 55 pounds) daytime operations within visual line-of-sight and at or below 400 feet above ground level in uncontrolled airspace. However, outside the limits of Part 107, requests for access to the NAS are subject to technical and operational assessments of the specific UAS operation in question. It has been a growing imperative within the UAS community, including public and civil users, to reduce these restrictions and support more routine access in order to improve and advance integration of UAS into the NAS. Therefore, validated operational standards and policies need to be established.

While the initial focus has been on integrating UAS into today's NAS, it is necessary to maintain awareness of how the NAS will evolve in NextGen. With the increased use of UAS, issues that are unique to these aircraft are likely to arise. Many UAS do not have the same performance characteristics as manned aircraft, communication procedures between air traffic control (ATC) and the UAS operator/pilot-in-command (PIC) are dissimilar, and lost link and other UAS-related off-nominal events can occur. These inevitable issues pose new challenges for air traffic control specialists (ATCS) and it is essential that they be addressed in order to maintain NAS safety, efficiency, and capacity, which are cornerstones of NextGen.

1.1. BACKGROUND

In order to reach full integration of UAS into the NAS, research is needed to validate operational standards, policies, and procedures, which can then be implemented for the ATC community and pilots.

---

[1] A UAS is the aircraft itself and all of the associated support equipment, control station, data links, telemetry, communications and navigation equipment, etc., necessary to operate the UA. The UA is flown by a pilot via a ground control system or autonomously using an on-board computer, communication links, and any additional equipment that is necessary for the UA to operate safely. A UA operates without the possibility of direct human intervention from within or on the aircraft.

Within the past few years the FAA has conducted several studies concerning UAS contingency conditions, such as lost link[2]. Most of the earlier studies were exploratory in nature and illustrated the potentially adverse effects that UAS contingency events can have on operations in the NAS and on ATCS. These studies also demonstrated the resiliency and capability of ATCS in mitigating these effects while still maintaining priority of safety above all other factors.

1.1.1. Previous Related Research – En Route Contingency Operations

A research activity entitled *Human Factors Investigation of UAS in the NAS: En Route Contingency Operations* was conducted to enable the development of contingency procedures and associated system requirements that will meet the needs of ATCS and unmanned aircraft operators. The goal of the research was to identify areas where standardization would make sense (driven by particular characteristics of specific en route sectors, airspace, UAS platforms, UAS operations, and other variables) and the information that ATCS may need during contingency events.

As part of this research, a literature review was conducted to gather information from related studies. One takeaway from the literature review was that although research has been conducted about the information that is required during a contingency event, further investigation was needed in this area in order to identify specific information requirements. Some information (e.g., lost link points, divert/contingency points) is currently provided in the flight plan but this information is not necessarily standard across facilities. Furthermore, the modality in which this information is presented is not standardized.

In the next stage of the research, a user needs analysis was performed to identify potential procedures and/or technological solutions that may help mitigate the impact of a UAS contingency event. The analysis was conducted via a series of knowledge elicitation sessions in which information was gathered from ATCS and UAS operators to identify their needs for both current and future operations in contingency situations. Ideally, ATCS would like contingency procedure and other pertinent information to be readily available to them in an easy-to-digest format. ATCS stated that they would like the ability to know if a UAS is lost link by a beacon code and would like for information pertaining to the lost link (e.g., important phone numbers, maps) to be easily accessible.

In general, ATCS exhibited a desire for predictability. During lost link events, ATCS would like to know if the UAS is lost link by the beacon code, how long a UA will be lost link before squawking 7400[3], where it will go if it loses link, what its intentions are, how long before it maneuvers once it squawks 7400, and what will happen if it does not regain link. Knowing that

---

[2] A partial list: Multi-UAS Operational Assessment: Class D Airspace Simulation (AKA Victorville; Buondonno et al., 2012), Initial NAS Integration Simulation (INI-Sim), UAS Operational Assessment: Contingency Operations (Pastakia et al., 2015). All of these studies were conducted at the William J. Hughes Technical Center.

[3] According to FAA Notice N JO 7110.24, there are two components to lost link: one if the uplink that transmits command and control (C2) instructions to the aircraft and the second is the downlink which relays the operation/status of onboard systems within the aircraft to the ground control station. If either link is disabled or malfunctions, the result is defined as 'lost link' and some aircraft transponders will automatically reset to code 7400, execute a pre-programmed flight profile, and controllers will react accordingly. NAS automation changes have been made to all NAS platforms to recognize the Mode 3 7400 code.

the UAS is going to follow a specific procedure or will execute a maneuver in a specified amount of time could help ATCS have more confidence in the predictability of the aircraft.

The research described above highlighted issues, both near and far term, of integrating UAS into the NAS. It identified some issues for which near-term solutions may be implemented:

1. To require UAS operators to file flight plans with ground-based navaids or fix-radial distances instead of latitudes/longitudes,
2. To have common and concise briefing sheets for ATCS readily available at the position so that they can access important information quickly, and
3. To develop and administer standard training for ATCS that work UAS.

A farther-term goal would be to integrate UAS contingency re-routing with En Route Automation Modernization (ERAM) and/or En Route Decision Support Tool (EDST) in order to better communicate what action the UAS will execute when it experiences a contingency event to ATCS.

1.1.2 Previous Related Research – ATC Receipt & Display of Contingency Information

In 2016-2017, a research effort entitled ATC Receipt and Display of Contingency Information (Pastakia et al., 2017) complemented and expanded on the En Route Contingency Operations Research. The project was initiated to explore contingency information and display needs. The objective of this task was to determine information and high-level display requirements for terminal and en route automation systems when an Instrument Flight Rules (IFR) flight plan is required.

The main activity in the research was the conduct of two cognitive walkthroughs, one for terminal and one for en route domains. During each cognitive walkthrough, variations of several scenarios (nominal and off-nominal) were presented to the participants. Researchers prompted participants to describe how they would handle each situation and what information and/or tools they would need to do so. Following each scenario and at the conclusion of the walkthrough participants completed questionnaires to provide data on their preferences for information, procedure, and display needs.

Overall results from both terminal and en route cognitive walkthroughs highlighted that ATCS need time to react after seeing that a UAS has lost link. Participants felt that there should be a standardized time (yet to be determined) between when the UAS squawks lost link and executes the lost link contingency maneuver. ATCS need this time to confirm what the contingency plan is and to coordinate with other aircraft and sectors to accommodate the plan.

As UAS operations become more commonplace and UAS procedures become more familiar, information needs of ATCS may potentially change. Perhaps these changes will require software upgrades and/or automated information exchange with UAS operators. Participants in the en route cognitive walkthrough felt that one potential requirement was that the UAS should appear on the scope, as with any aircraft, and that UAS should be denoted as such for ATCS' situation awareness. They felt that some type of signifier should be on the scope, a secondary display, or both; they identified a number of ways to note an aircraft as a UAS: color coding the data block, adding a particular letter/code to the aircraft identification in the data block, or adding some signifying element to a secondary display.

1.2. OBJECTIVE

The objective of the research presented in this document was to evaluate time lags (referred to as T-times) for providing UAS lost link intent and maneuver information to ATCS via three different delivery mechanisms (manual, automated, and voice) in an effort to determine acceptable time lags and compare the efficacy of the delivery mechanisms. Investigation of these time lags could lead to the development of UAS requirements and could help determine ATC automation requirements for management of UAS lost link events.

A real-time HITL simulation allowed researchers to immerse ATCS from the field into an ATC lab environment in order to present carefully designed scenarios to them to test specific variables, while assessing performance metrics. The HITL also allowed researchers to assess the impact of UAS lost link contingency events on the overall traffic operation in the en route environment and on ATCS.

1.3. SCOPE

The Validation of UAS Contingency Procedures and Requirements research supports the NAS Segment Implementation Plan (NSIP). The activities conducted under this research contribute to the Separation Management – UAS Concept Validation and Requirements Development. The research effort includes two HITL simulations, one in the terminal environment and one in the en route environment, that are designed to assess the impact of UAS integration into the NAS. The en route HITL simulation, described in this document, is a follow-on activity to the ATC Receipt and Display of Contingency Information and En Route Contingency Operations research activities.

2. METHOD

2.1. PARTICIPANTS

We recruited 10 ATCS from Level 10 through 12 ARTCCs across the country to participate in the HITL simulation. We required that CPCs currently be active. Previous experience with UAS was preferred but not a requirement. The task lead and sponsor from ANG-C2 coordinated with AJV-7, ATO Technical Labor, and NATCA to secure ATCS as simulation participants.

It is commonplace at many en route facilities to have a data controller, often referred to as the D-side, to assist the radar or R-controller. These two controllers often act as a team to manage flight operations. In an effort to better assess the overall impact to the R-side position, the D-side position was not staffed during the simulation.

During the simulation, each participant controlled traffic in the same sector (ZOA 30/33; see section 2.6). The setup for this HITL allowed researchers to have participants sit at different positions (see Figure 2) in the lab and run scenarios simultaneously. During some weeks only two participants were available, so the positions (A, B, or C) were chosen at the research team's discretion.

2.2. LABORATORIES AND EQUIPMENT

This research effort made use of the WJHTC's laboratory infrastructure to successfully complete all of the research activities. In particular, we utilized the FAA NIEC laboratory. The NIEC is the FAA's research platform to explore, integrate, and evaluate advanced aviation concepts through simulation activities resulting in concept maturation and requirements definition; it is a flexible and extensible environment that consists of real and simulation systems as well as infrastructure capabilities to support the environment. Characteristics of the NIEC include:

- Representation and integration across multiple NAS domains in one facility
- A real-time, rapid prototyping and simulation environment that simulates the NAS while integrating NextGen concepts
- Voice communications capabilities
- Audio, video, and data recording capabilities
- Flexibility to support multiple concurrent studies

2.2.1 Audio-Video Recording System

We used the NIEC's audio-video recording system to record communication during the simulation. Each controller position had a microphone nearby to record communications and ambient noise. Each position also had an overhead camera to record the scope, the workload assessment keypad (WAK), and controller actions. The audio-video recordings allowed us to review playbacks of the simulation as needed to provide clarity about specific situations (e.g., in cases where the WAK response was missing, we could review the video to determine if a response was attempted).

2.2.2. Communication

A simulated communication environment allowed for realistic air-ground and ground-ground voice communication between simulation participants, simulation pilots, and ghost controllers.

Each ATCS position had an Interim Voice Switch Replacement (IVSR) communication system (Figure 1). Additionally, each ghost controller position was equipped with an interphone and the ability to monitor the active frequencies.

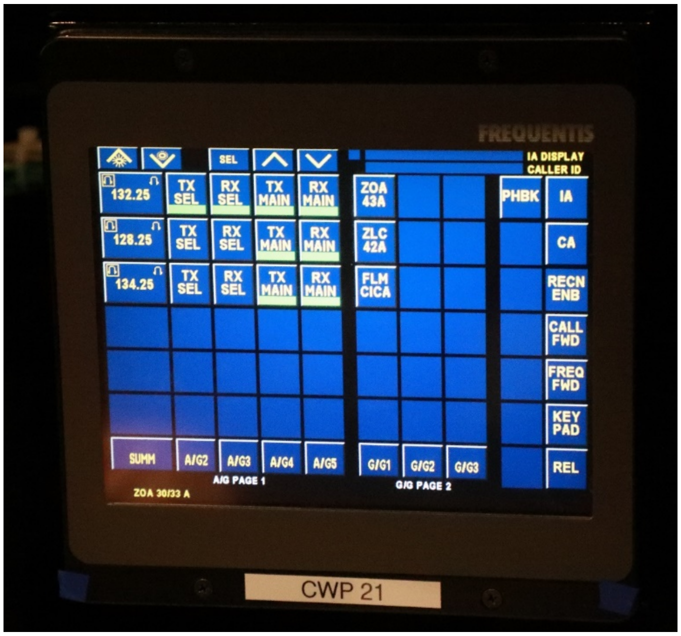


Figure 1. IVSR Communication System

2.3. HARDWARE

Air traffic controller workstations for the participants and ghost sector controllers were located in the NIEC laboratory. The lab configuration is depicted in Figure 2. Although not shown in the figure, researchers placed a partition between Participant A and Ghost A as well as Participant B and Ghost B. The purpose of the barrier was to encourage use of the communications systems between participants and ghost controllers as well as to reduce potential distractions between positions.

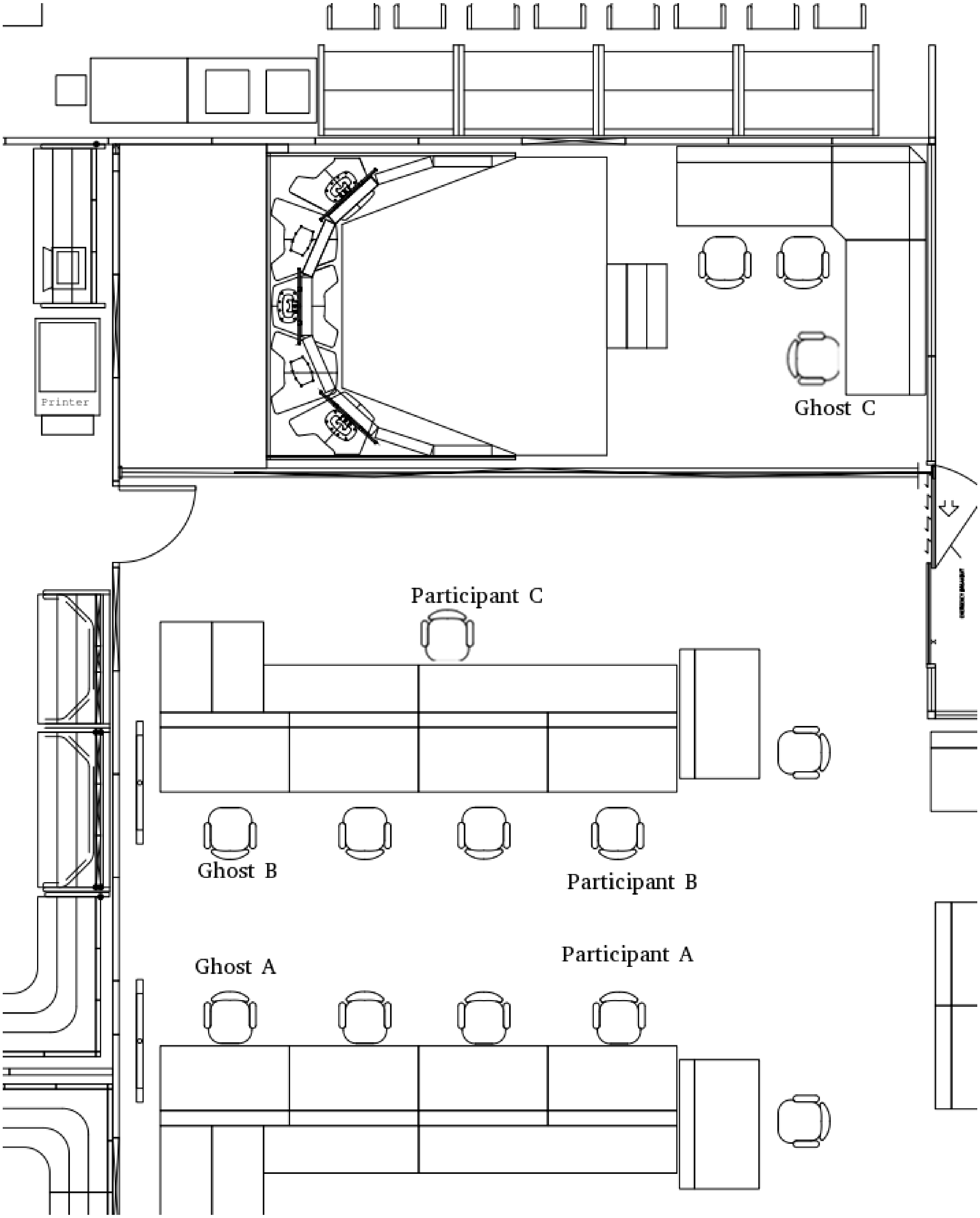


Figure 2. Simulation Participant and Ghost Controller Locations in NIEC Lab.

2.3.1. Air Traffic Control En Route Workstation Consoles

Each ATCS workstation was equipped with a Barco ISIS 2K x 2K Liquid Crystal Display (LCD), a keyboard and trackball, and an IVSR communications system. The Barco LCD was designed for ATCS use and provided the same resolution (2048 x 2048 pixels) and display size (19.83" x 19.83", 28.05" diagonal) as those used in the field. Each console was programmed to emulate ERAM, which is widely used at en route facilities across the country.

Figure 3. ATCS Workstation

2.3.2. User Request Evaluation Tool (URET)

Each ATCS workstation was also equipped with a tool that emulates the User Request Evaluation Tool (URET). URET uses the trajectories of all aircraft to check continuously for conflicts. A potential conflict is declared when the trajectories of two aircraft indicate that the horizontal and vertical separation will both decrease below corresponding thresholds. To determine which sector to notify, URET takes account of which sectors currently control each of the aircraft in a conflict as well as where the actual conflict is predicted to occur. In general, URET will notify the sector where the conflict occurs.

URET presents textual flight information on an Aircraft List (ACL) where one row of information is presented for each aircraft. Aircraft are automatically added to a sector ACL 20 minutes before projected entry into the sector, and they are automatically deleted after leaving the sector. The ATCS can receive alert information on the ACL. There are four relevant boxes at the beginning of each row. The first is a bookkeeping box in which the ATCS can enter a mark to indicate that they have examined the alerts for that aircraft. The other three boxes provide areas to indicate red, yellow, and blue alerts respectively. Conflicts with a predicted minimum horizontal separation of less than or equal to 5 nautical miles (nm) are coded in red. Those with a predicted minimum separation of greater than 5 nm but still within close approach criteria are coded in yellow. When URET determines that an aircraft trajectory will pass through a Special Use Airspace, it generates a blue alert (McFarland, 1997).

The version used in the NIEC lab is actually a prototype of URET (Figure 4) known as JEDI that was developed by MITRE. The NIEC lab adapted this version for use with their simulators and equipment.

✓ ACCEPT
AMENDMENT
JBU993/372

R Y A		Flight Id	PA	Type	Alt	Code	Hdg / Spd	H	Route
■■■1					330				
☐☐☐	189	ASA190	☐	B738/L	380T240	7502	/		KDEN./.TATOO..MONOH.OAKES2.KOAK/1411
☐☐☐	969	SWA5064	☐	B737/L	380	6074	/		KLAS.SHEAD1.KENNO..RUSME.RAZRR4.KSJC/0106
☐☐☐	731	AAL670	☐	B738/L	340	7636	/		KORD./.KU540..RUMPS..OAL..INYOE.DYAMD3.KSFO/0344
☐☐☐	287	NKS305	☐	A321/L	310	0432	/		KSFO./.TIPRE.Q126.EKR..DVV.J80.HCT.J80.JOT..ELX..SVM.
☐☐☐	401	SKW145	☐	CRJ9/L	390	6711	/		KRNO.ZEFFR6.PESKE..LIDAT.J92.BTY.SUNST4.KLAS/0057
☐☐☐	687	UPS9355	☐	SF50/L	270	5713	/		KSFO./.SYRAH.Q128.JSICA..DOTME..JNC..KX40E..BNA..TAZZ
N☐☐☐	439	CPZ5410(U42)	☐	CRJ7/L	320	1147	/		KDEN./.VOAXA.Q136.RUMPS..KNGRY.RAZRR4.KSJC/0200
N☐☐☐	074	SKW279(U42)	☐	CRJ7/L	340	3446	/		KXNA./.ILC.J80.OAL..INYOE.DYAMD3.KSFO/0311
☐☐☐	473	CPZ5518(U42)	☐	CRJ9/L	260	4655	/		KEWR./.KU51K..TATOO..MONOH.OAKES2.KOAK/0459
☐☐☐	394	JBU530(U42)	☐	A319/L	370	7503	/		KOAK./.JSICA..ILC..BCE..RSK..TXO..TURKI.VKTRY2.KDFW/0
☐☐☐	578	JBU368(43)	☐	H/B757/L	370	0630	/		KOAK.HUSSH2.SYRAH.Q128.JSICA..MLF..KD42S..CIM..IRW..L
N☐☐1	284	JBU214(43)	☐	A321/L	370	6632	/		KSFO.SSTIK3.NTELL.Q162.MYCAL..BLD.J72.TXO..TURKI.VKTR
☐☐☐	987	DAL990(43)	☐	A320/L	310	6310	/		KOAK./.NTELL.Q162.ESSAA..BTY.SUNST4.KLAS/0119

Figure 4. URET Prototype

2.3.3. Ghost Sector Air Traffic Control Workstations

Three ghost sector workstations were required during the conduct of the simulation. Each workstation simulated sectors adjacent to the participants' (ZOA 43 and ZLC 42). Like the other ATCS workstations, they included a monitor, keyboard, and trackball. Workstations also included audio communications, the ability to monitor simulated frequencies, and the ability to communicate directly with participants.

2.3.4. Simulation Pilot Workstations

Simulation pilot workstations were located in the TGF sim pilot laboratory, separate from the NIEC laboratory. Each workstation consisted of a computer, keyboard, monitor, and communication equipment. Simulation pilots controlled TGF-generated aircraft using commands on their respective workstations. These commands are typically comprised of pre-defined strings of alphanumeric characters that pilots entered using a standard keyboard.

Each simulation pilot workstation provided a plan view (2-D) display of traffic and a list of assigned aircraft. For each assigned aircraft, the workstation provided information regarding the current state and flight plan data.

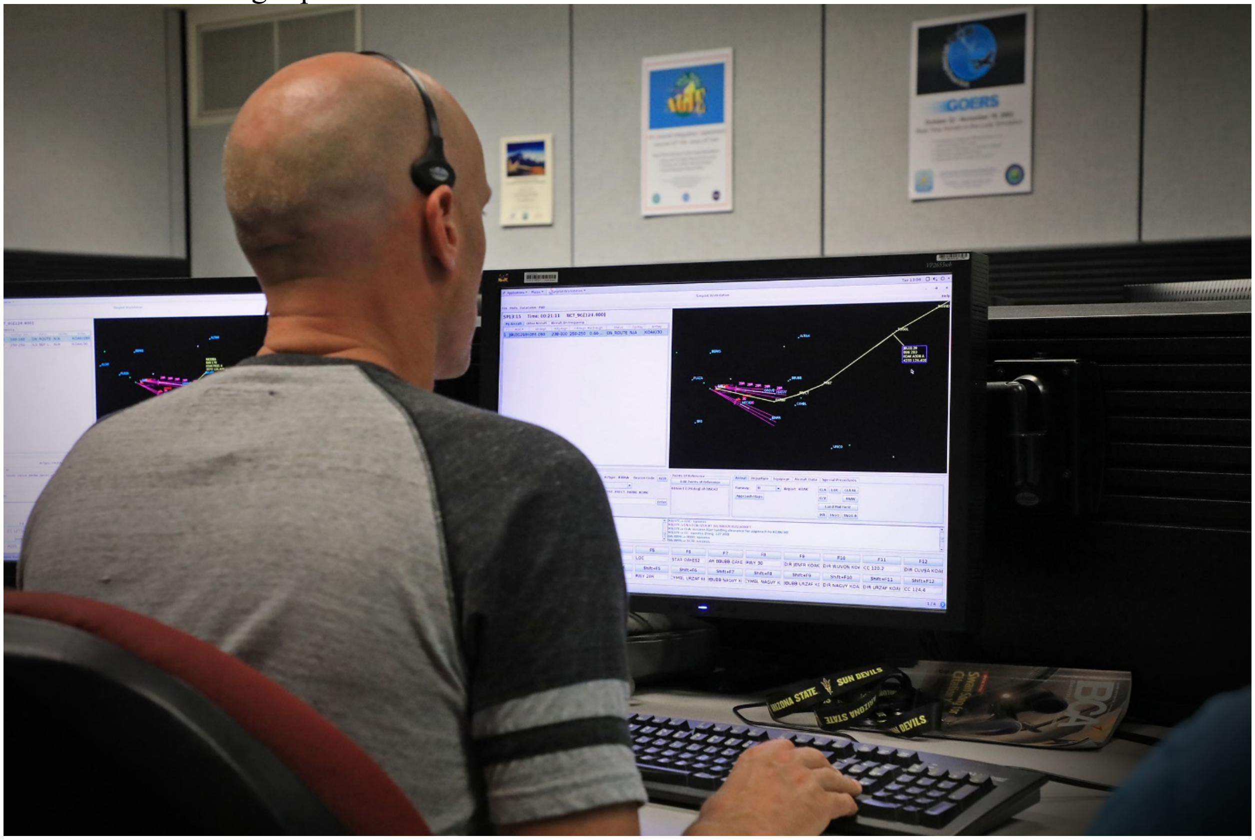


Figure 5. TGF Sim Pilot Workstations

2.4. SOFTWARE

The simulation utilized the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) and TGF simulation engines that are key in conducting air traffic simulations. The TGF is a crosscutting infrastructure that drives terminal, en route, and other research laboratories. DESIREE and TGF work together to immerse the research participant into a realistic environment that can emulate air traffic environments.

DESIREE consists of a series of interchangeable human-machine interfaces. It has the capability to emulate multiple ATC platforms and displays. Its purpose is to enable researchers to modify or add information and functionality to a variety of current ATC environments to allow for the evaluation of new concepts and procedures. DESIREE receives input from TGF that allows it to present information on a radar display (e.g., Standard Terminal Automation Replacement System (STARS), Display System Replacement (DSR), and ERAM), including radar tracks, flight data block (FDB), and sector maps. It also allows ATCS to perform the typical functions that they would perform in the appropriate ATC operational environment (e.g., performing handoffs, entering data into the host computer). DESIREE can also emulate ghost sector operations by providing automation to control these unstaffed sectors. When needed this automation can communicate with TGF to act as a simulation pilot for the aircraft. DESIREE has data collection capabilities and can collect information on all ATCS entries made during a simulation run.

TGF was designed to generate realistic aircraft trajectories and associated digital radar messages for aircraft in a simulated airspace environment. It uses preset flight plans and dynamic flight models to generate simulated surveillance tracks. The TGF algorithms control aircraft maneuvers so that they represent realistic climb, descent, and turn rates. In addition, TGF records information about aircraft trajectories, proximity, and other relevant data, which researchers use in subsequent analyses. Real-time information on each aircraft was output in DIS format to both the DESIREE displays and the Dynamic DSP Simulator (DDS) applications.

Use of a TGF aircraft model in simulation can sometimes be less costly than using simulators; however, it is a lower fidelity method of integrating a UAS into simulation. TGF models of UAS were either created in-house by TGF developers based on data from the UAS lab simulators or were based on NASA Base of Aircraft Data (BADA) models. In cases where a TGF model did not exist, we selected a surrogate aircraft model of a manned aircraft with flight characteristics that closely matched the desired flight characteristics of a particular UAS. TGF models emulated how a UAS flies (e.g., speed, turn rates) but not necessarily how it behaves. The responsiveness of an actual UAS and the TGF models differ; however, for the purposes of shakedown and simulation testing we scripted the scenarios in a manner such that we could fix the behavior of the UAS.

2.5. ASSUMPTIONS AND CONSTRAINTS

The FAA currently restricts how and where UAS can be operated; therefore, rules and procedures for operating outside Warning Areas or Restricted Airspace were defined according to current FAA regulations. Assumptions for this simulation were as follows:

- UAS squawked 7400 when a lost link occurred.
- In addition to loss of C2 link, two-way communications may be lost as well
- Participants were provided the following information prior to the start of a scenario: T1, T3, method of lost link delivery in the event of lost communications
- Lost link routings were provided to participants at a specified time after 7400 squawk (T1) via voice communications from PIC or if lost communications via manual (handout

delivered by FLM/CIC) or automated method (ABRR-like). UA platforms were programmed to comply with lost link routing accordingly

- Performance characteristics (e.g., altitude limitations, climb/descent rates) were known to participants
- A UAS lost link condition was not considered an emergency for the purposes of this simulation, however, participants were to advise appropriate adjacent sectors and/or FLM/CIC as necessary
- All FDX and UPS aircraft were UAS only. There were two types of UAS in the scenario and they had a ‘U’ prefix for the aircraft type and displayed on the fourth line of the flight data block. UPS were a single engine jet and FDX were a single engine turbo-prop
- UAS were deemed airworthy and were operating on an IFR flight plan
- Communication between participants and flight crew was by voice
- UAS PIC and flight crew were appropriately credentialed by the FAA
- UA was transponder and/or ADS-B equipped

2.6. AIRSPACE

The HITL utilized Sectors 30 and 33 of the Oakland Air Route Traffic Control Center (ZOA) airspace. These sectors were combined and slightly modified for the purposes of the simulation. Sectors 30 and 33 are the eastern most sectors in ZOA airspace with FL240 to FL600. These sectors are a major crossroad for air traffic in the western United States that arrive and depart the San Francisco Bay area and Sacramento Complex.

The simplification of airspace was necessary to reduce the training time needed for participants to become familiar with the operation in preparation for the simulation.

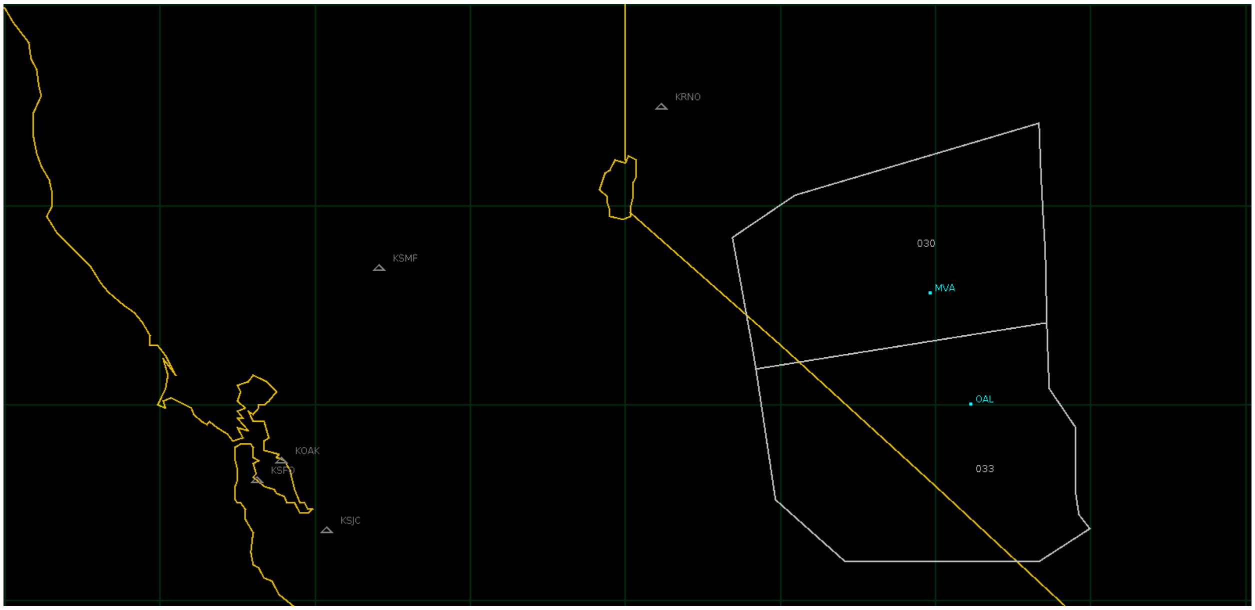


Figure 6. ZOA Airspace Sectors 30/33

2.7. UAS PLATFORMS

Scenarios used two conceptual UAS platforms similar to known airframes:

- Single engine turbo prop; Cessna 208-like with a service ceiling of FL250 and cruise speed of 214 knots
- Single engine jet; Cirrus Vision Jet-like with a service ceiling of FL280 and a cruise speed of 330 knots

For the purposes of the simulation, we referred to each UAS as being ‘like’ a known aircraft as they had flight characteristics similar to those known platforms; however, the simulation did not evaluate specific aircraft.

As UAS operations in the NAS become normalized, the ability to differentiate between manned aircraft and UAS will become increasingly important to ATCS. In this simulation, all aircraft with FDX and UPS call signs were unmanned aircraft; therefore, participants knew that the call sign was one way to identify the aircraft as a UAS. Additionally, the aircraft type was available on the fourth line of the ERAM flight data block.

Background traffic included 10-20% unmanned aircraft with the rest of the traffic consisting of manned aircraft.

2.8. MATERIALS

2.8.1. Informed Consent

Each participant read and signed an informed consent statement (Appendix A-1) before beginning the experiment. The informed consent statement described the study, the foreseeable risks, and the rights and responsibilities of the participants, including that their participation was voluntary. We protected all the information that the participant provided, including personally identifiable information (PII), from release except as may be required by statute. Signing the form indicated that the participant understood their rights as a participant in the study and gave their consent to participate.

The task lead offered a copy of FAA Order 1280.1B, “Protecting Personally Identifiable Information”, to any participant who requested it as well as to answer any questions concerning that order.

2.8.2. Data Collection Instruments

We collected objective and subjective data during the simulation. Objective system data from TGF and DESIREE included safety and efficiency metrics such as the number of operational errors and the number of delays to manned aircraft. These measures provided information regarding what occurred during the different test conditions. The simulation was also video and audio recorded so that the research team could review the scenarios as needed during data analysis.

2.8.3. Background Questionnaire

We collected biographical data from participants via tablet. Before beginning the simulation participants completed a Background Questionnaire (Appendix A-2). The Background Questionnaire included items about age, gender, experience, and attitudes about participating.

2.8.4. Post-Scenario Questionnaire

At the end of each scenario, participants completed a Post-Scenario Questionnaire (PSQ; Appendix A-3), also via tablet, that asked them to provide subjective ratings regarding their performance, workload, and reactions to the test conditions (e.g., easy/difficult). Workload ratings in particular were assessed using a standard 10-point scale where 1 referred to very low workload and 10 referred to very high workload. The 10-point workload scale has been validated on numerous occasions in aviation research (Roscoe & Ellis, 1990; Stein & Rosenberg, 1983; Parasuraman, Sheridan, & Wickens, 2008).

The first six questions on the PSQ consisted of the NASA Task Load Index (TLX). The TLX is a subjective workload assessment tool that allows users to perform subjective workload assessments when working with various human-machine interface systems. Hart and Staveland (1988) originally developed the TLX for the NASA Ames Research Center. The NASA TLX has become the gold standard for measuring subjective workload across a wide range of applications. The NASA TLX asks users to rate their experience on six dimensions:

- Mental Demand
- Physical Demand
- Temporal Demand
- Performance
- Effort
- Frustration

The NASA TLX has been successfully used around the world to assess workload in various environments such as aircraft cockpits; command, control, and communication (C3) workstations; supervisory and process control; and simulations and laboratory tests (So, 2018).

2.8.5. Exit Questionnaire

At the end of the simulation, participants completed an Exit Questionnaire (Appendix A-4) which asked participants to provide ratings about the realism of the ATCS displays, effectiveness of pre-simulation training, and performance compared to actual operations using 10-point scales. It also prompted the participants to provide additional responses about their experience in the simulation and their impressions regarding overall effects on safety, efficiency, workload, and other measures across all of the test conditions.

2.8.6. Workload Assessment Keypad

Participants also provided workload ratings in real time via the WAK (Figure 7). A WAK was present at each ATCS position (excluding the ghost positions) and data was collected and timestamped by DESIREE. The WAK consisted of a touch panel display with ten numbered buttons. At three-minute intervals, the WAK prompted the participants to press a button to provide their subjective workload rating. At the onset of the prompt the WAK emitted a brief tone and the background screen flashed yellow for 30 seconds or until the participant responded. The participant indicated their current workload by pressing one of the numbered buttons.

Figure 7. Workload Assessment Keypad (WAK)

2.9. PRE-TESTING

After thoroughly testing all the scenarios and equipment in the weeks prior, we conducted a validation session (referred to as shakedown) over three days. During shakedown, researchers and lab personnel worked through each of the test conditions and scenarios and collected the data from the system as planned for the simulation. Lab personnel also ensured that the system hardware and software were working correctly and identified any issues or problems for remediation and reevaluation prior to the simulation. During this time, we also ensured all data collection devices were functioning as required and that data files contained the information necessary for later analysis. Shakedown also provided an opportunity to confirm that the chosen T1 and T2 times were appropriate.

3. EXPERIMENTAL DESIGN

3.1 SCENARIOS

This study included 13 scenarios in which lost links occurred (one of them a ‘continue on flight plan’ baseline condition) plus one no-lost-link baseline scenario.

Table 1. Experimental Design

Condition	T1 Lag	T2 Lag	T3 Total Time	Delivery Mechanism	Two-way Communications	Route Length
1	Short (1 min)	Short (2 min)	3 min	Manual	Lost	Fix Based 7 Elements
2				Automated	Lost	
3				Voice	Present	
4	Short (1 min)	Long (4 min)	5 min	Manual	Lost	
5				Automated	Lost	
6				Voice	Present	
7	Long (4 min)	Short (2 min)	6 min	Manual	Lost	
8				Automated	Lost	
9				Voice	Present	
10	Long (4 min)	Long (4 min)	8 min	Manual	Lost	
11				Automated	Lost	
12				Voice	Present	
13 Continue	1 or 4 minutes	No turn	No turn	Voice	Present	Continue
14 Baseline	No lost link	NA	NA	NA	NA	NA

To clarify, Figure 8 shows that T1 is the time from when the lost link UAS squawks 7400 to when the UAS intent is provided to participants. T2 is defined as the time from when UAS intent information is provided to participants to the aircraft's first maneuver off the original flight plan. T3 is the total time (T1 + T2) from the 7400 squawk until the aircraft maneuvers off its original flight plan.

Once the T1 time expired, participants were provided with the UAS intent, which consisted of a seven-element amendment to the original route of flight. The amendment included both fixes and airways but did not include any altitude change. The UAS intent was delivered either manually, by automation, or by voice, which is further described below.

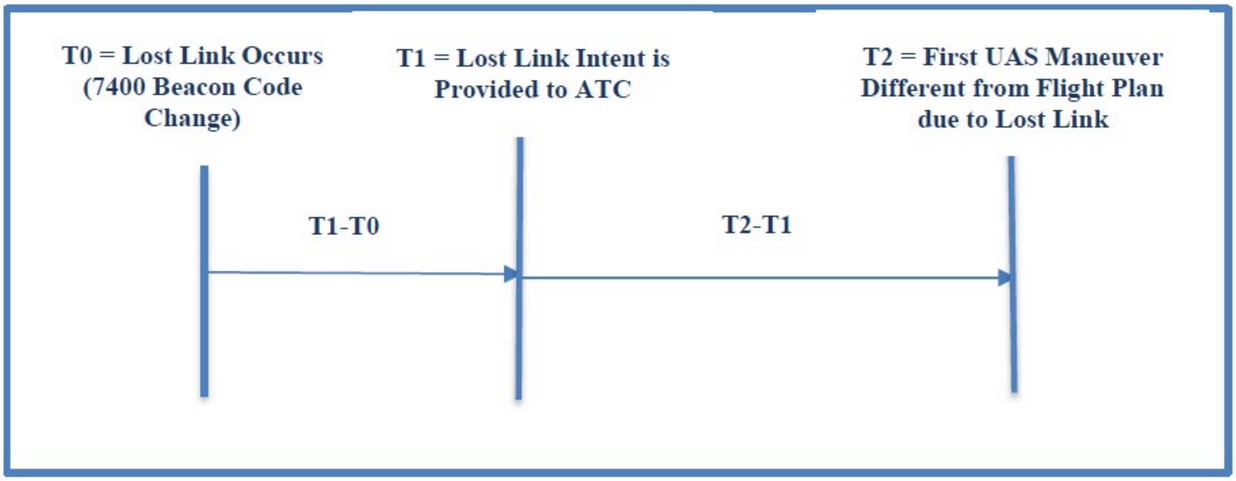


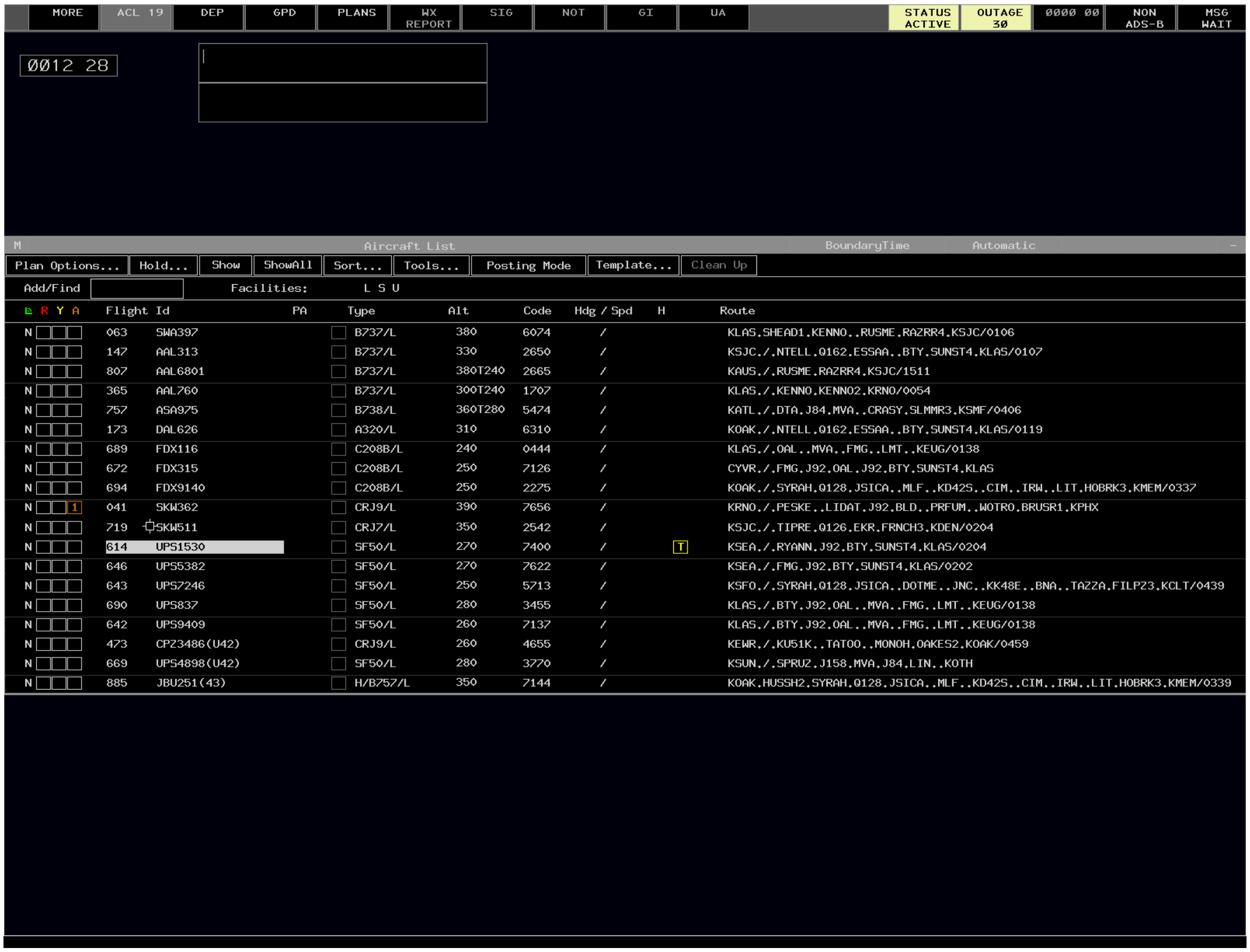
Figure 8. Timeline of Lost Link Events

3.1.1 Manual Delivery Method

The manual delivery method consisted of the FLM/CIC providing the participant with a written copy of the UAS lost link routing. Prior to the study, we briefed that the FLM/CIC may have received this information via landline with the UAS PIC or by some other unspecified method. Two-way communication with the lost link UAS was not available during this delivery method.

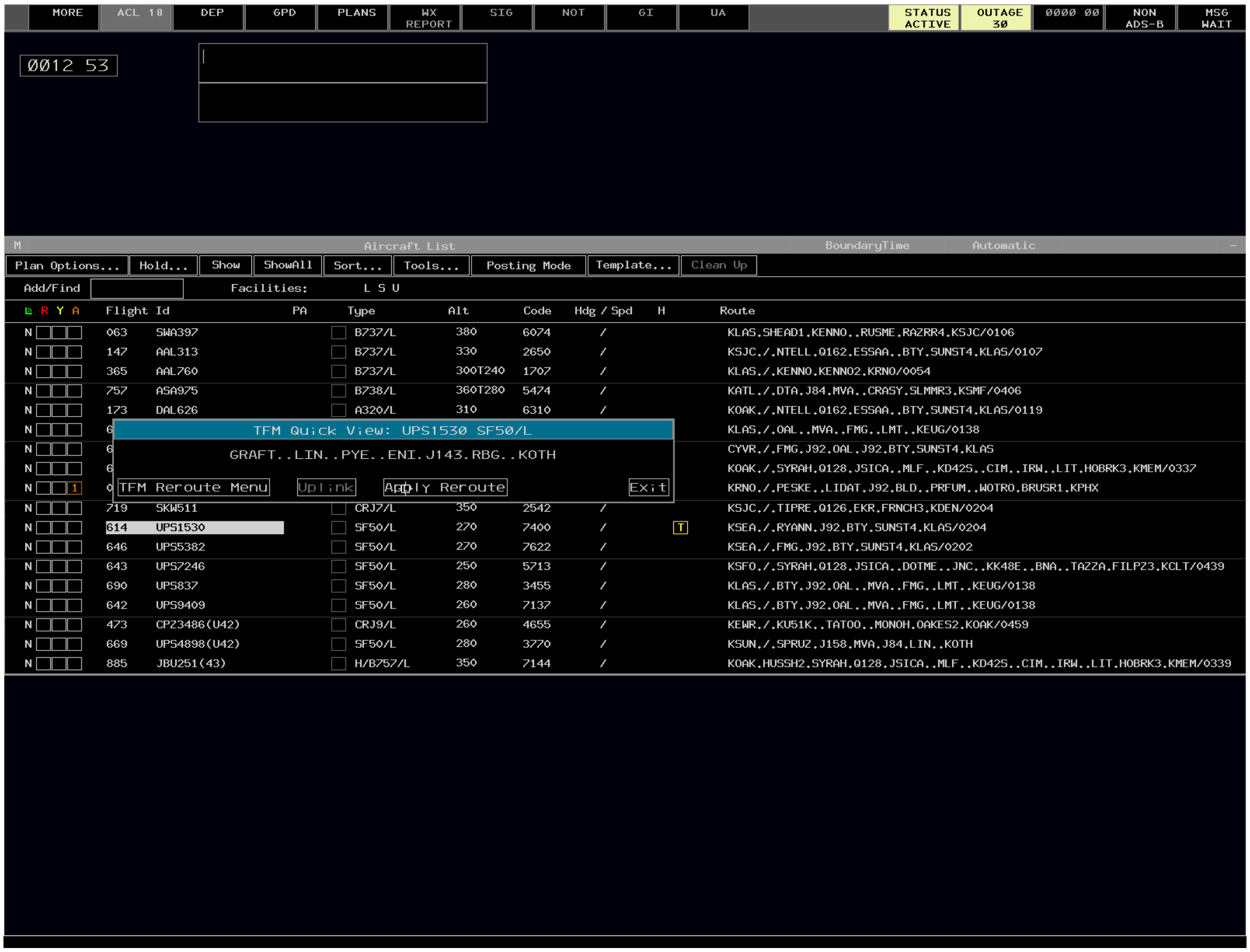
3.1.2. Automated Delivery Method

The automated delivery method consisted of the UAS lost link routing being pushed to the participant similar to the current ABRR functionality of ERAM. After the expiration of the T1 time, the UAS lost link routing was pushed to the ACL on the D-side position and depicted with a boxed ‘T’ signifying new routing was available (Figure 9, top). The R-side position depicted a ‘T’ next to the aircraft target as well. Once the participant clicked on the ‘T’ on the ACL, the UAS lost link routing was depicted with an option to ‘apply reroute’ or ‘exit’ (Figure 9, middle). Once the new route was successfully applied, an accept message was received (Figure 9, bottom).



Facilities: L S U

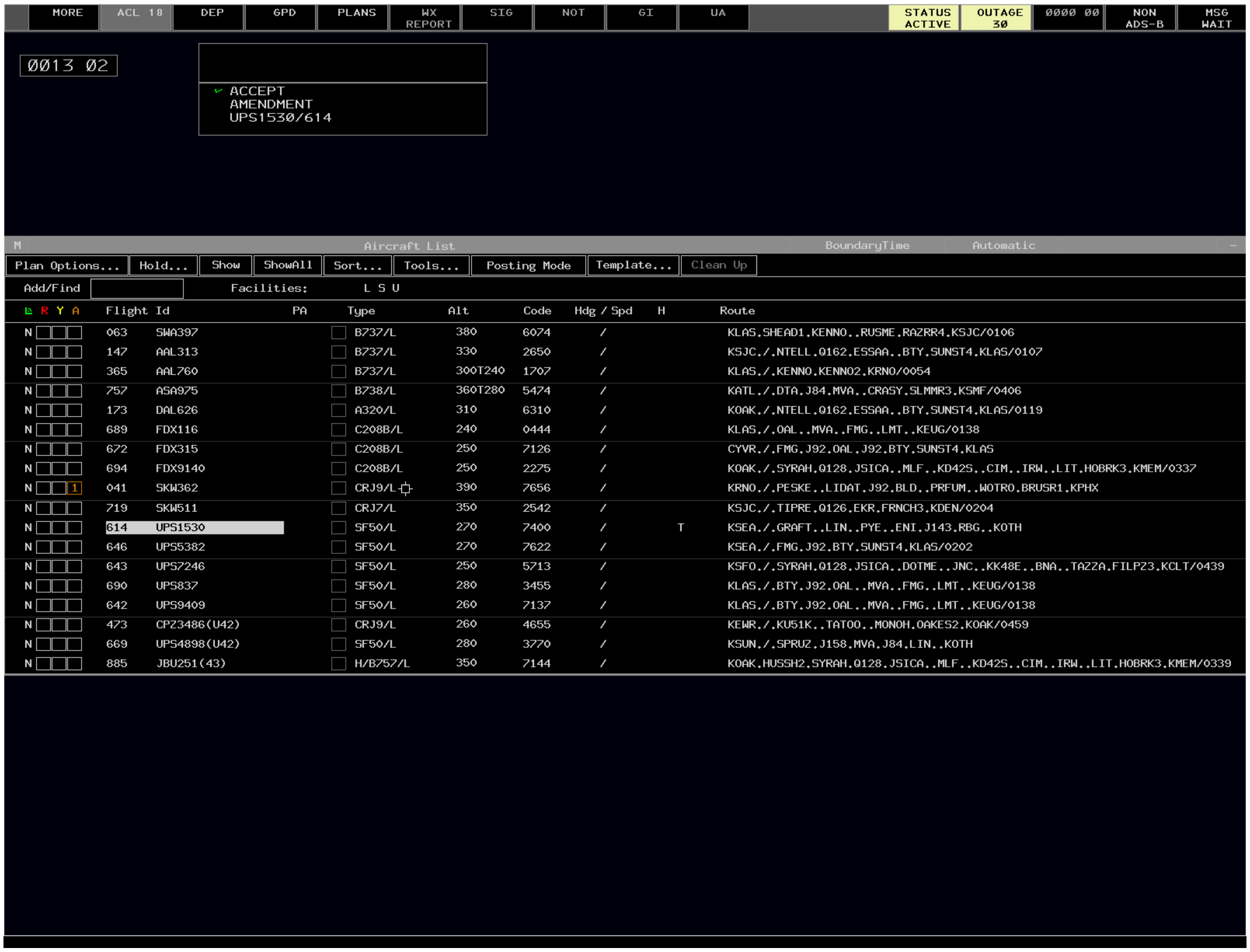
R Y A	Flight Id		PA	Type	Alt	Code	Hdg / Spd	H	Route
N ☐☐☐	063	SWA397		☐ B737/L	380	6074	/		KLAS.SHEAD1.KENNO..RUSME.RAZRR4.KSJC/0106
N ☐☐☐	147	AAL313		☐ B737/L	330	2650	/		KSJC./.NTELL.Q162.ESSAA..BTY.SUNST4.KLAS/0107
N ☐☐☐	807	AAL6801		☐ B737/L	380T240	2665	/		KAUS./.RUSME.RAZRR4.KSJC/1511
N ☐☐☐	365	AAL760		☐ B737/L	300T240	1707	/		KLAS./.KENNO.KENNO2.KRNO/0054
N ☐☐☐	757	ASA975		☐ B738/L	360T280	5474	/		KATL./.DTA.J84.MVA..CRASY.SLMMR3.KSMF/0406
N ☐☐☐	173	DAL626		☐ A320/L	310	6310	/		KOAK./.NTELL.Q162.ESSAA..BTY.SUNST4.KLAS/0119
N ☐☐☐	689	FDX116		☐ C208B/L	240	0444	/		KLAS./.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	672	FDX315		☐ C208B/L	250	7126	/		CYVR./.FMG.J92.OAL.J92.BTY.SUNST4.KLAS
N ☐☐☐	694	FDX9140		☐ C208B/L	250	2275	/		KOAK./.SYRAH.Q128.JSICA..MLF..KD42S..CIM..IRW..LIT.HOBRK3.KMEM/0337
N ☐☐1	041	SKW362		☐ CRJ9/L	390	7656	/		KRNO./.PESKE..LIDAT.J92.BLD..PRFUM..WOTRO.BRUSR1.KPHX
N ☐☐☐	719	SKW511		☐ CRJ7/L	350	2542	/		KSJC./.TIPRE.Q126.EKR.FRNCH3.KDEN/0204
N ☐☐☐	614	UPS1530		☐ SF50/L	270	7400	/	T	KSEA./.RYANN.J92.BTY.SUNST4.KLAS/0204
N ☐☐☐	646	UPS5382		☐ SF50/L	270	7622	/		KSEA./.FMG.J92.BTY.SUNST4.KLAS/0202
N ☐☐☐	643	UPS7246		☐ SF50/L	250	5713	/		KSFO./.SYRAH.Q128.JSICA..DOTME..JNC..KK48E..BNA..TAZZA.FILPZ3.KCLT/0439
N ☐☐☐	690	UPS837		☐ SF50/L	280	3455	/		KLAS./.BTY.J92.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	642	UPS9409		☐ SF50/L	260	7137	/		KLAS./.BTY.J92.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	473	CPZ3486(U42)		☐ CRJ9/L	260	4655	/		KEWR./.KU51K..TATOO..MONOH.OAKES2.KOAK/0459
N ☐☐☐	669	UPS4898(U42)		☐ SF50/L	280	3770	/		KSUN./.SPRUZ.J158.MVA.J84.LIN..KOTH
N ☐☐☐	885	JBU251(43)		☐ H/B757/L	350	7144	/		KOAK.HUSSH2.SYRAH.Q128.JSICA..MLF..KD42S..CIM..IRW..LIT.HOBRK3.KMEM/0339



TFM Quick View: UPS1530 SF50/L

GRAFT..LIN..PYE..ENI.J143.RBG..KOTH

R Y A	Flight Id		PA	Type	Alt	Code	Hdg / Spd	H	Route
N ☐☐☐	063	SWA397	☐	B737/L	380	6074	/		KLAS.SHEAD1.KENNO..RUSME.RAZRR4.KSJC/0106
N ☐☐☐	147	AAL313	☐	B737/L	330	2650	/		KSJC./.NTELL.Q162.ESSAA..BTY.SUNST4.KLAS/0107
N ☐☐☐	365	AAL760	☐	B737/L	300T240	1707	/		KLAS./.KENNO.KENNO2.KRNO/0054
N ☐☐☐	757	ASA975	☐	B738/L	360T280	5474	/		KATL./.DTA.J84.MVA..CRASY.SLMMR3.KSMF/0406
N ☐☐☐	173	DAL626	☐	A320/L	310	6310	/		KOAK./.NTELL.Q162.ESSAA..BTY.SUNST4.KLAS/0119
N ☐☐☐	6								KLAS./.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	6								CYVR./.FMG.J92.OAL.J92.BTY.SUNST4.KLAS
N ☐☐☐	6								KOAK./.SYRAH.Q128.JSICA..MLF..KD42S..CIM..IRW..LIT.HOBRK3.KMEM/0337
N ☐☐1	0								KRNO./.PESKE..LIDAT.J92.BLD..PRFUM..WOTRO.BRUSR1.KPHX
N ☐☐☐	719	SKW511	☐	CRJ7/L	350	2542	/		KSJC./.TIPRE.Q126.EKR.FRNCH3.KDEN/0204
N ☐☐☐	614	UPS1530	☐	SF50/L	270	7400	/	T	KSEA./.RYANN.J92.BTY.SUNST4.KLAS/0204
N ☐☐☐	646	UPS5382	☐	SF50/L	270	7622	/		KSEA./.FMG.J92.BTY.SUNST4.KLAS/0202
N ☐☐☐	643	UPS7246	☐	SF50/L	250	5713	/		KSFO./.SYRAH.Q128.JSICA..DOTME..JNC..KK48E..BNA..TAZZA.FILPZ3.KCLT/0439
N ☐☐☐	690	UPS837	☐	SF50/L	280	3455	/		KLAS./.BTY.J92.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	642	UPS9409	☐	SF50/L	260	7137	/		KLAS./.BTY.J92.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	473	CPZ3486(U42)	☐	CRJ9/L	260	4655	/		KEWR./.KU51K..TATOO..MONOH.OAKES2.KOAK/0459
N ☐☐☐	669	UPS4898(U42)	☐	SF50/L	280	3770	/		KSUN./.SPRUZ.J158.MVA.J84.LIN..KOTH
N ☐☐☐	885	JBU251(43)	☐	H/B757/L	350	7144	/		KOAK.HUSSH2.SYRAH.Q128.JSICA..MLF..KD42S..CIM..IRW..LIT.HOBRK3.KMEM/0339



0013 02

✓ ACCEPT
AMENDMENT
UPS1530/614

Facilities: L S U

R Y A	Flight Id		PA	Type	Alt	Code	Hdg / Spd	H	Route
N ☐☐☐	063	SWA397	☐	B737/L	380	6074	/		KLAS.SHEAD1.KENNO..RUSME.RAZRR4.KSJC/0106
N ☐☐☐	147	AAL313	☐	B737/L	330	2650	/		KSJC./.NTELL.Q162.ESSAA..BTY.SUNST4.KLAS/0107
N ☐☐☐	365	AAL760	☐	B737/L	300T240	1707	/		KLAS./.KENNO.KENNO2.KRNO/0054
N ☐☐☐	757	ASA975	☐	B738/L	360T280	5474	/		KATL./.DTA.J84.MVA..CRASY.SLMMR3.KSMF/0406
N ☐☐☐	173	DAL626	☐	A320/L	310	6310	/		KOAK./.NTELL.Q162.ESSAA..BTY.SUNST4.KLAS/0119
N ☐☐☐	689	FDX116	☐	C208B/L	240	0444	/		KLAS./.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	672	FDX315	☐	C208B/L	250	7126	/		CYVR./.FMG.J92.OAL.J92.BTY.SUNST4.KLAS
N ☐☐☐	694	FDX9140	☐	C208B/L	250	2275	/		KOAK./.SYRAH.Q128.JSICA..MLF..KD42S..CIM..IRW..LIT.HOBRK3.KMEM/0337
N ☐☐1	041	SKW362	☐	CRJ9/L	390	7656	/		KRNO./.PESKE..LIDAT.J92.BLD..PRFUM..WOTRO.BRUSR1.KPHX
N ☐☐☐	719	SKW511	☐	CRJ7/L	350	2542	/		KSJC./.TIPRE.Q126.EKR.FRNCH3.KDEN/0204
N ☐☐☐	614	UPS1530	☐	SF50/L	270	7400	/	T	KSEA./.GRAFT..LIN..PYE..ENI.J143.RBG..KOTH
N ☐☐☐	646	UPS5382	☐	SF50/L	270	7622	/		KSEA./.FMG.J92.BTY.SUNST4.KLAS/0202
N ☐☐☐	643	UPS7246	☐	SF50/L	250	5713	/		KSFO./.SYRAH.Q128.JSICA..DOTME..JNC..KK48E..BNA..TAZZA.FILPZ3.KCLT/0439
N ☐☐☐	690	UPS837	☐	SF50/L	280	3455	/		KLAS./.BTY.J92.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	642	UPS9409	☐	SF50/L	260	7137	/		KLAS./.BTY.J92.OAL..MVA..FMG..LMT..KEUG/0138
N ☐☐☐	473	CPZ3486(U42)	☐	CRJ9/L	260	4655	/		KEWR./.KU51K..TATOO..MONOH.OAKES2.KOAK/0459
N ☐☐☐	669	UPS4898(U42)	☐	SF50/L	280	3770	/		KSUN./.SPRUZ.J158.MVA.J84.LIN..KOTH
N ☐☐☐	885	JBU251(43)	☐	H/B757/L	350	7144	/		KOAK.HUSSH2.SYRAH.Q128.JSICA..MLF..KD42S..CIM..IRW..LIT.HOBRK3.KMEM/0339

Figure 9. Automated lost link delivery. See description in text for explanation of screens.

3.1.3. Voice Delivery Method

The voice delivery method consisted of the UAS PIC providing the participant with the lost link routing at the expiration of the T1 time. Two-way communication with the lost link UAS was maintained throughout the duration of the scenario.

3.1.4. Starting Point for UAS Lost Link Routing

During the planning phase of the simulation, we considered two options for a starting point on the UAS lost link route: fix-based and time-based.

The fix-based option assumed that a UAS experiencing lost link would make the first turn off the original route at a specified fix on the existing route. This would provide the ATCS with a predictable, unambiguous starting point for the lost link route. The known starting point would allow a flight plan amendment to be made immediately at the discretion of the ATCS. The flight plan amendment would provide accurate and automated notification to impacted adjacent sectors of the incoming lost link UAS. The fix-based option would also allow UAS operators to appropriately plan for avoidance of any special use airspace potentially impacted by a lost link event.

The alternative was a time-based approach, which assumed that a UAS experiencing lost link would make the first turn off the original route at the expiration of a specific amount of time after squawking 7400. This option would require some type of timer to alert the ATCS of when the lost link UAS would start the first turn off the original route. The exact point for the first turn is somewhat ambiguous and could require several flight plan amendments to accurately reflect the programmed lost link trajectory. This ambiguity has the potential to negatively affect the automated notification to impacted adjacent sectors. The time-based option could also prove challenging for UAS operators to ensure avoidance of special use airspace (or other potential hazards) when programming lost link routing.

After thorough consideration, researchers selected the fix-based option to be used in the simulation. The 7400 squawk times were manipulated to ensure the T1/T2 test conditions for the experimental design were achieved.

3.1.5. Scenario Run Order

During the simulation, two or three scenarios (one for each participant available each week) were run simultaneously. Scenarios were counterbalanced so that although some participants may have run different scenarios at the same time, at the end of the day the participants had completed the same scenarios. This allowed participants to discuss the same scenarios during the end of the day debrief while still counterbalancing for scenario order. In circumstances where participants did not run the same scenarios during a day (such as when a participant had to be rescheduled), only the scenarios completed by all participants were discussed.

3.2. INDEPENDENT VARIABLES

We examined the following independent variables in this HITL: T1 time, T2 time, lost link intent delivery method, and availability of communications. We also examined UAS platform (jet versus turbo-prop) but did not expect, nor find, any significant effects based on previous research.

3.3. DEPENDENT VARIABLES

We collected the following dependent measures during the simulation:

1. **Safety Measures**
   Number of losses of separation involving manned and UAS aircraft
   Duration of loss of separation
   Closest point of approach during loss of separation

2. **Efficiency Measures**
   Time and distance flown in sector

3. **Communication Measures**
   Number of ATCS-pilot communications
   Duration of communications

4. **ATCS Workload Measures**
   Number of sim pilot commands entered
   WAK
   NASA-Task Load Index
   Amendment entry times

5. **ATCS Situation Awareness Measures**
   Subjective situation awareness on PSQ

6. **ATCS Preference Measures**
   Subjective preferences on various topics on PSQ

3.4. PROCEDURE

3.4.1. Schedule of Events/Timeline

The table below is an example of the weekly schedule of simulation scenarios.

Table 2. Weekly Schedule of Events

Time	Day 1 Tuesday	Time	Day 2 Wednesday	Time	Day 3 Thursday
8:00	Welcome, In Brief	8:00	Scenario 3	8:00	Scenario 10
8:30		8:30		8:30	
9:00		9:00	Scenario 4	9:00	Scenario 11
9:30		9:30		9:30	
10:00	Practice Scenario 1	10:00	Scenario 5	10:00	Scenario 12
10:30		10:30		10:30	
11:00	Practice Scenario 2	11:00	Scenario 6	11:00	Scenario 13
11:30		11:30		11:30	
12:00	LUNCH	12:00	LUNCH	12:00	LUNCH
12:30		12:30		12:30	
1:00	Practice Scenario 3	1:00	Scenario 7	1:00	Scenario 14
1:30		1:30		1:30	
2:00	Practice Scenario 4	2:00	Scenario 8	2:00	End of Simulation Debrief
2:30		2:30		2:30	
3:00	Scenario 1	3:00	Scenario 9	3:00	
3:30		3:30		3:30	
4:00	Scenario 2	4:00	Daily Debrief	4:00	
4:30		4:30		4:30	

3.4.2. In-Brief

On the first day of the study, the research task lead introduced the research team to participants and then briefed them on the background and objectives of the study. An ATCS SME then gave a brief overview of the airspace, procedures, and simulated, laboratory environment the participants would experience. After the briefings the participants completed an informed consent form along with a background questionnaire (see Section 2.8).

3.4.3. Training/Practice Scenarios

Following the in-brief, we gave participants an overview of the laboratory equipment. The participants then controlled traffic in practice scenarios as ATCS SMEs observed and were on standby to answer questions. The practice runs served to familiarize participants with the test environment, interactions with the simulation pilots, and usage of the WAK. Practice scenarios used the same airspace as the experimental scenarios and featured traffic in the same patterns as the experimental scenarios; however, the traffic density was lower. We confirmed with participants that they were sufficiently prepared to begin data collection runs after completing the practice scenarios.

3.4.4. Data Collection Procedure

The simulation ran over a four-week period. Each week consisted of two travel days for the participants and three days of training and simulation. The HITL contained 14 data collection scenarios over the course of the three days. Each scenario lasted approximately 30 minutes followed by a break for the participants.

Before each scenario, researchers provided the participants with a ‘cheat sheet’ indicating the T1 time, T3 time, and potential lost-communications delivery mechanism (e.g., ‘if a lost link occurs, the lost link route will be provided by the pilot or ABRR’) for the upcoming scenario. Pilots therefore knew the T1 and T3 times for the scenario but did not know whether communications would be present or absent or if a lost link would occur. They also knew what delivery method would be used in the event of lost two-way communications. This arrangement was meant to allow each scenario to simulate a potential future state of the NAS where T1 time, T2 time, and delivery mechanism are standardized.

We collected data in the form of audio and video recordings of ATCS, telemetry data from TGF, questionnaires, workload assessment via the WAK, and other forms of operational and performance data via DESIREE.

4. DATA ANALYSIS METHODOLOGY

We used Bayesian methods to model the impact of the independent variables on the dependent variables (see Sections 3.2 and 3.3). Each dependent variable was modeled separately using a Bayesian multilevel generalized linear regression (see Appendix F for a detailed description of the models). The one exception was loss of separation data, which included the number of losses, the point of closest approach during a loss, and duration of loss of separation. After SME review, only three loss of separation events occurred during the simulation, which does not lend itself to statistical analysis.

4.1. HOW RESULTS ARE REPORTED

Most readers are likely more familiar with frequentist statistical methods that result in a p-value and sometimes a confidence interval (CI). The p-value is the probability of a statistical result given a null hypothesis. The CI (commonly set to 95%, thus termed the 95% CI) gives the probability that a given CI would include the real value of, say, a mean if the experiment were repeated many, many times and the CI calculated for each. A statistical result is commonly termed 'significant' if the p-value is less than .05. The CI and level of significance for the p-value are related such that if the p-value is less than .05, the 95% CI will also exclude zero. In contrast, Bayesian methods do not use a p-value and instead of a CI use what is called a high-density interval (HDI). The HDI is interpreted in the way that many people intuitively think a CI should be (but is not) interpreted: a 95% HDI tells us the range of values that we can be 95% certain the actual value falls in.

Following the typical convention but with Bayesian statistics, this report uses 95% HDIs and refers to comparisons as 'significant' if the 95% HDI excludes zero. It also reports a single-number 'confidence' in a comparison, similar to a p-value, by examining how much of the Bayesian posterior probability falls to one side or the other of a comparison point. For example, say that the distance flown by aircraft in condition A is compared to the distance flown in condition B. The 95% HDI may find that the difference between conditions ranges from 0.5 to 3.1 nautical miles (nm; e.g., condition A's mean distance flown is higher than condition B's), but the model also finds a 1% chance that the difference could be negative (i.e., condition B actually has a higher mean than condition A if we had access to the population instead of a sample). This would be reported as a significant difference between A and B with 99% confidence in the difference (the complement of the 1% chance that B is truly higher than A) and a 95% HDI of (0.5, 3.1) nm.

Aside from the description of participant demographic data and sector differences, the results are organized according to our four independent variables (i.e., delivery method, T1, T2, and communications availability). Within each condition, results are broken out for each type of dependent variable (e.g., safety). Only statistically significant results (95% confidence or greater) are reported along with some select results that are consistent with statistically significant results but did not reach significance themselves. An exhaustive summary of results can be found in the appendices. Please note that for non-normally distributed data (e.g., WAK data, survey responses; see Appendices C, D, and E) the HDI values may not be on the same scale as the data.

5. RESULTS

5.1. PARTICIPANT BACKGROUND QUESTIONNAIRE

Participants were active controllers from seven different facilities across the contiguous United States. Participant ages ranged from 32 to 55 years old with a median age of 43.5. All participants were male. Participants' years of experience as ATCS (including possible military time) ranged from 9 to 35 (median of 24.5) years with a range of 9 to 35 (median of 16) years specifically in an en route facility. Only one participant controlled traffic for less than all of the previous year; that participant reported seven months of active work. All but one of the participants had prior experience working UAS at their position, with the most frequent response being that they worked UAS daily or weekly. The participants largely worked Global Hawks and Predators. Most participants reported neutral (5 on a scale from 1 = negative to 9 = positive) feelings as to how UAS impact ATC services, although three gave a slightly negative (3 or 4) rating.

5.2. STATISTICAL RESULTS

5.2.1. Loss of Separation

Only three loss of separation events occurred during the study. All three happened during automatic delivery trials but none involved the lost link aircraft. Given the small number of events, we did not attempt any statistical analysis and we do not consider them meaningful to the outcome of the study.

5.2.2. Factor 1: NAS Delivery Mechanism

The delivery mechanism describes how lost link routes were delivered to participants. Recall that there were four levels of delivery: a no lost link baseline scenario with no delivery, scenarios where communications were maintained and the UAS pilot told the participant the lost link route (*pilot* or *voice*), scenarios where communications were lost and a researcher gave the participant a handout with the lost link route (*manual*), and scenarios where communications were lost and an ABRR-like feature showed the lost link route and the participant accepted it (*automatic*). Only statistically significant results are reported. If a dependent variable or data source (e.g., PSQ) is not listed, then there were no statistically significant differences with regard to that variable. Please note that for non-normally distributed data (e.g., WAK data, survey responses; see Appendices C, D, and E), the confidence interval values may not be on the same scale as the data.

5.2.2.1. Safety

Post-Scenario Questionnaires

When participants rated the overall safety of operations (on the scale 1 = extremely low to 10 = extremely high), we observed a three-way interaction between T1, T2, and delivery method. If T1 was long and T2 was short there was a general decrease in ratings from automatic to manual to pilot (automatic mean 9.3, manual mean 8.7, pilot mean 8.5), but manual was similar to pilot if T2 was long (automatic mean 9.5, manual mean 8.3, pilot mean 8.5). If T1 and T2 were short there was a precipitous drop across the delivery conditions (automatic mean 9.6, manual mean 8.7, pilot mean 7.3) but if T1 was short and T2 was long automatic and manual were similar and pilot was rated especially poorly (automatic mean 9, manual mean 9, pilot mean 6.5).

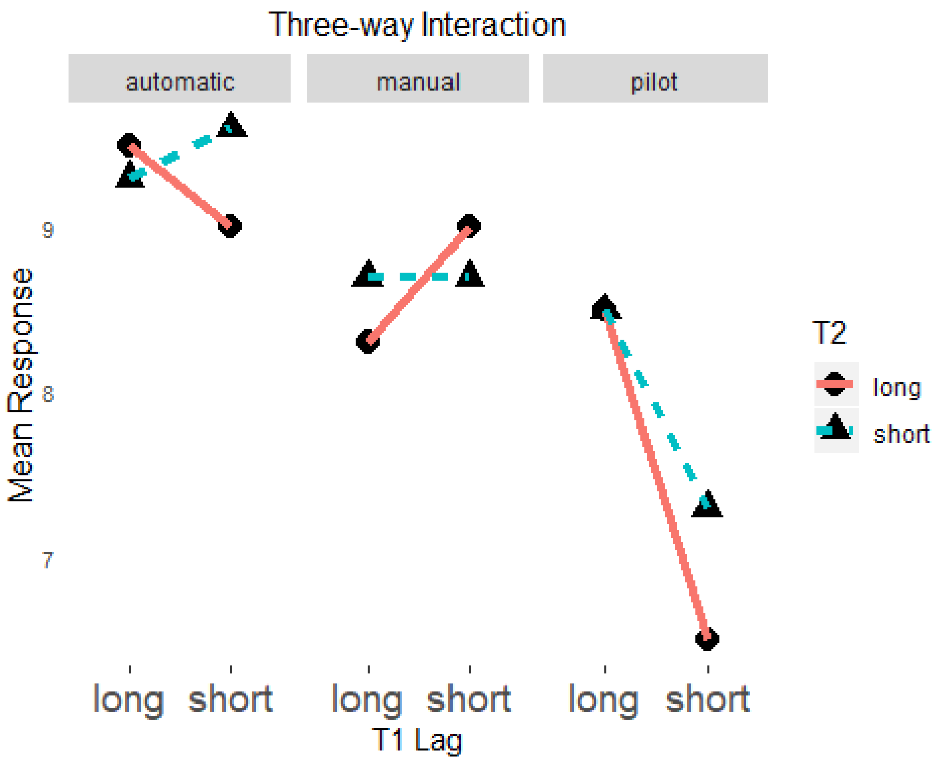


Figure 9. PSQ Overall Safety of Operations ratings

While complicated by the three-way interaction just described, we also observed a significant two-way interaction between T1 and delivery method. Ratings for manual and pilot delivery were similar if T1 was long (8.5 for pilot, 8.5 for manual, and 9.4 for automatic) but pilot was notably lower if T1 was short (6.9 versus 8.9 for manual and 9.3 for automatic). Also, as can be seen in the figure above, there is a main effect of delivery such that each of the three methods were significantly different from each other (95% HDI for pilot compared to manual [1.08, 2.28]; 95% HDI for pilot compared to automatic [1.85, 3.00]; 95% HDI for manual compared to automatic [0.27, 1.26]; mean rating for pilot 7.7, manual 8.7, automatic 9.4).

We observed an interaction between T1 and delivery method when participants were asked to rate how safely they resolved the lost link event (on the scale 1 = extremely low to 10 = extremely high). The interaction occurred because the pilot/manual difference was relatively small after a long T1 (pilot mean 8.9, manual mean 8.2) but was larger (and reversed) after a short T1 (pilot mean 7.5, manual mean 9.4). Ratings were high after automatic delivery regardless of T1 (9.6 for long T1, 9.4 for short). These effects were largely driven by low ratings in the short T1 – short T2 – pilot condition (mean of 6.9, all other conditions 8.0 or higher), but the three-way interaction was not statistically significant. We also observed a main effect of delivery method where each of the three methods were significantly different from each other (95% HDI for pilot compared to manual [0.35, 1.07]; 95% HDI for pilot compared to automatic [0.74, 1.92]; 95% HDI for manual compared to automatic [0.02, 1.34]; mean rating for pilot 8.2, manual 8.8, automatic 9.5).

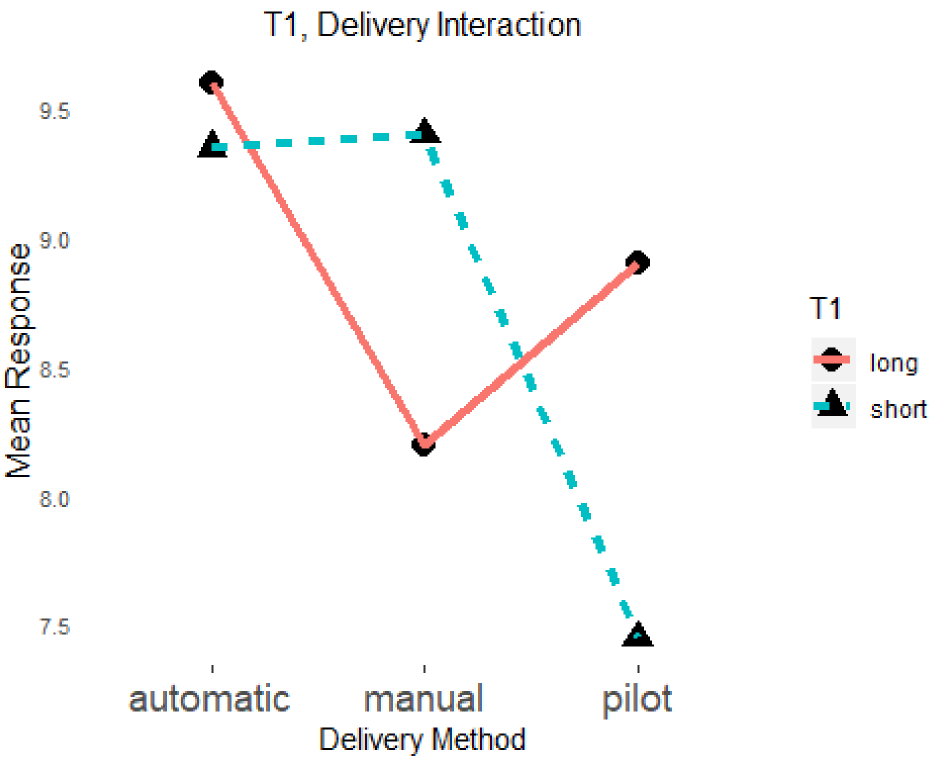


Figure 10. PSQ How Safely was the Lost Link Resolved ratings

5.2.2.2. Efficiency

Post-Scenario Questionnaires

When participants were asked to rate the overall efficiency of operations (on the scale 1 = extremely low to 10 = extremely high), we observed an interaction between T1 and delivery method. Ratings for manual and pilot delivery were similar if T1 was long (8.1 for pilot, 8.2 for manual, and 9.3 for automatic) but pilot was notably lower if T1 was short (7.2 versus 8.5 for manual and 8.6 for automatic). We also observed a main effect. We found that each of the three methods were significantly different from each other (95% HDI for pilot compared to manual [0.01, 1.79]; 95% HDI for pilot compared to automatic [0.92, 2.06]; 95% HDI for manual compared to automatic [0.06, 1.22]; mean rating for pilot 7.7, manual 8.4, automatic 8.9).

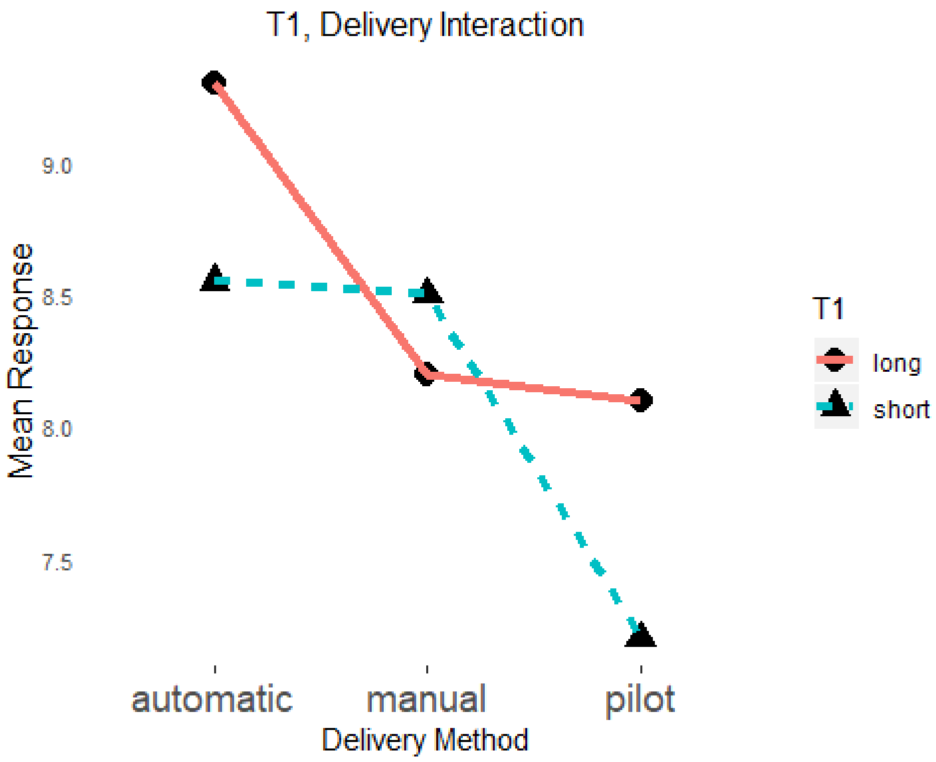


Figure 11. PSQ Overall Efficiency of Operations ratings

We observed a three-way interaction between T1, T2, and delivery method when participants were asked to rate how efficiently they resolved the lost link event (on the scale 1 = extremely low to 10 = extremely high). While complicated by the three-way interaction, we also observed a two-way interaction between T1 and delivery method. There was a small difference between manual (mean rating of 8.3) and pilot (8.5) if T1 was long but a larger difference (7.6 for pilot versus 9.2 for manual) if T1 was short. Ratings for automatic delivery were high regardless of T1 lag, means of 9.4. In addition, we observed a main effect of delivery method: both manual and automatic delivery methods had higher ratings than pilot but the manual-automatic difference did not reach significance (95% HDI for pilot compared to manual [0.28, 1.74]; 95% HDI for pilot compared to automatic [0.65, 2.44]; mean rating for pilot 8, manual 8.7, and automatic 9.4).

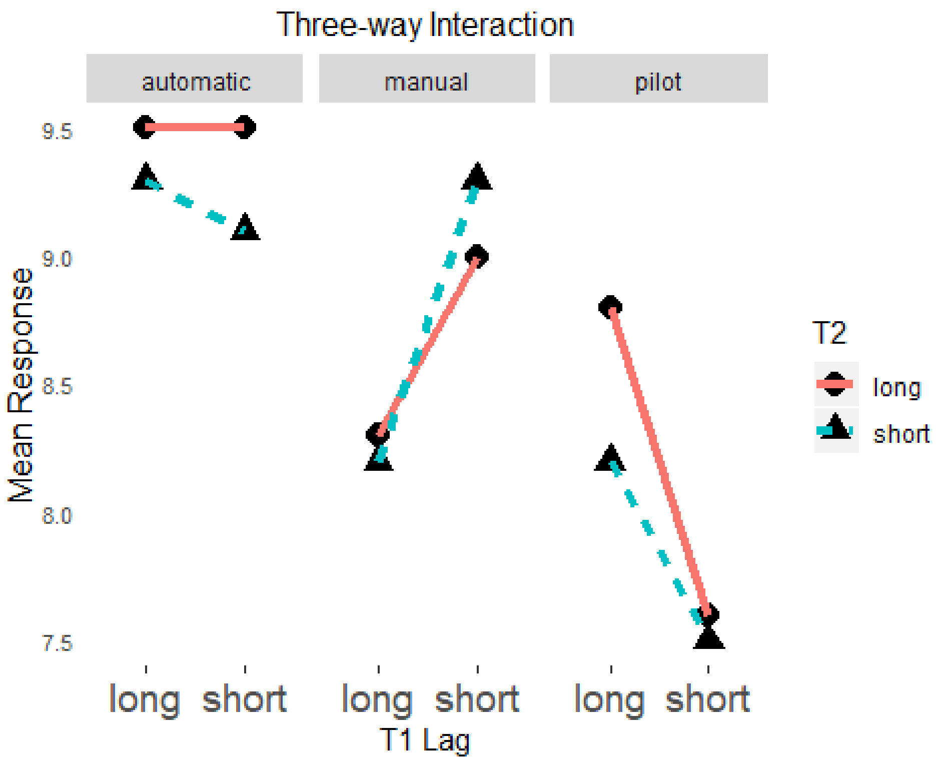


Figure 12. PSQ Efficiency in Resolving Lost Link ratings

Time and Distance Flown in Sector

There was no significant effect of delivery method on overall flight times or distances. However, we did observe an effect when limiting the data to UAS. We observed that average flight times were longer in the automatic than either the pilot or (at trend) manual conditions (95% HDI for pilot compared to automatic [0.14, 0.83]; 95% HDI for manual compared to automatic [-0.03, 0.65]; pilot mean 1105 seconds, manual mean 1113, automatic mean 1130). There was a similar result for average flight distance, although all three conditions were significantly different from each other (95 % HDI for pilot compared to manual [0.17, 0.39]; 95% HDI for pilot compared to automatic [0.56, 0.79]; 95% HDI for manual compared to automatic [0.28, 0.51]; pilot mean 85.6, manual mean 86.3, automatic mean 87.5). This effect was not due to the lost link UAS, as lost link UAS actually flew shorter times and distances in the automatic condition.

Time and distance flown in sector are important indicators of efficiency: in general, aircraft use less fuel when the travel shorter distances and ATCS have lower workloads when aircraft spend less time in their sector. While there are statistically significant reductions in time and distance flown by UAS, they are small in terms of practical significance (around one nautical mile and 30 seconds per aircraft). Thus, we do not place much importance on these results.

5.2.2.3. Situation Awareness

Post-Scenario Questionnaires

We observed a similar T1 – delivery method interaction as seen in the previous results when participants rated their overall situation awareness (on the scale 1 = extremely low to 10 = extremely high). There was a minimal effect of pilot/manual delivery when T1 was long (8.2 for pilot and 8.1 for manual, 9.4 for automatic) but ratings were lower for pilot delivery when T1 was short (7.3 versus 8.7; mean rating for automatic was 9.5). There was also a main effect of delivery such that the average rating for situation awareness was higher in the automatic condition compared to either manual or pilot (95% HDI for pilot compared to automatic [0.76, 2.94]; 95% HDI for manual compared to automatic [0.76, 1.82]; pilot mean rating 7.7, manual mean rating 8.4, automatic mean rating 9.4).

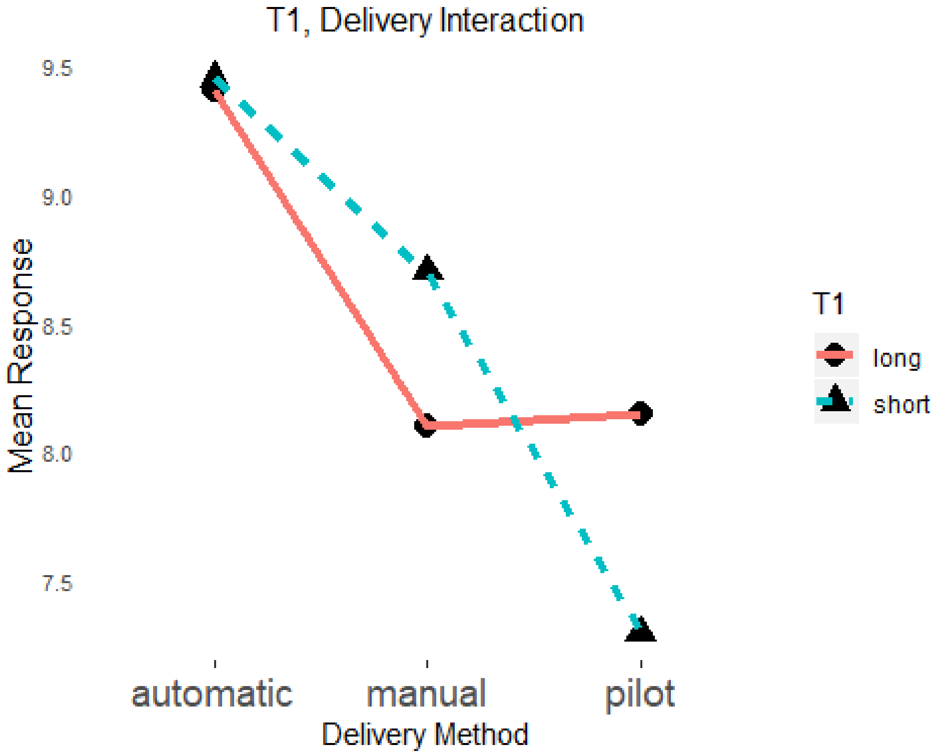


Figure 13. PSQ Overall Situation Awareness ratings

We observed similar results when participants were asked to rate their situation awareness for UAS progress along the lost link route (on the scale 1 = extremely low to 10 = extremely high), with an interaction between T1 and delivery method as well as a main effect of delivery method. There was little difference between pilot and manual delivery after a long T1 (8.3 for pilot, 7.7 for manual) but a larger (and reversed) difference after a short T1 (7.4 and 8.7; automatic delivery was higher in each case, 9.3 after long T1 and 9.2 after short). All three delivery methods were significantly different from each other (95% HDI for pilot compared to manual [0.01, 1.48]; 95% HDI for pilot compared to automatic [0.83, 2.63]; 95% HDI for manual compared to automatic [0.35, 2.1]; pilot mean rating 7.9, manual mean rating 8.2, automatic mean rating 9.2).

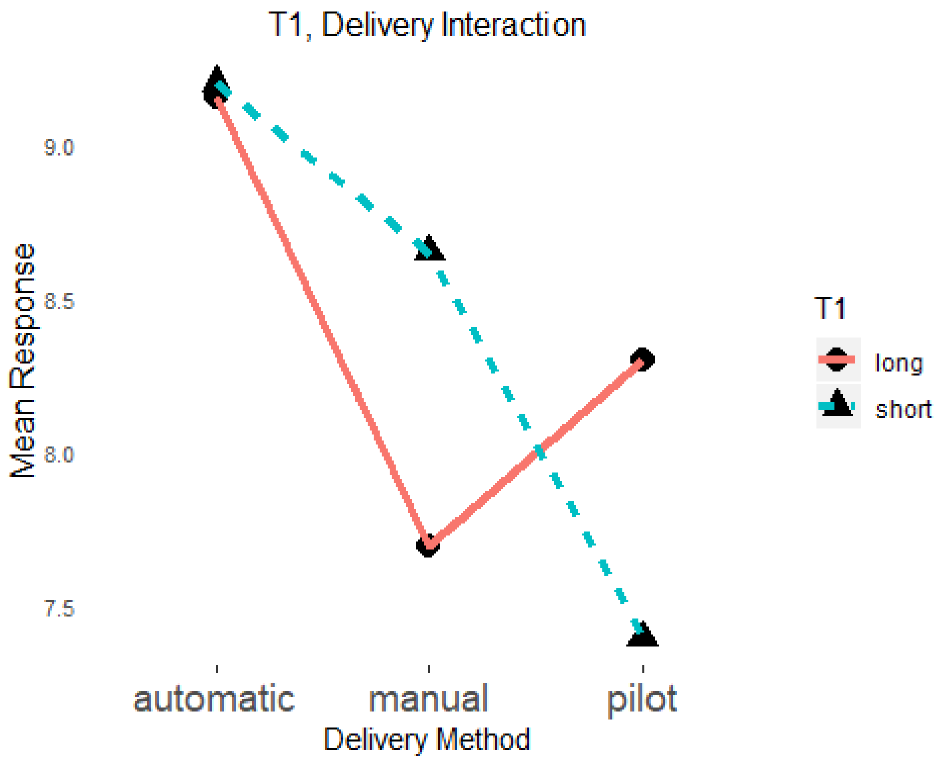


Figure 14. PSQ Situation Awareness for UAS on Lost Link Route ratings

5.2.2.4. Workload

Workload Assessment Keypad (WAK) responses

On the WAK prompt that immediately followed the delivery of lost link intent, there was no statistically reliable difference in responses between any of the delivery mechanisms according to our model (Appendix E). However, ratings were generally higher in the pilot conditions than the automatic conditions. As can be seen in Figure 15, this difference begins after the lost link event occurs (i.e., the earliest lost link events occur around 9 minutes into the scenario and the three conditions are similar up to that point).

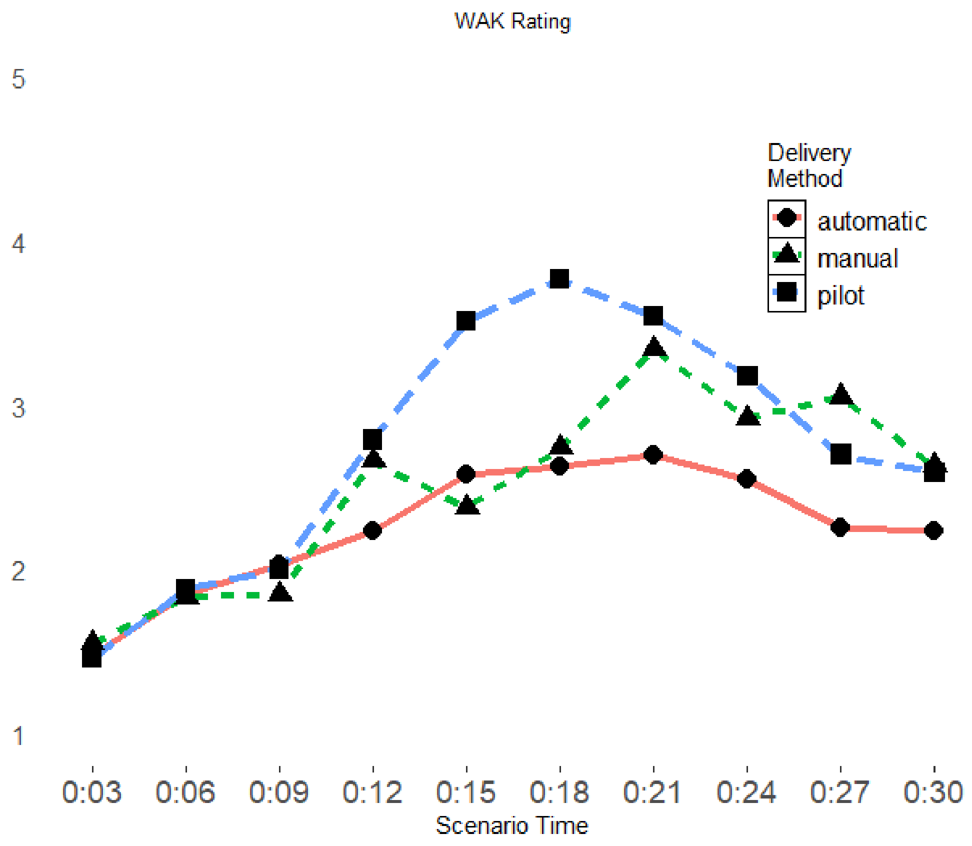


Figure 15. Mean WAK response timeline for each delivery condition

Number and Duration of Communications

The total number of push-to-talk communications was significantly higher during pilot delivery conditions than either automatic or manual (95% HDI comparing pilot to manual [0.01, 0.14]; 95% HDI comparing pilot to automatic [0.00, 0.13]; pilot mean number of communications 74, manual mean number 68.4, automatic mean number 69.6).

NASA-TLX

We observed generally lower responses in the automatic delivery condition, and some evidence for increased workload in the pilot delivery condition. The pilot condition systematically had higher workload ratings than the manual condition (and, in turn, the automatic condition), but was only significantly different for three of the six questions.

When rating their mental demand, participants gave a significantly higher rating in the manual condition than the automatic condition (95% HDI [0.29, 1.40]; pilot mean 5.4, manual mean 5.0, automatic mean 4.1).

When rating their physical demand, participants gave significantly higher ratings in the manual and pilot conditions than the automatic condition (95% HDI comparing manual to automatic

[0.41, 1.10]; 95% HDI comparing pilot to automatic [0.29, 1.59]; pilot mean 3.7, manual mean 3.5, automatic mean 2.5).

When rating their temporal demand, participants gave significantly higher ratings in the manual and pilot conditions than the automatic condition (95% HDI comparing manual to automatic [0.05, 1.04]; 95% HDI comparing pilot to automatic [0.19, 1.99]; pilot mean 5.0, manual mean 4.8, automatic mean 3.7).

When rating their performance, participants gave significantly lower ratings in the manual and pilot conditions than the automatic condition (95% HDI comparing manual to automatic [0.32, 1.54]; 95% HDI comparing pilot to automatic [0.51, 1.46]; pilot mean 8.6, manual mean 8.7, automatic mean 9.4).

When rating their effort, participants gave significantly different ratings to all three conditions (95% HDI comparing pilot to manual [0.11, 1.69]; 95% HDI comparing pilot to automatic [1.12, 2.07]; 95% HDI comparing manual to automatic [0.22, 1.43]; pilot mean 5.4, manual mean 4.6, automatic mean 3.5).

When rating their frustration, participants gave significantly different ratings to all three conditions (95% HDI comparing pilot to manual [0.84, 1.37]; 95% HDI comparing pilot to automatic [1.18, 2.08]; 95% HDI comparing manual to automatic [0.21, 0.83]; pilot mean 4.2, manual mean 3.0, automatic mean 2.1).

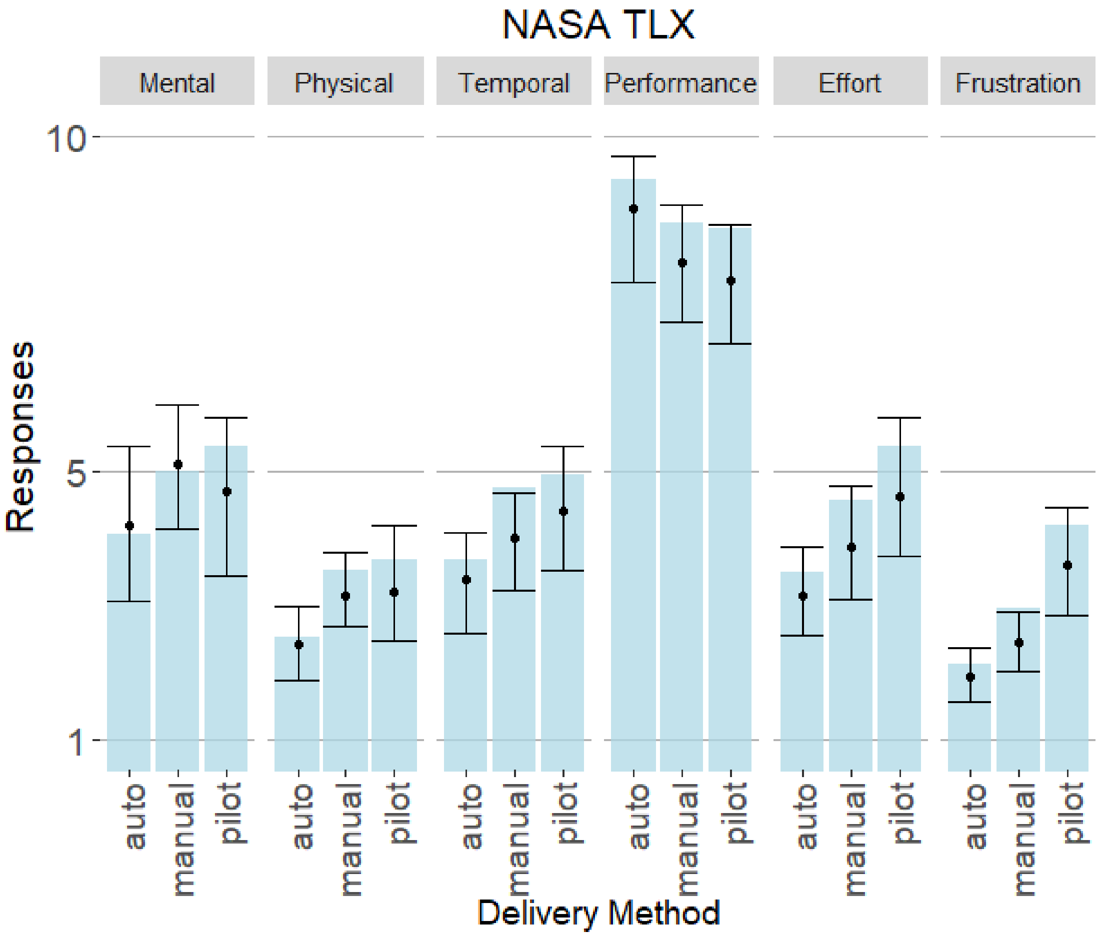


Figure 16. NASA TLX responses summarized by delivery method. The bar heights correspond to the mean of the participants' responses. The dot and error bars indicate the modeled mean and 95% HDI

Post-Scenario Questionnaires

We observed a main effect of delivery when participants were asked to rate their workload due to the lost link event (on the scale 1 = extremely low to 10 = extremely high). All three conditions were significantly different from each other (95% HDI comparing pilot to manual [0.05, 1.08]; 95% HDI comparing pilot to automatic [1.78, 2.5]; 95% HDI comparing manual to automatic [1.17, 2.15]; pilot mean 5.8, manual mean 5.4, automatic mean 3.5).

Amendment Entry Times

Delivery method had a notable impact on amendment entry times. First amendments were entered significantly later after lost link route delivery in the pilot condition compared to either the manual or automatic conditions (95% HDI comparing pilot to manual [0.61, 1.17]; 95% HDI comparing pilot to automatic [0.42, 1.03]; pilot mean 105 seconds, manual mean 43, automatic mean 58). Final amendment entry was also later in the pilot condition compared to the other two, and additionally automatic entry was earlier (95% HDI comparing pilot to manual [0.46, 1.06];

95% HDI comparing pilot to automatic [1.12, 1.73]; 95% HDI comparing manual to automatic [0.39, 1.00]; pilot mean 228 seconds, manual mean 110, automatic mean 63). As seen in Figure 17, a number of amendments were not entered until after the T2 time had passed, meaning that the UAS had already began turning onto the lost link route. This occurred particularly often when T2 was short (120 seconds/2 minutes) and when the pilot delivered the lost link routing.

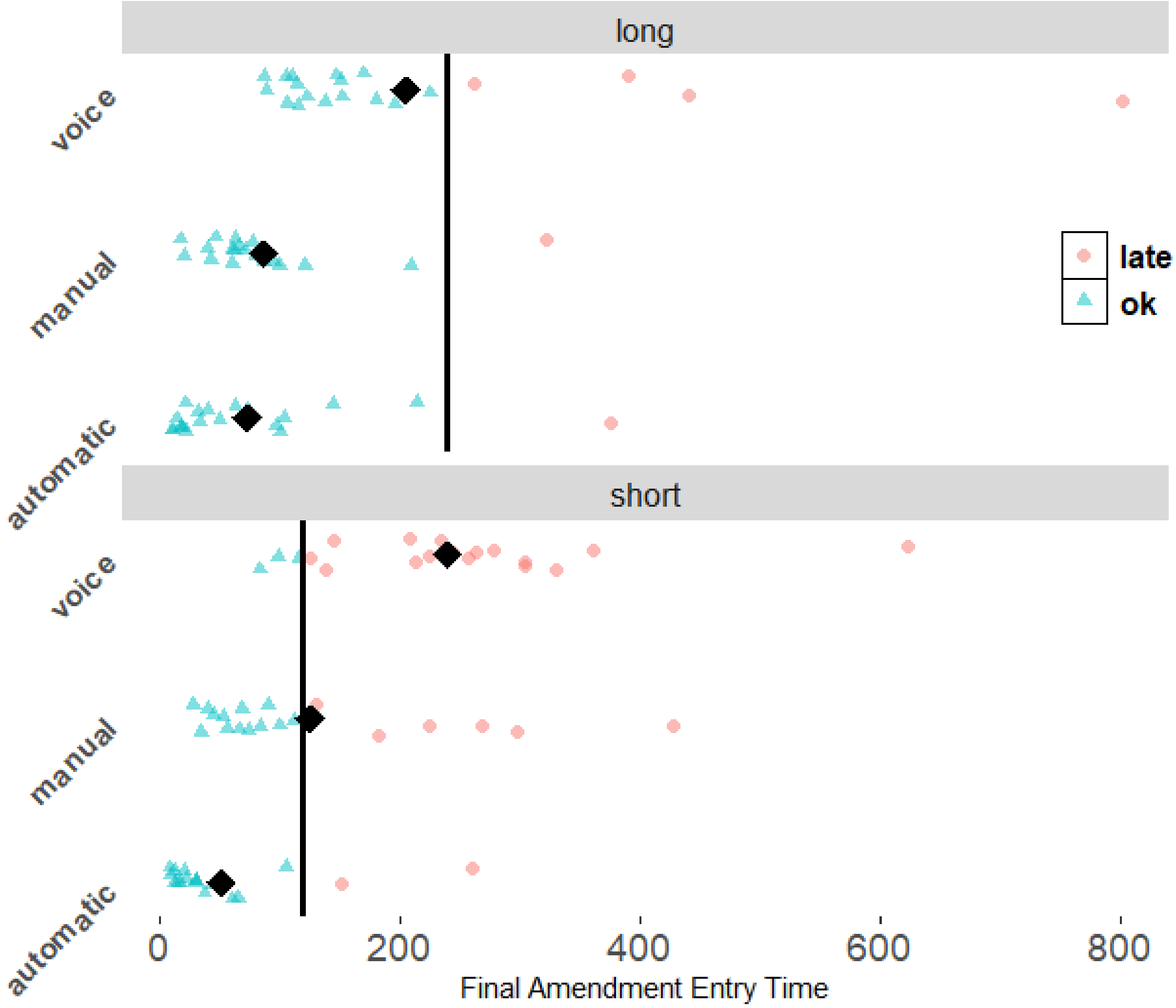


Figure 17. Amendment times relative to T2 time for each delivery method. Each dot represents an individual final amendment entry time. The dots are shape and color coded as to if they occurred before or after the appropriate T2 time had passed (T2 time marked by the vertical black line; long T2 conditions on top, short T2 conditions on bottom). The black diamond marks the average time for the condition.

As follows from the previous two results, all three conditions were significantly different from each other when looking at the time between first and last amendment entry (95% HDI comparing pilot to manual [0.25, 1.14]; 95% HDI comparing pilot to automatic [2.56, 3.48]; 95% HDI comparing manual to automatic [1.89, 2.79]; pilot mean 123 seconds, manual mean 66, automatic mean 4).

5.2.2.5. Additional Questions

When asked on the PSQ to rate the effectiveness of the delivery method (1 meaning very ineffective, 10 meaning very effective), all three methods were significantly different from each other with a large preference for automatic (95% HDI for manual versus pilot [0.12, 1.74]; 95% for automatic versus pilot [3.54, 4.79]; 95% HDI for manual versus automatic [2.34; 3.89]; pilot mean 5, manual mean 5.6, automatic mean 8.8). Automatic delivery was also greatly preferred when participants rated the effectiveness of entering the lost link route (95% for automatic versus pilot [3.58, 4.94]; 95% HDI for manual versus automatic [3.09, 4.88]; pilot mean 5, manual mean 5.5, automatic mean 9.2).

On the exit questionnaire, participants rated pilot delivery as slightly negative (mean response 4), manual delivery as positive (mean response 7.2), but automatic delivery as very positive (mean response 9.9; 95% HDI for automatic versus pilot [-5.0, -7.5]; 95% HDI for automatic versus manual [-2.4, -4.9]; 95% HDI for pilot versus manual [-1.1, -4.7]). They also rated manual entry of the lost link route as slightly negative (mean response 3.7) but automatic entry as positive (mean response 9.8; 95% HDI [-5.4, -7.8]).

5.2.3. Factor 2: Communications Availability

Communication availability describes if UAS pilots maintained normal air-ground contact with the participant after a lost link occurred. Recall that there were two levels of availability: scenarios in which communications stayed available (sometimes marked as ‘yes’) and scenarios in which communications were lost (sometimes marked as ‘no’). Alternatively, the communications variable can be seen as a comparison between pilot delivery (when communications were maintained) and the combination of manual and automatic delivery (when communications were lost after lost link). Due to the overlap between delivery method and communications availability, we report the communications results as an abbreviated version of the previous section.

5.2.3.1. Safety

Post-Scenario Questionnaire

When participants rated the overall safety of operations (on the scale 1 = extremely low to 10 = extremely high), we observed a two-way interaction between communications and T1. There was a small effect of communications after a long T1 (8.5 available and 9 unavailable) but a larger effect after a short T1 (6.9 available, 9 unavailable).

We observed a similar interaction when participants rated their performance in safely resolving the lost link event (on the scale 1 = extremely low to 10 = extremely high). When T1 was long there was no effect of communications (available mean 8.9, unavailable mean 8.8) but when T1 was short ratings went down for communications available (available mean 7.5, unavailable mean 9.4).

5.2.3.2. Efficiency

Post-Scenario Questionnaires

When participants were asked to rate the overall efficiency of operations (on the scale 1 = extremely low to 10 = extremely high), we observed an interaction between T1 and communications availability. There was a small effect, if any, of communications after a long T1 (8.1 available, 8.8 unavailable) but a larger effect after a short T1 (7.2 available and 8.5 unavailable).

We observed a similar interaction when participants rated their performance in efficiently resolving the lost link event (on the scale 1 = extremely low to 10 = extremely high). The effect of communications was larger after a short T1 (mean rating of 7.6 available versus 9.2 unavailable) than after a long T1 (8.5 versus 8.8).

5.2.3.3. Situation Awareness

Post-Scenario Questionnaires

When participants were asked to rate their overall situation awareness (on the scale 1 = extremely low to 10 = extremely high), we observed an interaction between T1 and communications availability. There was a small communications effect when T1 was long (mean rating 8.2 available versus 8.8 unavailable) but a larger effect when T1 was short (7.3 versus 9.1).

We observed similar results when participants were asked to rate their situation awareness for UAS progress along the lost link route (on the scale 1 = extremely low to 10 = extremely high). There was a small effect of communications after a long T1 (mean rating 8.3 available versus 8.4 unavailable) but a larger effect after a short T1 (7.4 versus 8.9).

5.2.3.4. Workload

Workload Assessment Keypad (WAK) responses

Communications availability did not significantly affect WAK responses according to our model (Appendix E). However, ratings were generally higher in the communications available than the communications unavailable conditions. As can be seen in the figure below, this difference begins after the lost link event occurs (i.e., the earliest lost link events occur around 9 minutes into the scenario and the three conditions are similar up to that point).

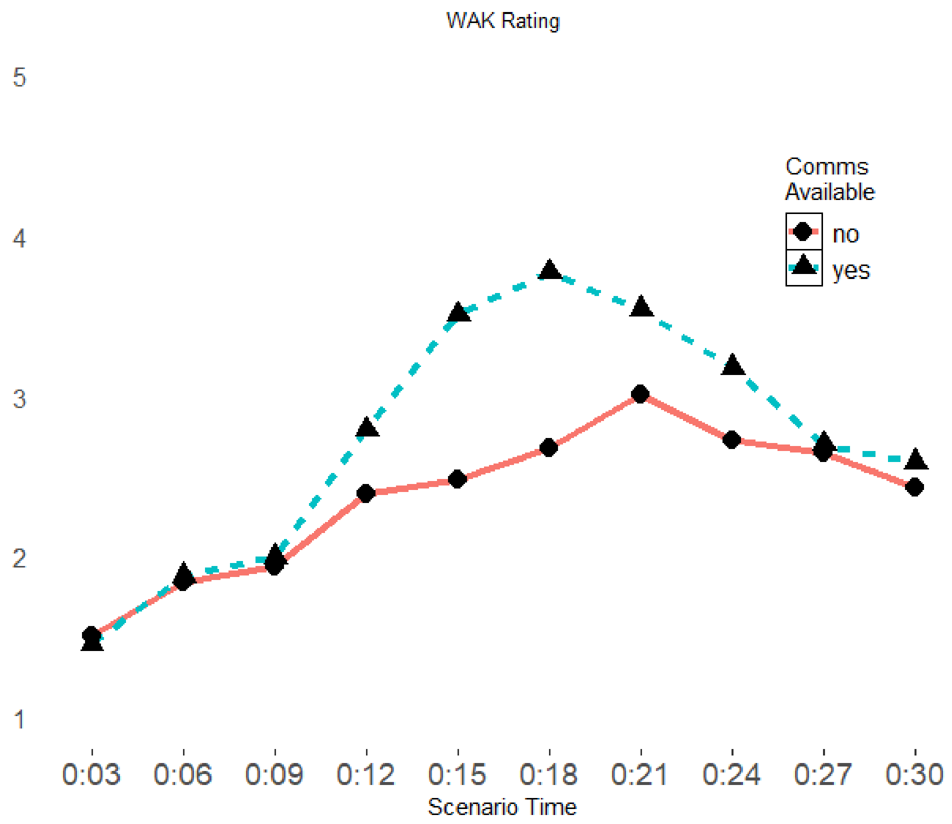


Figure 18. Mean WAK response timeline for each communications condition

Number and Duration of Communications

Although only one aircraft ever lost communications during a scenario (the lost link UAS, after it lost link), this was sufficient to affect the overall number of communications made during scenarios. The total number of push-to-talk communications was significantly higher during communications available conditions than communications unavailable conditions (mean 74 versus 69).

NASA-TLX

We observed a main effect of communications on workload. Ratings were significantly higher (i.e., higher workload) for communications available for each of the six questions except mental demand (which was still numerically higher).

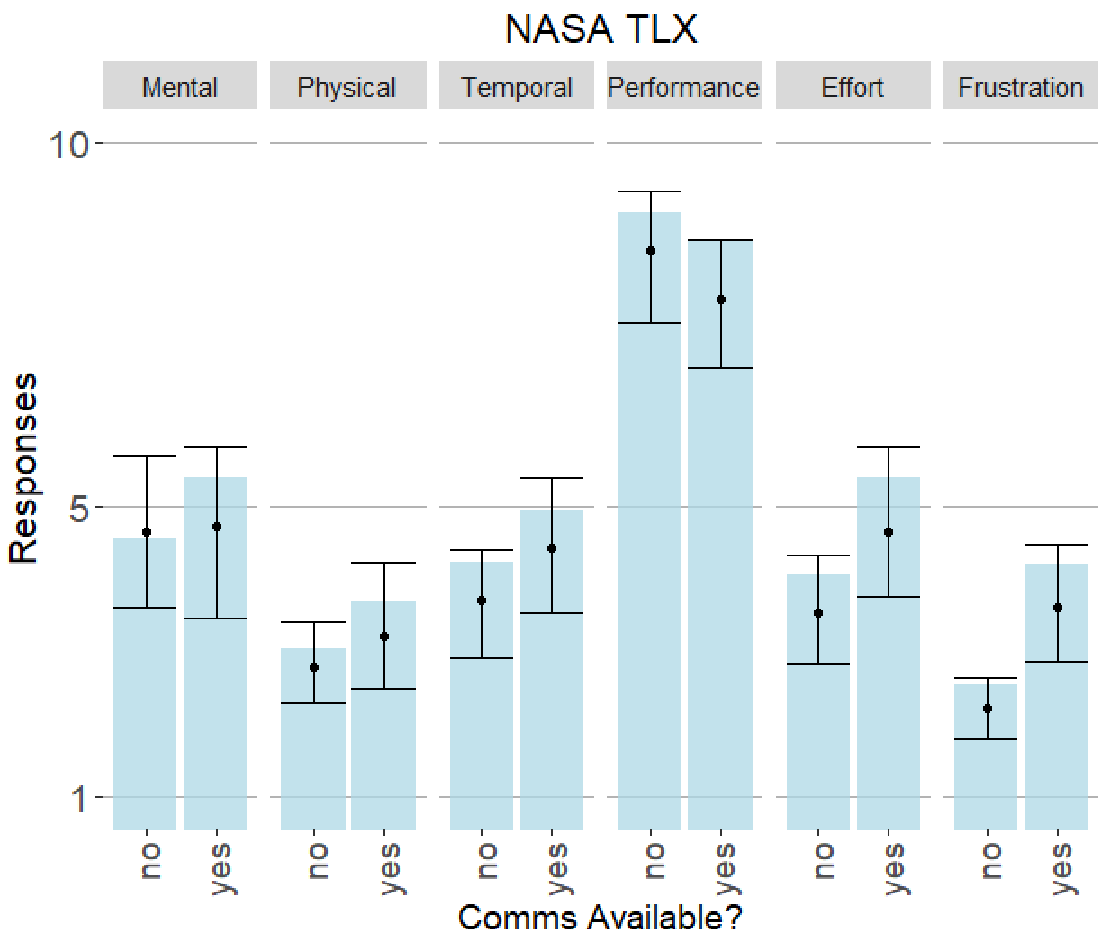


Figure 19. NASA TLX responses summarized by communications availability. The bar heights correspond to the mean of the participants' responses. The dot and error bars indicate the modeled mean and 95% HDI

Post-Scenario Questionnaires

We observed a main effect of communications availability when participants were asked to rate their workload (on the scale 1 = extremely low to 10 = extremely high) due to the lost link event (mean 5.8 available versus 4.5 unavailable).

Amendment Entry Times

Communications availability had a notable impact on amendment entry times. Participants entered first amendments later after lost link route delivery when communications were available, as was the case for final amendments. The time between first and final amendment was also longer if communications were available.

5.2.3.5. Additional Questions

On the exit questionnaire, participants were asked to rate the impact (1 = very negative to 10 = very positive) of losing pilot communications during a lost link event. They gave a mean response of 4.3, which was not significantly different from a neutral response of 5.

5.2.4. Factor 3: T1 Lag

The T1 lag is the amount of time between when a lost link UAS squawks 7400 and the lost link route is provided to the participant. Recall that there were three levels of T1 lag: a no lost-link baseline with no T1 value, scenarios with a T1 lag of 1 minute (short), and scenarios with a T1 lag of 4 minutes (long). Only statistically significant results are reported. If a particular dependent variable or data source (e.g., post-scenario questionnaires) is not listed, then there were no statistically significant differences with regard to that variable. Please note that for non-normally distributed data (e.g., WAK data, survey responses; see Appendices C, D, and E), the confidence interval values may not be on the same scale as the data.

5.2.4.1. Safety

Post-Scenario Questionnaires

When participants rated the overall safety of operations (on the scale 1 = extremely low to 10 = extremely high), we observed a three-way interaction between T1, T2, and delivery method. If the lost link route was delivered automatically there was at best a small effect of T1 or T2 (all ratings 9.0 or higher); if the route was delivered manually there was no effect of T1 if T2 was short (both ratings 8.7) but a lower rating after a long T1 if T2 was short (long T1 mean 8.3, short T1 mean 9.0); if the route was delivered by the pilot there were higher ratings for long T1 than short T1 but this difference was larger after a long T2 (both long T1 conditions mean 8.5, short T1 short T2 mean 7.3, short T1 long T2 mean 6.5).

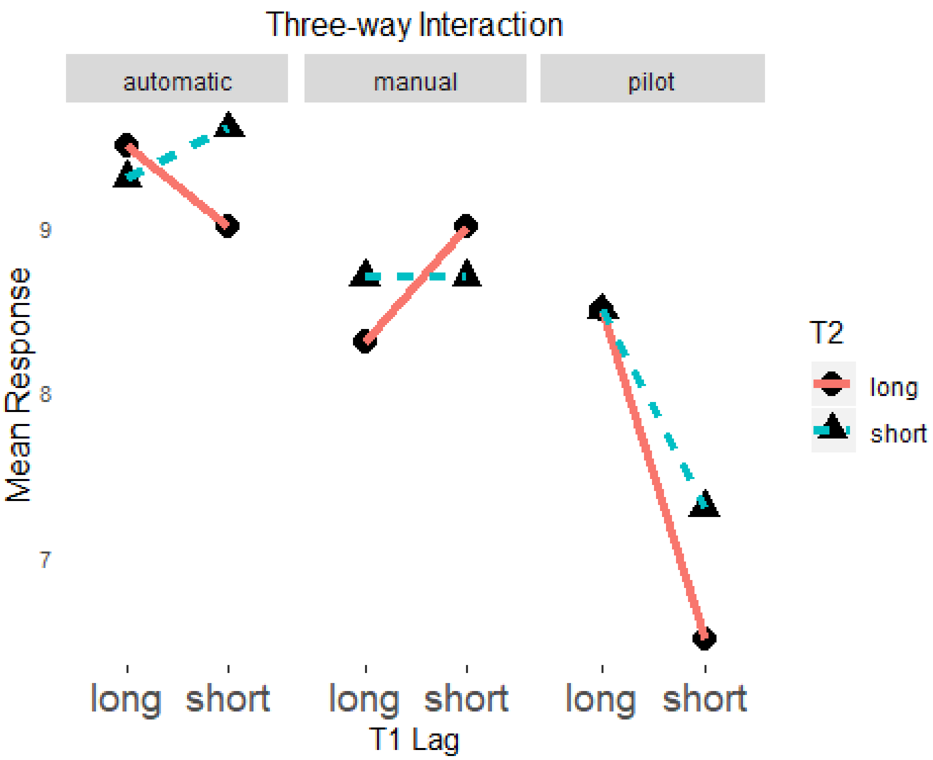


Figure 20. PSQ Overall Safety of Operations ratings

While complicated by the three-way interaction just described, we also observed a significant two-way interaction between T1 and delivery method. Ratings for long and short T1 were similar for manual or automatic delivery (manual: long T1 mean 8.5, short T1 mean 8.9; automatic: long T1 mean 9.4, short T1 mean 9.3) but long T1 was better under pilot delivery (long T1 mean 8.5, short T1 mean 6.9). The particularly low ratings in the short T1-pilot conditions cause an overall main effect of T1 (long T1 mean 8.8, short T1 mean 8.4).

We observed an interaction between T1 and delivery method when participants were asked to rate how safely they resolved the lost link event (on the scale 1 = extremely low to 10 = extremely high). When the lost link route was delivered by the pilot ratings were higher after a long T1 (long T1 mean 8.9, short T1 mean 7.5) but when it was delivered manually ratings were higher after a short T1 (long T1 mean 8.2, short T1 mean 9.4). Ratings were high after automatic delivery regardless of T1 (9.6 for long T1, 9.4 for short). These effects were largely driven by low ratings in the short T1 – short T2 – pilot condition (mean of 6.9, all other conditions 8.0 or higher), but the three-way interaction was not statistically significant.

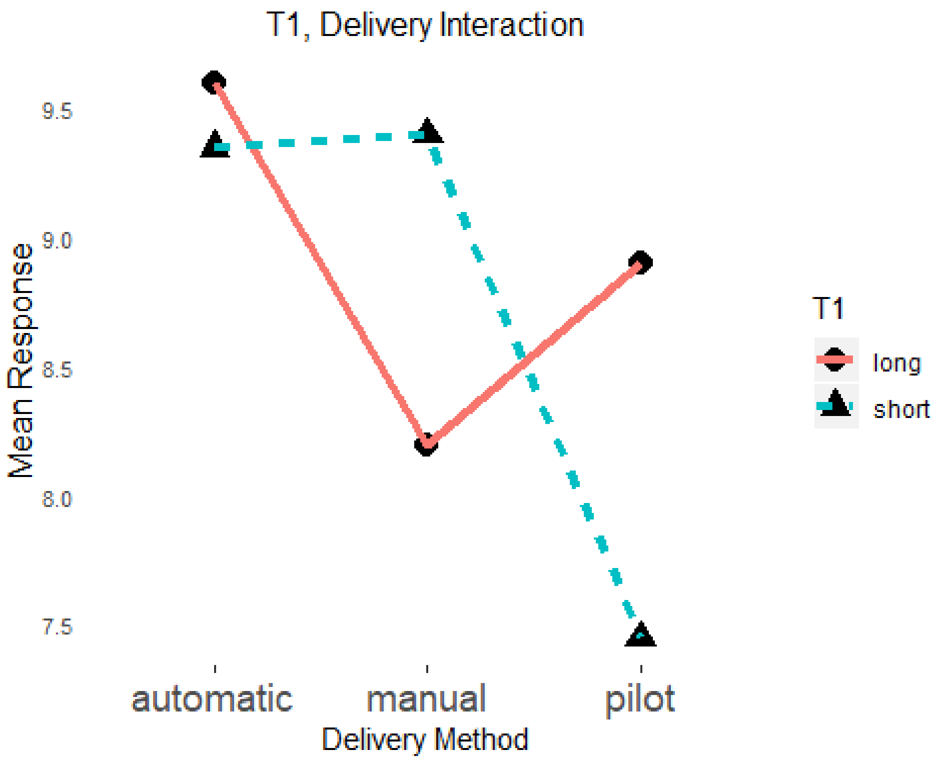


Figure 21. PSQ Safety in Resolving Lost Link ratings

<u>5.2.4.2. Efficiency</u>

Post-Scenario Questionnaires

When participants were asked to rate the overall efficiency of operations (on the scale 1 = extremely low to 10 = extremely high), we observed an interaction between T1 and delivery method. Long T1 was better if the pilot delivered the lost link route (long T1 mean 8.1, short T1 mean 7.2) but short T1 was slightly better if the route was delivered manually (long T1 mean 8.2, short T1 mean 8.5). The effect of T1 with automatic delivery (long T1 mean 9.3, short T1 mean 8.6) was similar to the pilot effect.

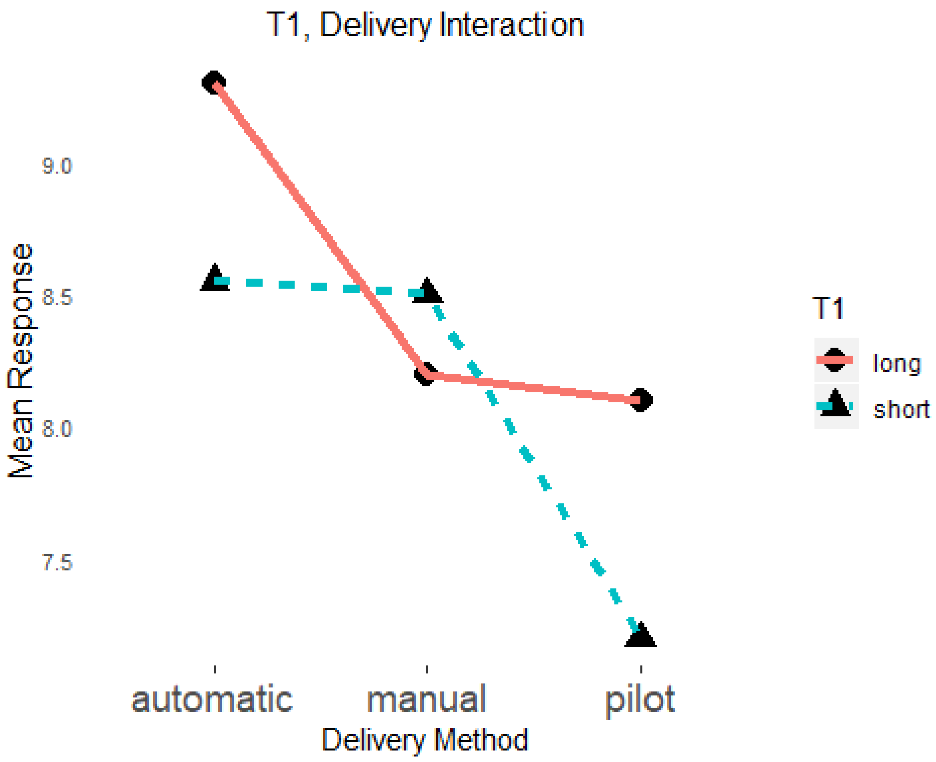


Figure 22. PSQ Overall Efficiency of Operations ratings

We also observed a significant interaction between T1 and T2 (95% HDI [0.5, 2.7]). If T2 is short, a long T1 is better (long T1 mean 8.8, short T1 mean 8) but if T2 is long there is not much difference (long T1 mean 8.3, short T1 mean 8.1).

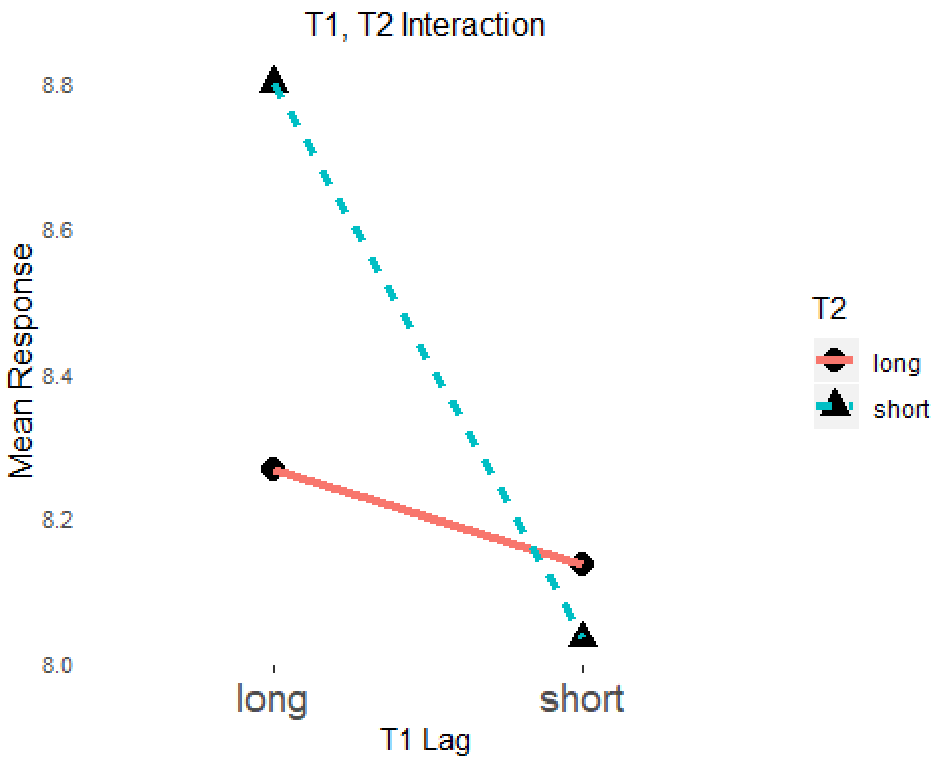


Figure 23. PSQ Overall Efficiency of Operations ratings

We observed a three-way interaction between T1, T2, and delivery method when participants were asked to rate how efficiently they resolved the lost link event (on the scale 1 = extremely low to 10 = extremely high). While complicated by the three-way interaction, we also observed a two-way interaction between T1 and delivery method. Short T1 was better if the route was delivered manually (long T1 mean 8.3, short T1 mean 9.2) but long T1 was better if the route was delivered by the pilot (long T1 mean 8.5, short T1 mean 7.6). Ratings for automatic delivery were high regardless of T1 lag, means of 9.4.

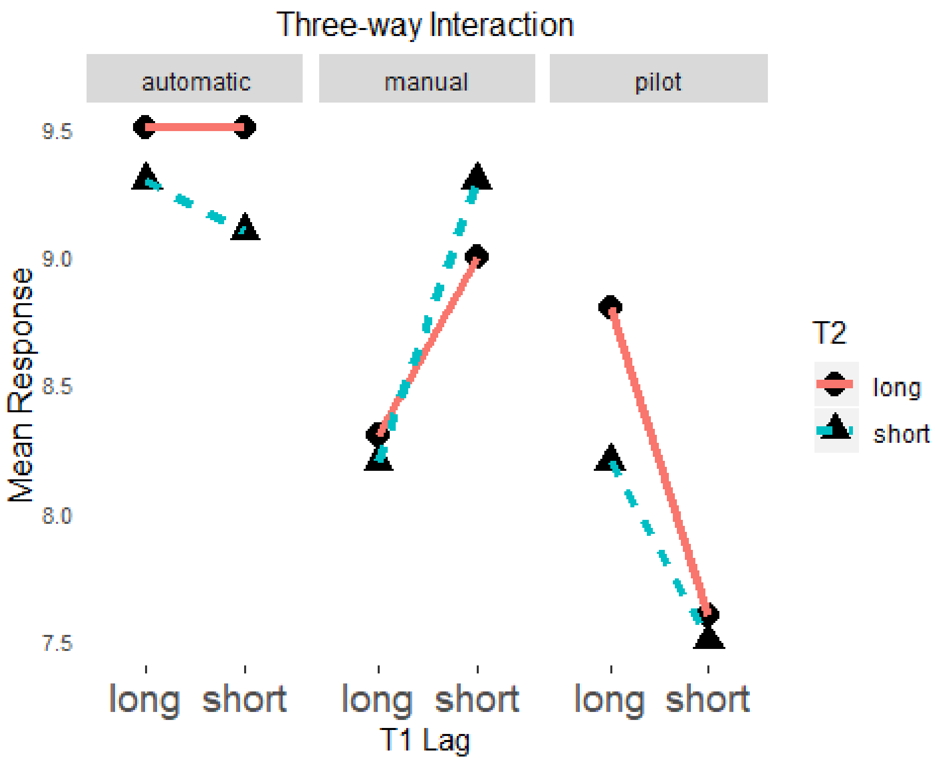


Figure 24. PSQ Efficiency in Resolving Lost Link ratings

Time and Distance Flown in Sector

We observed a small but statistically significant effect of T1 such that average flight distance was shorter when T1 is long (95% HDI [0.02, 0.34]; short T1 mean 75 nautical miles, long T1 mean 74.6 nm).

Time and distance effects were much more pronounced when limited to UAS. UAS average flight times were shorter by over 30 seconds when T1 was long (95% HDI [0.4, 0.94], short T1 mean 1134 seconds, long T1 mean 1098 seconds). This was matched by a similar effect for average flight distance (95% HDI [0.74, 0.93], short T1 mean 87.7 nautical miles, long T1 mean 85.3 nm). These results were not caused by the lost link aircraft as the time and distance effects are comparable or larger when the lost link UAS is removed from the data (difference of 1.9 nm average distance and 62 seconds average time).

As before, we note that while these results were statistically significant the differences are not large enough to be practically significant.

5.2.4.3 Situation Awareness

Post-Scenario Questionnaires

We observed a similar T1 – delivery method interaction as seen in the previous results when participants rated their overall situation awareness (on the scale 1 = extremely low to 10 =

extremely high). Long T1 was better with pilot delivery (long T1 mean 8.2, short T1 mean 7.3) but short T1 was better with manual delivery (long T1 mean 8.1, short T1 mean 8.7). Ratings were high for either T1 lag with automatic delivery (long T1 mean 9.4, short T1 mean 9.5).

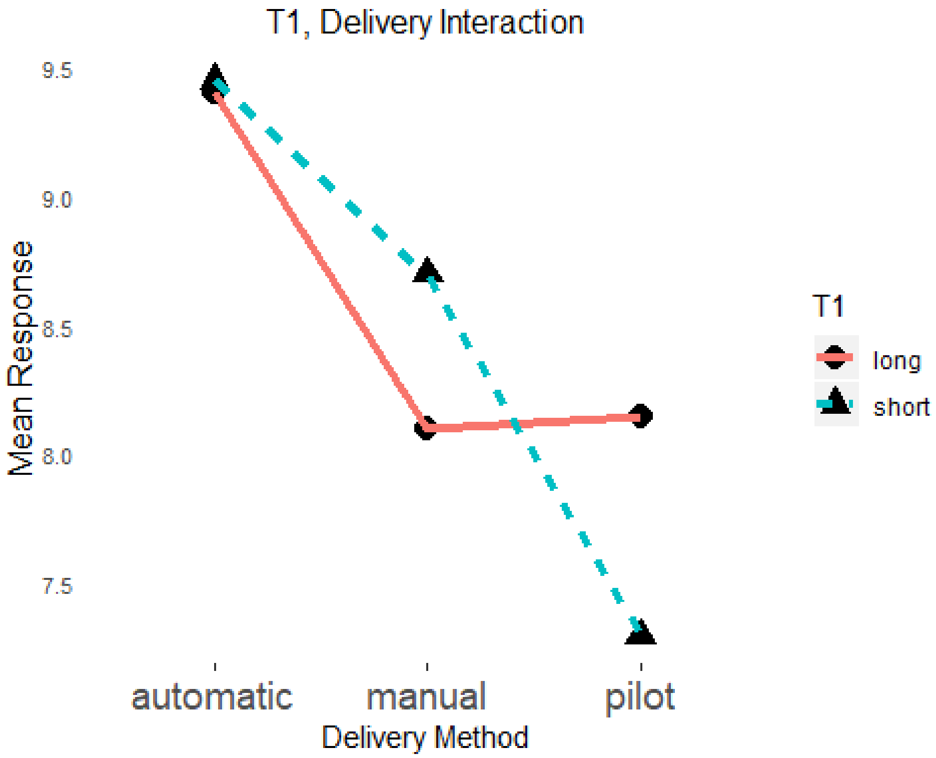


Figure 25. PSQ Overall Situation Awareness ratings

We observed similar results when participants were asked to rate their situation awareness for UAS progress along the lost link route (on the scale 1 = extremely low to 10 = extremely high). Long T1 was better with pilot delivery (long T1 mean 8.3, short T1 mean 7.4) but short T1 was better with manual delivery (long T1 mean 7.7, short T1 mean 8.7). Ratings were high with either T1 lag with automatic delivery (long T1 mean 9.3, short T1 mean 9.2).

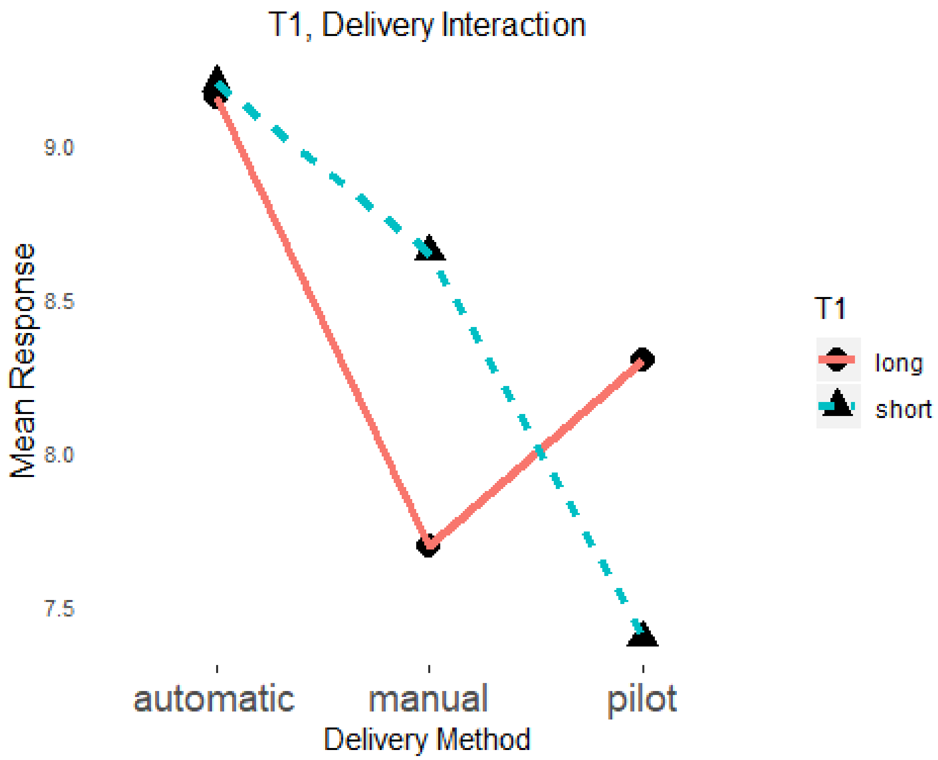


Figure 26. PSQ Situation Awareness for UAS on Lost Link Route ratings

5.2.4.4. Workload

Workload Assessment Keypad (WAK) responses

The T1 lag was not associated with higher WAK responses than baseline. In other words, controlling a lost link aircraft with an unknown intent did not reliably affect subjective workload.

NASA-TLX

We observed a slight trend for workload ratings to be higher after a long T1 compared to a short T1. Ratings of temporal demand were significantly higher (95% HDI [0.08, 0.61]; short T1 mean 4.3, long T1 mean 4.6) as were ratings of effort (95% HDI [0.08, 0.7]; short T1 mean 4.4, long T1 mean 4.6), although both effects are small. Ratings were also numerically but not statistically significantly higher for the mental demand and frustration components of the NASA TLX.

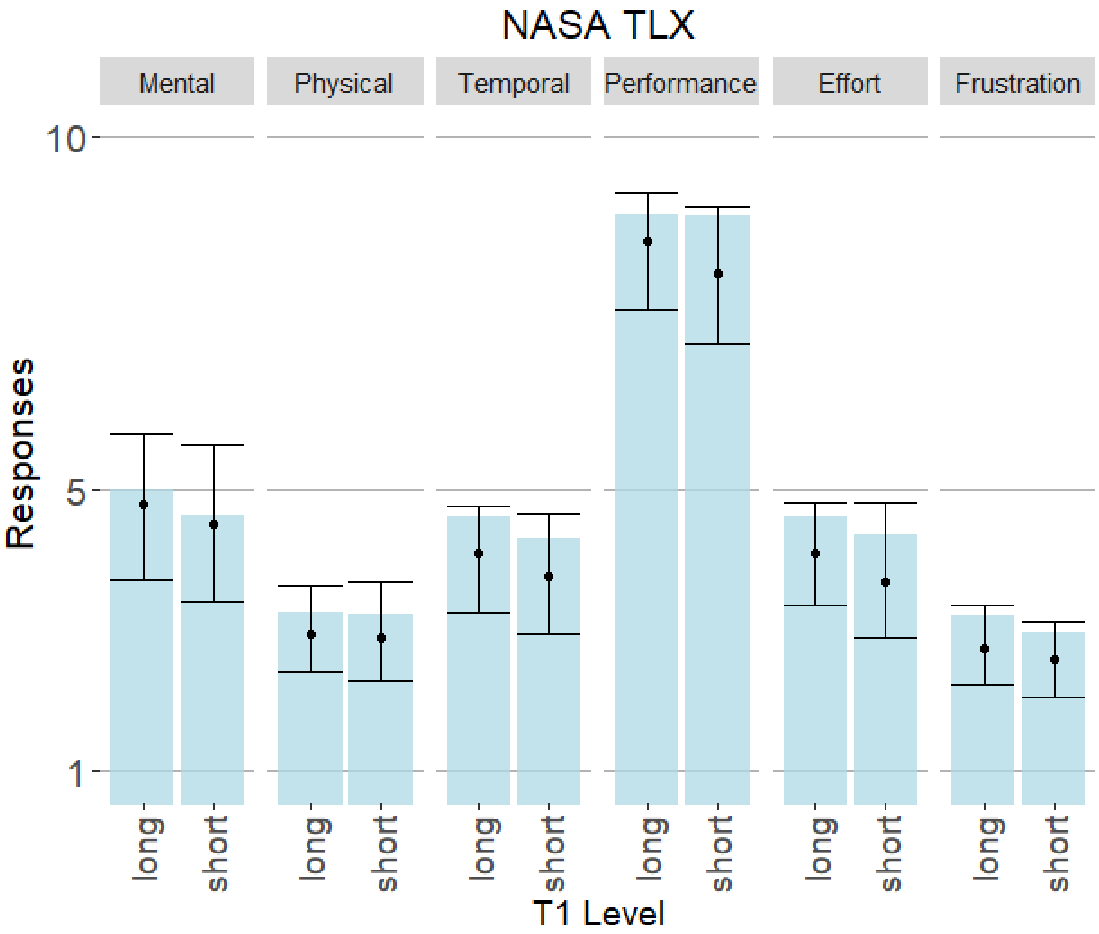


Figure 27. NASA TLX responses summarized by T1 lag length. The bar heights correspond to the mean of the participants' responses. The dot and error bars indicate the modeled mean and 95% HDI

Post-Scenario Questionnaires

We observed a significant interaction between T1 and T2 when participants were asked to rate their workload due to the lost link event (on the scale 1 = extremely low to 10 = extremely high) (95% HDI [0.09, 2.26]). Short T1 had slightly higher workload if T2 was short (long T1 mean 4.9, short T1 mean 5.4) but long T1 had slightly higher workload if T2 was long (long T1 mean 4.8, short T1 mean 4.4).

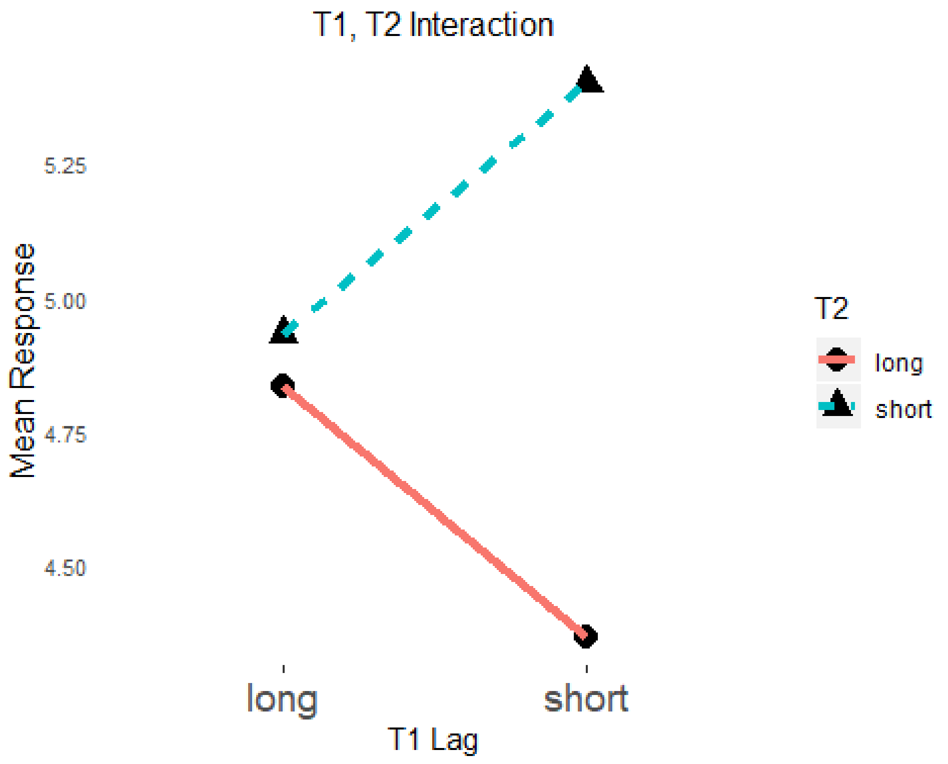


Figure 28. PSQ Workload Due to Lost Link ratings

5.2.4.5. Additional Questions

When asked on the PSQ to rate the duration of the T1 lag (1 meaning too short, 10 meaning too long), there was an interaction between T1 and T2 lag (95% HDI [0.1, 1.4]). The long T1 lag was rated as being somewhat too long, and equally so for both the short and long T2 conditions (short T2 mean 6.2, long T2 mean 6.3), while the short T1 was rated as being somewhat too short in the short T2 condition but about right in the long T2 condition (short T2 mean 4.5, long T2 mean 5.2).

There was also a main effect of T1 lag. Ratings were higher after a long T1 than a short T1 (95% HDI [0.5, 2.0]; long T1 mean 6.2, short T1 mean 5). This means that the long T1 was rated as being somewhat too long overall.

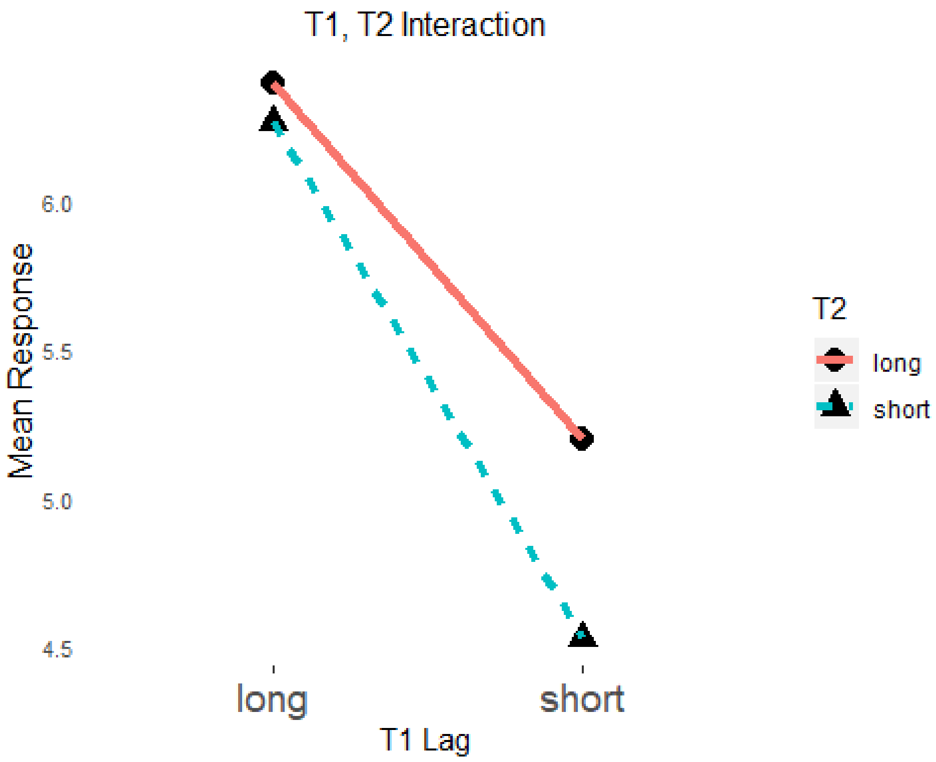


Figure 29. PSQ Rate the T1 Duration ratings

On the exit questionnaire, participants were asked to rate the impact (1 = very negative, 10 = very positive) of the short T1 lag. They gave a mean rating of 8.4, which was statistically significantly higher than a neutral rating of 5. They were also asked to rate the impact of the long T1 lag, giving a mean response of 3.7. This was statistically significantly lower than 5. Additionally, the difference between the two responses was statistically significant (95% HDI [3.3, 5.7]).

5.2.5. Factor 4: T2 Lag

The T2 lag is the amount of time between when the lost link route is provided to the participant and when the UAS makes its first maneuver off its previous route. Recall that there were three levels of T2 lag: baseline scenarios with no T2 value (either because no lost link occurred or because the UA continued on its previous route), scenarios with a T2 lag of 2 minutes (short), and scenarios with a T2 lag of 4 minutes (long). If a particular dependent variable or data source (e.g., post-scenario questionnaires) is not listed, then there were no statistically significant differences with regard to that variable. Please note that for non-normally distributed data (e.g., WAK data, survey responses; see Appendices C, D, and E), the confidence interval values may not be on the same scale as the data.

5.2.5.1. Safety

Post-Scenario Questionnaires

When participants rated the overall safety of operations (on the scale 1 = extremely low to 10 = extremely high), we observed a three-way interaction between T1, T2, and delivery method. If the lost link route was delivered automatically there was at best a small effect of T1 or T2 (all ratings 9.0 or higher); if the route was delivered manually there were opposite effects of T2 depending on T1 (long T2 better if T1 is short, long T2 mean 9, short T2 mean 8.7; short T2 better if T1 is long, long T2 mean 8.3, short T2 mean 8.7); and if the route was delivered by the pilot there was no effect of T2 when T1 was long (both means 8.5) but short T2 was better if T1 was short (long T2 mean 6.5, short T2 mean 7.3).

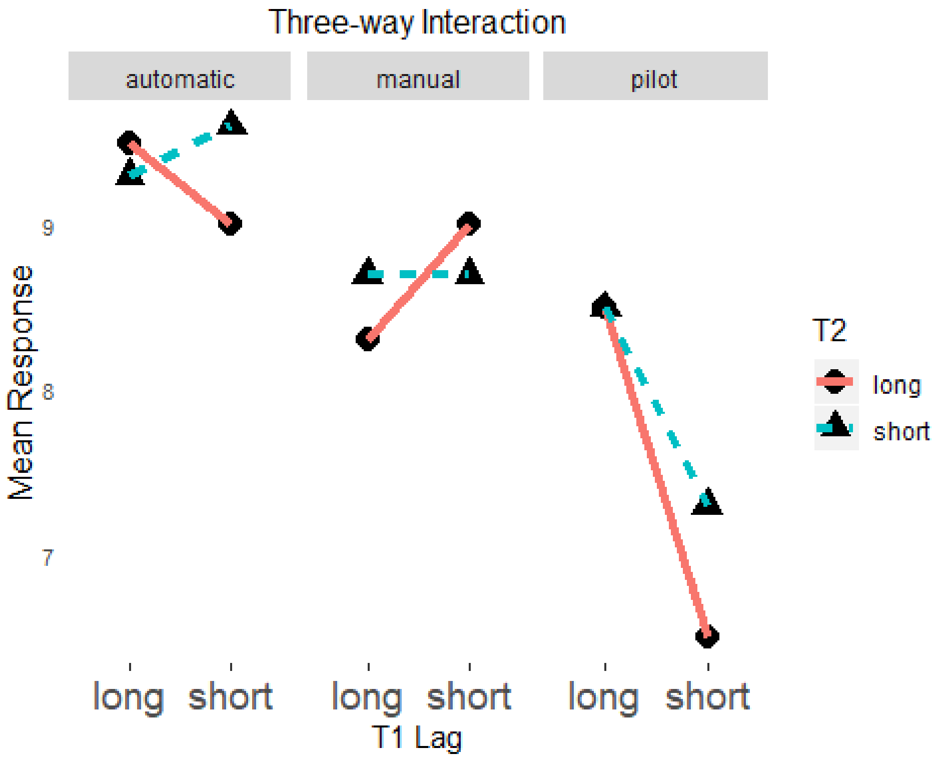


Figure 30. PSQ Overall Safety of Operations ratings

While complicated by the three-way interaction just described, we also observed a significant main effect of T2. The long T2 condition was rated as being less safe overall (95% HDI [0.04, 0.85], long T2 mean 8.4, short T2 mean 8.7).

When asked to rate how safely they resolved the lost link event (on the scale 1 = extremely low to 10 = extremely high), participants gave higher ratings after a long T2 than a short T2 (95% HDI [0.15, 0.94]; long T2 mean 9.0, short T2 mean 8.7).

5.2.5.2. Efficiency

Post-Scenario Questionnaires

When participants were asked to rate the overall efficiency of operations (on the scale 1 = extremely low to 10 = extremely high), we observed a significant interaction between T1 and T2

(95% HDI [0.5, 2.7]). If T1 is short there is little effect of T2 (short T2 mean 8.0, long T2 mean 8.1) but if T1 is long efficiency is rated higher when T2 is short (short T2 mean 8.8, long T2 mean 8.3).

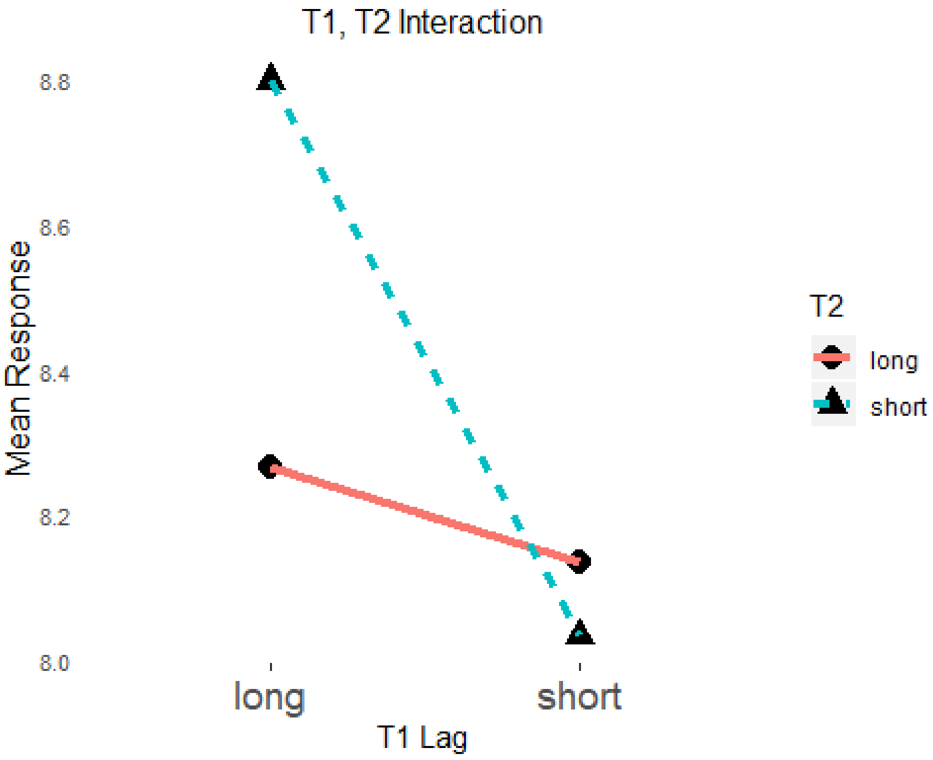


Figure 31. PSQ Overall Efficiency of Operations ratings

We observed a three-way interaction between T1, T2, and delivery method when participants were asked to rate how efficiently they resolved the lost link event (on the scale 1 = extremely low to 10 = extremely high).

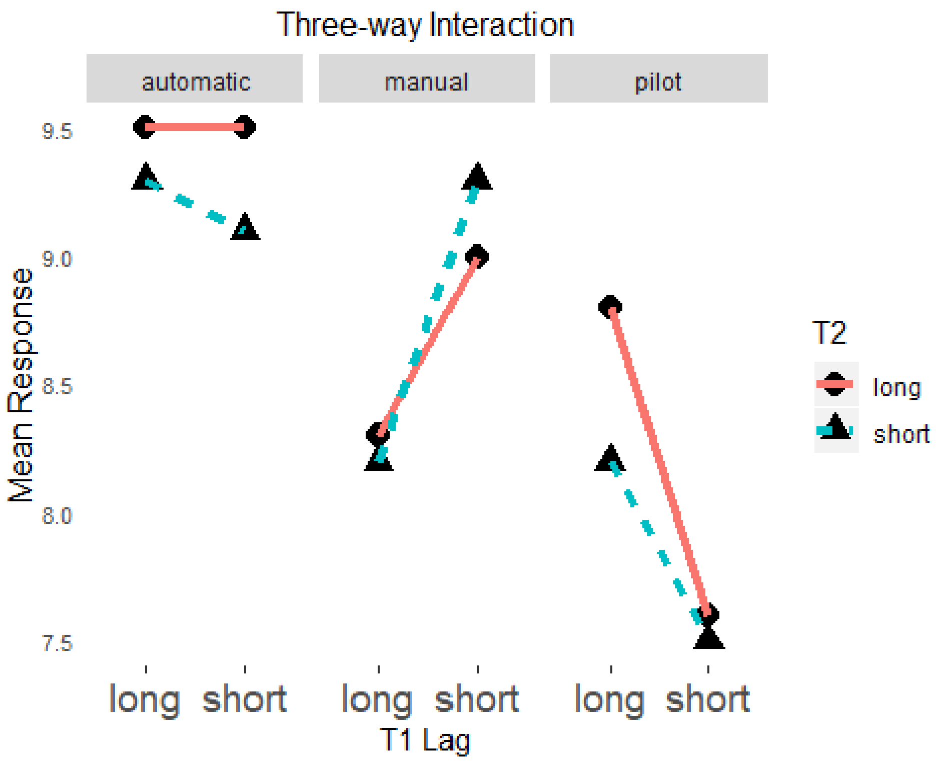


Figure 32. PSQ Efficiency in Resolving Lost Link ratings

Time and Distance Flown in Sector

We observed a small but statistically significant effect of T2 such that average flight distance was shorter when T2 is long (95% HDI [0.13, 0.45]; short T2 mean 75.2 nautical miles, long T2 mean 74.5 nm). There was an accompanying trend (91% confidence in difference; 95% HDI [-0.1, 0.55]) for average flight time to be shorter when T2 was long (short T2 mean 754 seconds, long T2 mean 748 seconds).

Time and distance effects were much more pronounced when limited to UAS. UAS average flight times were shorter by nearly 30 seconds when T2 was long (95% HDI [0.2, 0.75], short T2 mean 1129 seconds, long T2 mean 1103 seconds). This was matched by a similar effect for average flight distance (95% HDI [0.49, 0.67], short T2 mean 87.3 nautical miles, long T2 mean 85.6 nm). These results were not caused by the lost link aircraft as the time and distance effects are comparable or larger when the lost link UAS is removed from the data (difference of 3.1 nm average distance and 42 seconds average time). In fact, lost link aircraft showed the opposite effects numerically, with longer flight times and distances after a long T2 (difference of 10 nm average distance and 103 seconds average time).

As before, we note that while these effects were statistically significant, we do not consider them large enough to be practically significant.

5.2.5.3. Situation Awareness

Post-Scenario Questionnaires

We observed a significant interaction between T2 and delivery when participants were asked to rate their situation awareness for the UAS' progress along the lost link route (on the scale 1 = extremely low to 10 = extremely high) (95% HDI [0.02, 2.68]). When T2 was short there was a small difference between pilot and manual delivery and a large jump to automatic (pilot mean 8, manual mean 8.3, automatic mean 9.4) but when T2 was long there was a larger difference between pilot and manual but a smaller jump up to automatic (pilot mean 7.7, manual mean 8.1, automatic mean 9).

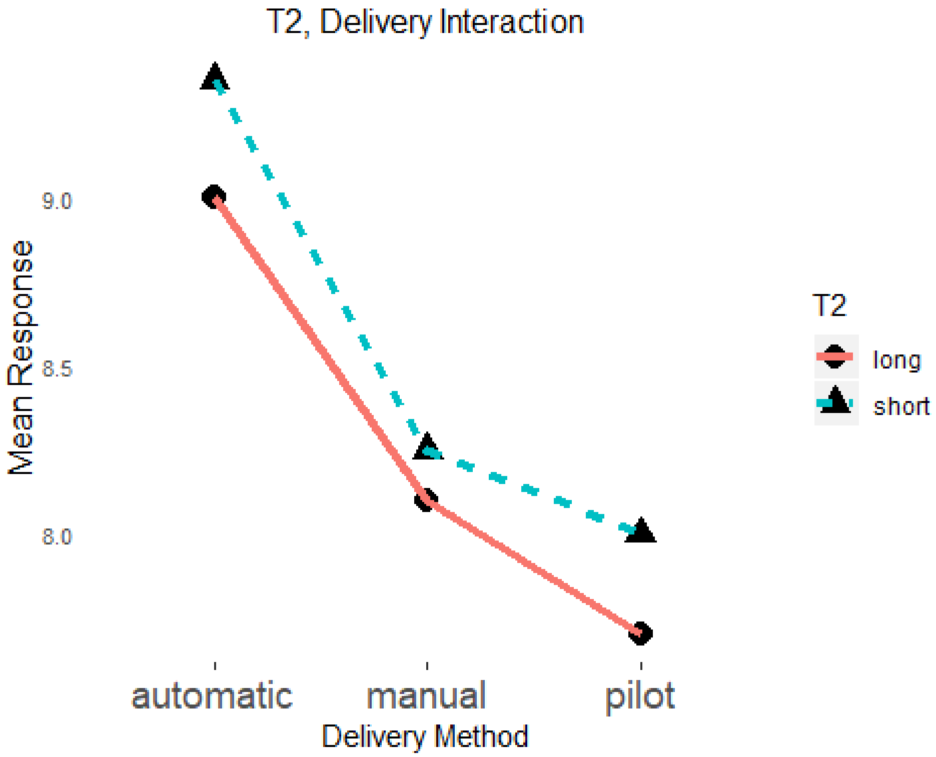


Figure 33. PSQ Situation Awareness for UAS Lost Link Route ratings

5.2.5.4. Workload

Workload Assessment Keypad (WAK) responses

The T2 lag was not associated with different WAK responses than baseline. However, we did observe significantly lower workload ratings after the UAS deviated and began following its lost link flight plan (95% HDI [-1.12, -0.04], 98% confidence), which suggests that resolving a lost link event is beneficial to participants.

Post-Scenario Questionnaires

We observed a significant interaction between T1 and T2 when participants were asked to rate their workload due to the lost link event (on the scale 1 = extremely low to 10 = extremely high) (95% HDI [0.09, 2.26]). There was little effect of T2 if T1 was long (short T2 mean 4.9, long T2 mean 4.8) but a higher response for short T2 than long T2 if T1 was short (short T2 mean 5.4, long T2 mean 4.4). There was also a main effect of T2 lag due to higher ratings when T2 was short (95% HDI [0.14, 0.85]; short T2 mean 5.2, long T2 mean 4.6).

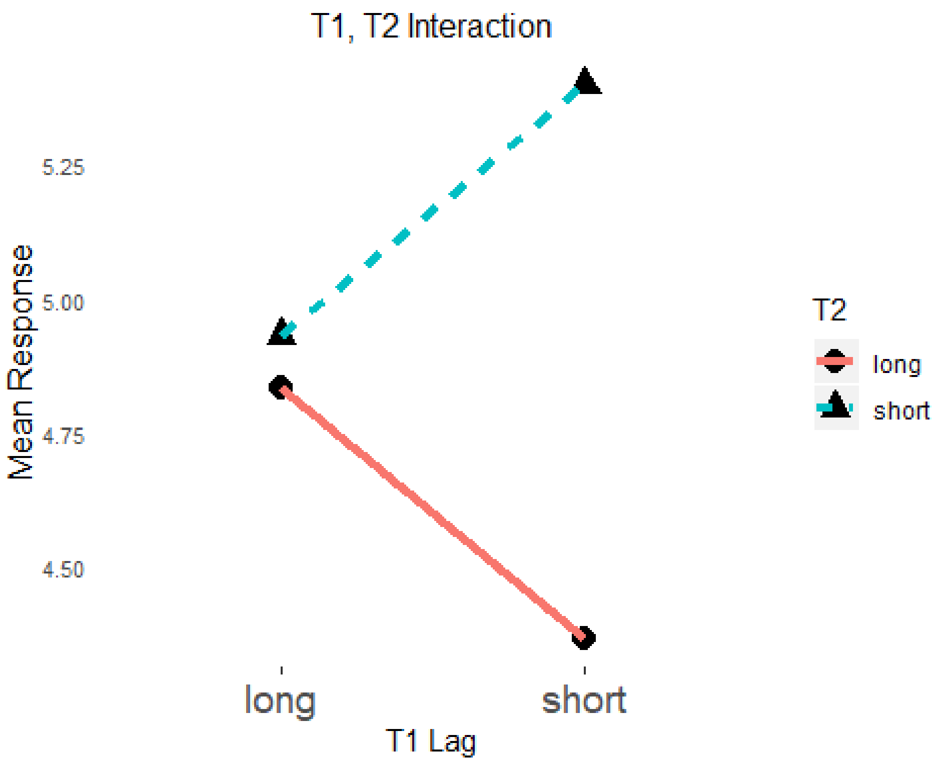


Figure 34. PSQ Workload Due to Lost Link ratings

5.2.5.5. Additional Questions

When asked on the PSQ to rate the duration of the T2 lag (1 meaning too short, 10 meaning too long), there was an interaction between T1 and delivery (95% HDI [0.3, 1.5]). Participants rated the lag as a bit too long in the automatic condition if T1 was long, and as somewhat too short in the automatic and pilot conditions if T1 was short. There was also an overall main effect of delivery due to the relatively lower (i.e., somewhat too short) ratings for automatic (at trend) and pilot compared to manual (manual versus pilot 95% HDI [0.2, 1.1]; manual versus automatic 95% HDI [-0.04, 1.48]; manual mean 5.0, pilot mean 4.4, automatic mean 4.9).

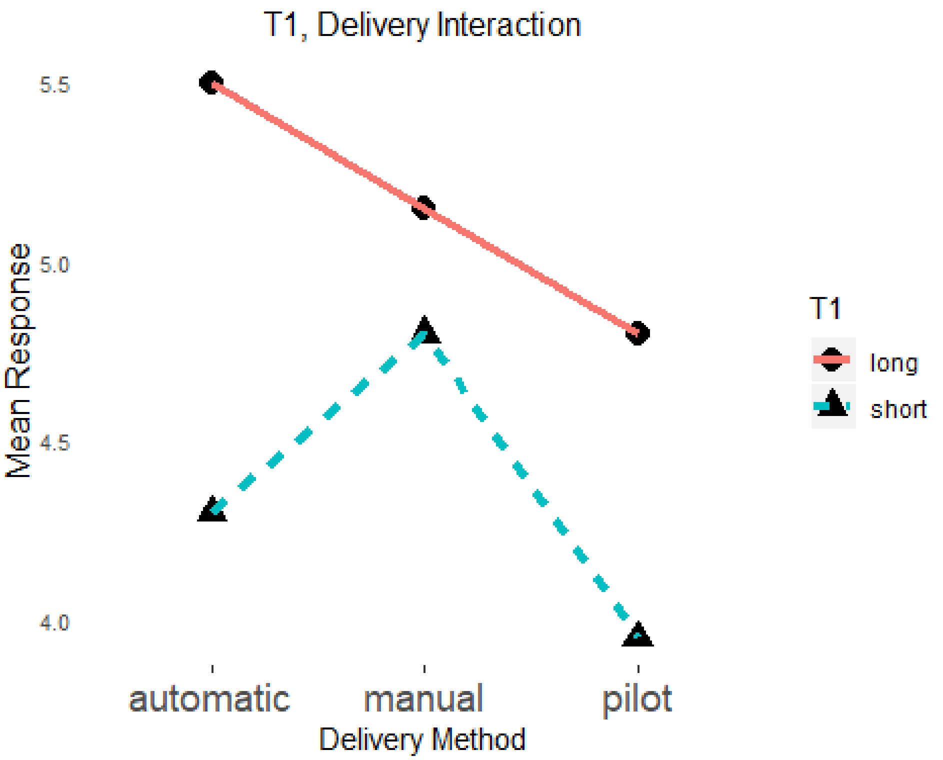


Figure 35. PSQ Rate Duration of T2 Lag ratings

On the exit questionnaire, participants rated the impact (1 = very negative, 10 = very positive) of the short T2 lag. They gave a mean rating of 3.8, which was close to (but only a trend) statistically significantly lower than a neutral rating of 5. They also rated the impact of the long T2 lag, giving a mean response of 5.4 that was not statistically different from 5. The difference between the two responses was not statistically significant.

5.3. POST-SIMULATION DEBRIEF SUMMARY

At the conclusion of all simulation scenarios each week, the researchers and participants held post-simulation debrief discussions. These debriefs gave the research team an opportunity to ask the participants questions about their overall feelings of the operations that they worked and about specific situations they may have encountered in the simulation. It also allowed the participants to offer their insight about the research objectives and to ask questions of the research team. Although these types of responses do not allow for any statistical analyses, they do offer some insight into the thoughts of the participants (see Appendix L). Notable participant comments are included throughout the document where they are relevant.

5.4. SUMMARY OF RESULTS

5.4.1. Factor 1: NAS Delivery Mechanism

The clearest result from the study is the large impact of delivery mechanism. We investigated three methods of delivering a lost link UAS' intent to ATCS: an automated, ABRR-like mechanism; a manual method where a researcher gave the ATCS a printed handout; and normal pilot-to-ATCS communication. Virtually all of the PSQ responses revealed a preference for automated delivery of intent over the other mechanisms, and most revealed a preference for the handout over pilot communication.

Amendment entry times were shortest when lost link intent was sent to ATCS via the automated mechanism and longest when delivered via pilot communication. Notably, amendment entry time was slow enough that some amendments were not even started by the time the UAS started turning in the pilot condition, and participants were often still adjusting the amendments after the turn in the pilot and manual conditions. In contrast, amendments were made very quickly in the automatic condition.

The automation preference in the PSQ results on efficiency, safety, and workload are bolstered by direct questions about preferences on the PSQ and exit questionnaire. When asked after each scenario to rate the effectiveness of the delivery method, all three methods were statistically significantly different from each other with a large preference for automatic. Participants also greatly preferred automatic delivery when rating the effectiveness of entering the lost link route. On the exit questionnaire, participants rated (on a scale of 1 = very ineffective to 10 = very effective) pilot delivery as slightly ineffective (mean response 4), manual delivery as somewhat effective (mean response 7.2), but automatic delivery as very effective (mean response 9.9). They also rated manual entry of the lost link route as slightly ineffective (mean response 3.7) but automatic entry as very effective (mean response 9.8).

The participants said, in comments on the questionnaires as well as in the debriefing, that the pilot delivery simply had the greatest chance of human error and led to too much heads-down time. The pilot of the lost link aircraft had to tie up the communications frequency to read off the new route and the participant had to confirm with a read-back; this was reflected in an increase in the number of push-to-talk events. There were also concerns about the possibility of the participant making an error while writing down the new route. Finally, the participant had to enter the new route as an amendment, which again required heads-down time and the chance of human error. The manual delivery method eliminated the pilot communication issues but still required amendment input. The automatic delivery method was very simple for the participant to use.

Overall, participants had a great preference for automatic delivery and entry of the lost link route. Manual delivery is less preferred but acceptable, and pilot delivery was least preferred but not unacceptable. As seen in the other results, delivery method was a big enough factor that it also interacted with the other factors of interest, for example changing preferences between a short or long T1 lag.

5.4.2. Factor 2: Communications Availability

We did not study two-way communications availability as a completely independent factor. For the scenarios in which the lost link intent was delivered via automation or a handout, communication between the UAS pilot and the participant was not available. When the lost link procedure was delivered via the pilot, communications were maintained. As such, the communication results overlap with those of delivery method: participants had a large preference for lost (unavailable) communications compared to maintained pilot communication. This would be a surprising result for normal air traffic operations, but (as noted above) the communications available condition has greater opportunities for UAS lost link route readback-hearback errors in comparison to the communications unavailable conditions. When asked directly on the exit questionnaire, participants gave only a slightly negative response when rating the impact of losing pilot communications (mean response 4.3 on a scale of 1 = very negative to 10 = very positive).

5.4.3. Factor 3: T1 Lag

The T1 lag was the period between the loss of link (when the UAS began squawking 7400) and the delivery of its intention to the participant. When directly comparing the two, or asked during the debriefing, participants rated the short lag (1 minute) as better than the long lag (4 minutes), wanting to know what the aircraft was going to do as soon as possible. However, this preference is qualified by ratings on the PSQ when the pilot delivered the lost link route; in this case, responses revealed that the short T1 lag was rated lower on safety, efficiency, and situation awareness.

Participants specifically rated the T1 lag on the exit questionnaire. They gave a positive rating (1 meant a negative impact and 10 meant a positive impact) to the impact of the short T1 lag (mean response 8.4) but a somewhat negative rating to the impact of the long T1 lag (mean response 3.7). Overall, while the participants generally preferred a shorter T1 lag, the long T1 lag was not unacceptable.

It is unclear why the PSQ ratings were higher for a long T1 with pilot delivery given the general preference for a short lag, although there are some hints. Some participants commented that the short T1 did not allow enough time to prepare after becoming aware of the 7400 squawk. The seven-element lost link routing was relatively long and contributed to significant frequency congestion during a busy portion of the scenario. The long T1 may have allowed the participant to become better prepared for the exchange of routing information that was about to take place.

5.4.4. Factor 4: T2 Lag

The T2 lag was the period between the delivery of a lost link UAS' intent to the participant and its deviation from its filed flight plan. Participants commented that they would prefer to have more time before the lost link aircraft turns (i.e., a long T2 lag) instead of less. More time would allow them to better manage traffic while entering the new route and coordinating with other sectors.

In their PSQ ratings, no clear preference for short versus long T2 lags emerged. Ratings of workload were higher and the safety of resolving a lost link were lower when the T2 lag was short (2 minutes). On the other hand, participants rated overall safety as being lower for the long T2 lag (4 minutes). Please see results in Section 5.2.5.1.

Participants also specifically rated the T2 lag after each scenario as well as in the exit questionnaire. Ratings of the short and long T2 lag after each scenario were not statistically different (long T2 mean 5.2, short T2 mean 4.4). Similarly, responses to the two lags on the exit questionnaire (from 1 = too short to 10 = too long) were not statistically different (long T2 mean 5.4, short T2 mean 3.8).

Some participants commented that the traffic patterns in the study did not pressure them as much as they could in the field, which may explain why their stated preferences were different from their ratings.

5.4.5. Identification of UAS

Because of ATCS variability in usage of the data block's fourth line, we did not put a large emphasis on a special UAS indicator in this study (in contrast to the terminal airspace version of this research). We simplified the situation for participants by making all UAS fly as either FedEx or UPS aircraft. However, we did have a 'U' indicator in the data block that participants could refer to, and we used a blinking red 7400 indicator and 'LLNK' entry to denote when a UAS was lost link. Participants gave high ratings on the exit questionnaire when asked about the effectiveness of the display system for managing UAS operations (mean response 8.7) and the effectiveness of the 7400 squawk for making them aware of the lost link (mean response 8.9).

6. RECOMMENDATIONS

This section includes a summary and recommendations based on the results of the HITL simulation.

Referring back to our research questions:

1. How long can ATCS wait before receiving a lost link UAS' procedure during a lost link?
    a. En route ATCS could wait up to 4 minutes, but a shorter time is preferred

2. How long do ATCS need after receiving the lost link procedure to be prepared (e.g., clear traffic, coordinate with other sectors) for the aircraft to maneuver off the previous route?
    a. En route ATCS need at least 4 minutes after receiving the lost link procedure to be prepared for the aircraft to maneuver off the previous route; however, longer would be better, especially in high/complex traffic environments

3. Given questions 1 and 2, what kind of automation is needed or would assist ATCS (i.e., which delivery mechanism is preferred)?
    a. The ABRR-like mechanism was greatly preferred by en route ATCS. Some kind of handout (e.g., the manual condition) is acceptable but may not be practical given desires for a short T1.

4. How does the availability of communications affect lost link operations?
    a. Available communications were a negative in this study due to the possibility of hearback-readback errors and extra time using the frequency compared to the

other two delivery methods. If lost link information was delivered by some other method, maintaining communications would likely be acceptable.

In addition, we propose the following takeaways based on the results of the HITL simulation:

- **A T1 time as long as 4 minutes is acceptable but shorter would be better.** However, the 1-minute lag posed difficulty for participants, potentially due to issues related to communicating with the pilot shortly after learning of the lost link.

- **A T2 time longer than 4 minutes is advisable**. Shorter times may be acceptable if automated delivery methods are used, but traffic density/complexity will be an important factor. Notably, the *average* amendment entry time was as long or longer than the short T2 when the lost link route was delivered manually or by the pilot, meaning that many amendments were not entered until after the UAS had already turned on to its new route. The longest amendment entry time was over 10 minutes but nearly all were completed in 7 minutes or less, suggesting that a T2 time in this range may give the best opportunity for ATCS to be prepared for the change in route.

- **Automated delivery of lost link route information would greatly reduce ATCS workload**, allowing them to work without a D-side (in the context of our study). Manual delivery, particularly if a D-side is present, would also be acceptable.

- **Automation should include some manner of capturing ATCS attention.** A beep or blinking indicator in addition to the T marker used in the study would be beneficial. One participant never entered the amendment on an automatic delivery trial because he did not notice the T marker, and some participants commented that in the field they would like to be explicitly told that a reroute is available.

- **The 7400 squawk and flashing LLNK indicator are effective for bringing attention to a lost link aircraft**, but participants commented that it would likely benefit from additional saliency (see Appendix L). A change in color may be useful.

6.1. RECOMMENDATIONS FOR FUTURE RESEARCH

As with any study, the results of this HITL simulation are based solely on the design as described in the report. Some parameters such as traffic level and pilot responses are variable and could lead to different recommendations if they served as the focus of study. The following are suggestions for topics that could be explored in future research:

- Traffic levels and complexity are an important factor in determining what T1 and T2 times would be acceptable. Additional research should examine more complex traffic than used in this study.

- Related, the current study only ever had a single lost link UAS at a time. Additional research should examine if multiple lost link UAS, especially if they have different and/or conflicting procedures, cause increased problems for ATCS.

- Additional display features should be tested to see if they notably improve ATCS response times to important events, such as a UAS going lost link or new routing information becoming available.

- Additional lost link procedures should be tested. In the previous terminal airspace study, lost link procedures were built into existing approach/departure procedures; a similar tact could be investigated for en route airspace. Procedures could also include altitude changes (which were avoided in this study but resulted in lost link UA flying at improper altitudes) and more severe turns such as for a return to base.

- Additional/alternative automation methods should be tested. We modified the current ABRR feature, but other methods for displaying a new route and entering it into ERAM could be better. In addition, practical matters like how the lost link information will be communicated to ATC systems need to be determined.

- On the UAS pilot side, studies should be done to determine how lost link information can be found quickly so that the pilot can communicate it to ATCS within the T1 window (e.g., via radio communication or some automated transmission), and how to determine if a maneuver initiated by air traffic control instruction will comply with lost link procedure restrictions (e.g., if a turn would result in the current lost link procedure sending the UAS through restricted airspace).

REFERENCES

Buondonno, K., Gilson, S., Pastakia, B., & Speulvado, L. (2012). *Multi-UAS Operational Assessment: Class D Airspace* (DOT/FAA/TC-TN 12/45). Atlantic City, NJ: Federal Aviation Administration William J. Hughes Technical Center

Guttman, J.A., & Stein, E.S. (1997). *En Route Generic Airspace Evaluation* (DOT/FAA/CT-TN 97/7). Atlantic City, NJ: Federal Aviation Administration William J. Hughes Technical Center

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.), *Human Mental Workload.* Amsterdam: North Holland Press.

McFarland, A.L. (1997). *A Conflict Probe to Provide Early Benefits for Airspace Users and Controllers* (Rep.). McLean, VA: Center for Advanced Aviation System Development, The MITRE Corporation.

Parasuraman, R., Sheridan, T., & Wickens, C. (2006). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, 30* (3), 286-297.

Pastakia, B., Konkel, A., Thompson, L., & Sollenberger, R. (2017). *Air traffic control receipt and display of contingency information*. Unpublished manuscript.

Pastakia, B., Won, J., Sollenberger, R., Aubuchon, D., Entis, S., & Thompson, L. (2015). *UAS Operational Assessment: Contingency Operations* (DOT/FAA/TC-TN 15/55). Atlantic City, NJ: Federal Aviation Administration William J. Hughes Technical Center

Rehmann, J.T., Stein, E.S. & Rosenberg, B.L. (1983). Subjective pilot workload assessment. *Human Factors*, *25*, 297-307.

Roscoe, A.H. & Ellis, G.A. (1990). *A Subjective Rating Scale for Assessing Pilot Workload in Flight: A decade of Practical Use* (RAE TR 90019). Farnborough, Hampshire, UK: Ministry of Defence

So, P. (2018, April 27). NASA TLX Task Load Index. Retrieval October 22, 2018, from https://humansystems.arc.nasa.gov/groups/TLX/

Stein, E.S., & Rosenberg, B.L. (1983). *The Measurement of Pilot Workload* (DOT/FAA/CT-82/23). Atlantic City, NJ: Federal Aviation Administration Technical Center

LIST OF ACRONYMS

ABRR	Airborne Reroute
ACL	Aircraft List
ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
ATCS	Air Traffic Control Specialist
C2	Command and Control
CFR	Code of Federal Regulations
CIC	Controller In Charge
CPC	Certified Professional Controller
COA	Certificate of Authorization
DAA	Detect and Avoid
DSR	Display System Replacement
EDST	En Route Decision Support Tool
ERAM	En Route Automation Modernization
FAA	Federal Aviation Administration
FDB	Flight Data Block
FLM	Front Line Manager
HITL	Human-in-the-Loop
IFR	Instrument Flight Rules
IVSR	Interim Voice Switch Replacement
NAS	National Airspace System
NATCA	National Air Traffic Controllers Association
NIEC	NextGen Integration and Evaluation Capability
NM	Nautical Miles
PIC	Pilot-in-Command
PSQ	Post Scenario Questionnaire
SME	Subject Matter Expert
STARS	Standard Terminal Automation Replacement System
TRACON	Terminal Radar Approach Control (facility)
UA	Unmanned Aircraft
UAS	Unmanned Aircraft System
URET	User Request Evaluation Tool
WAK	Workload Assessment Keypad
WJHTC	William J. Hughes Technical Center

APPENDIX A—1. INFORMED CONSENT

I, ______________________________, understand that this study entitled "Validation of Unmanned Aircraft Systems (UAS) Contingency Procedures and Requirements En Route Human-in-the-Loop (HITL) Simulation" is sponsored by the Federal Aviation Administration (FAA) and is being directed by Lacey Thompson.

Nature and Purpose

I have been recruited to volunteer as a participant in this research study. The purpose of this study is to conduct a HITL simulation to determine how UAS integration may affect air traffic operations in the en route environment. The results of the study will be used to identify future research and to help the FAA establish Air Traffic Control (ATC) standards and procedures for integrating UAS into the National Airspace System.

Experimental Procedures

The controllers will participate in three days of simulation from 8am to 4pm each day. On the first day, a group of researchers and ATC subject-matter experts will provide an overview of the project's research goals and familiarize participants with UAS operations. The participants will then complete a Background Questionnaire about their ATC experience. Next, researchers will provide the participants with guidelines of how the simulation will be conducted and participants will have an opportunity to familiarize themselves with the airspace and handle traffic in practice scenarios. Over the following three days of simulation, participants will work traffic scenarios, in which they will provide ratings of subjective workload. An automated data collection system will record system operations and generate a set of standard measures related to safety, efficiency, and communications. Scenarios will be audio-video recorded so that researchers can derive additional measures and reexamine any important events. At the end of each scenario, the participants will be asked to complete a Post-Scenario Questionnaire that will be used to gather their thoughts on the scenario just presented. Finally, after all the scenarios have been presented and discussed, participants will complete an Exit Questionnaire to gather their overall thoughts about their simulation experience.

Discomfort and Risks

I understand that I will not be exposed to any foreseeable risks or intrusive measurement techniques. I understand that I will not be exposed to any foreseeable risks beyond what I usually experience in my every day job.

Anonymity and Confidentiality

My participation in this study is strictly confidential. All information that I provide will be anonymized by the experimenters. All data collected in the study will be used for scientific purposes only and no names or identities will be released in any reports. Laboratory personnel will not disclose or release any Personally Identifiable Information (PII), except as may be required by statute. I understand that any Personally Identifiable Information [PII] will be protected according to FAA Order 1370.121 – FAA Information Security and Privacy Program & Policy.

Benefits

I understand that the only benefit to me is that I will be able to provide the researchers with valuable feedback and insight regarding UAS integration into the NAS. My data will help the FAA to identify the human factors issues with UAS integration and help the FAA to establish ATC standards and procedures for UAS integration.

Participant Responsibilities

I am aware that to participate in this study I must be an ATC professional. I will answer any questions asked during the study to the best of my ability. I will not discuss the content of the experiment with other potential participants until the study is completed.

Participant Assurances

I understand that my participation in this study is completely voluntary and I can withdraw at any time without penalty. I also understand that the researchers in this study may terminate my participation if they believe it is in my best interest. I understand that if new findings develop during the course of this research that may relate to my decision to continue participation, I will be informed. I have not given up any of my legal rights or released any individual or institution from liability for negligence.

Lacey Thompson has adequately answered all the questions I have asked about this study, my participation, and the procedures involved. I understand that Lacey Thompson or another member of the research team will be available to answer any questions concerning procedures throughout this study. If I have questions about this study or need to report any adverse effects from the research procedures, I will contact Lacey Thompson at (609) 485-8429.

Compensation and Injury

I agree to immediately report any injury or suspected adverse effect to Lacey Thompson at (609) 485-8429.

Audio-Video Recording

The entire simulation will be audio-video recorded in case the researchers need to review any events for data evaluation purposes. The researchers may show the simulation recordings to our study sponsors, but only with the explicit consent of the participants.

☐ Yes, I agree to allow my simulation recordings to be shown to sponsors.

☐ No, I do **not** agree to allow my simulation recordings to be shown to sponsors.

Signature Lines

I have read this informed consent statement. I understand its contents, and I freely consent to participate in this study under the conditions described. I understand that, if I want to, I may have a copy of this form.

Research Participant:____ Date: __________
Investigator: ____ Date: __________
Witness: ____ Date: __________

APPENDIX A – 2. BACKGROUND QUESTIONNAIRE

Instructions

The Background Questionnaire is designed to obtain information about your background and experience as an air traffic control specialist. The researchers will only use this information to describe the participants in this study as a group. Your identity will remain anonymous.

1. What is your **gender**?	☐ Male ☐ Female

2. What is your **age**?	______ years

3. What is the **En Route facility** where you work?	______ facility

4. How many **years of experience** do you have in ATC? (include all FAA and military experience)	______ years

5. How many **years of experience** do you have in **the En Route environment**?	______ years

6. How many of the **past 12 months** have you actively controlled traffic at your facility?	______ months

7. How **frequently do you work UAS** at your facility?	☐ I have **never** worked any UAS ☐ About 1 UAS every **year** ☐ About 1 UAS every **month** ☐ About 1 UAS every **week** ☐ About 1 UAS every **day** ☐ About 1 UAS every **hour** ☐ More than 1 UAS every **hour** How many? ____

8. Please list the **UAS types** that operate at your facility, if any e.g., RQ-4 Global Hawk, MQ-1 Predator

____

____

____

9. In your experience, **how do UAS operations affect ATC services in your sector**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ | None	Positive Effect

Comments:

Appendix A – 3. Post-Scenario Questionnaire

Instructions

The Post-Scenario Questionnaire is designed to draw upon your thoughts about the specific events presented in the scenario you just completed.

1. **Mental Demand** – How much mental and perceptual activity was required? (Did the scenario require a lot of thinking, deciding, calculating, remembering, looking, searching, etc.? Was the task easy or demanding, simple or complex, exacting or forgiving?)	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

2. **Physical Demand** – How much physical activity was required? (Did the scenario require any pushing, pulling, turning, controlling, activating, etc.? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?)	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

3. **Temporal Demand** – How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? (Was the pace slow and leisurely or rapid and frantic?)	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

4. **Performance** – How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)?	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

5. **Effort** – How hard did you have to work (mentally and physically) to accomplish your level of performance?	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

6. **Frustration** – How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the scenario?	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

7. Rate your **performance for safely resolving the lost link event(s)** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

8. Rate your **performance for efficiently resolving the lost link event(s)** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

9. Rate the **overall safety of operations** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

10. Rate the **overall efficiency of operations** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

11. Rate your **workload due to the lost link event(s)** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

12. Rate the **duration of the time interval** between **UAS reporting lost link** and the **lost link procedure being delivered** to you.	Too Short	①②③④⑤⑥⑦⑧⑨⑩	Too Long

Comments:

____

____

____

13. Rate the **duration of the time interval** between **lost link procedure being delivered** and the **UAS making its first diversion** from the previous route.	Too Short	①②③④⑤⑥⑦⑧⑨⑩	Too Long

Comments:

____

____

____

14. Rate the **effectiveness of the method for delivering the lost link procedure**.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

____

____

____

15. Rate **the effectiveness of the method for entering the lost link procedure**.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

____

____

____

16. Rate your **overall level of situation awareness** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

____

____

____

17. Rate your **situation awareness for UAS progress along the lost link contingency procedure** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

18. Rate the **overall performance of the simulation pilots** in terms of responding to control instructions, phraseology, and providing readbacks.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

19. What aspects of this scenario were hardest to work with? Why?

Comments:

20. Are there any display features or functions that could be provided to make this scenario more manageable? Please explain.

Comments:

21. Are there any procedures that could be provided to make this scenario more manageable? Please explain.

Comments:

22. Do you have any additional comments about the scenario just completed?

Comments:

Appendix A – 4. Exit Questionnaire

Instructions Please answer the following questions with regard to the entire series of UAS scenarios presented in this study.

1. Rate the **overall realism of the simulation experience** compared to actual ATC operations.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

2. Rate the **realism of the simulation hardware** compared to actual equipment.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

3. Rate the **realism of the simulation software** compared to actual functionality.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

4. Rate the **realism of the airspace** compared to actual NAS airspace.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

5. Rate the **realism of the simulation traffic scenarios** compared to actual NAS traffic.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Comments:

____

____

____

6. To what extent did the **WAK online workload rating technique interfere** with your ATC performance?	Not at All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

Comments:

____

____

____

7. How effective was the **airspace training**?	Extremely Ineffective	①②③④⑤⑥⑦⑧⑨⑩	Extremely Effective

Comments:

____

____

____

8. How effective was the **UAS training**?	Extremely Ineffective	①②③④⑤⑥⑦⑧⑨⑩	Extremely Effective

Comments:

____

____

____

9. How effective were the **procedures used to manage UAS operations**?	Extremely Ineffective	①②③④⑤⑥⑦⑧⑨⑩	Extremely Effective

Comments:

____

____

____

10. How effective was the **display system for managing UAS operations**?	Extremely Ineffective	①②③④⑤⑥⑦⑧⑨⑩	Extremely Effective

Comments:

11. How effective was the **7400 squawk in making you aware of the lost link**?	Extremely Ineffective	①②③④⑤⑥⑦⑧⑨⑩	Extremely Effective

Comments:

12. What was the impact of having the **shorter time interval (1 min)** between **UAS reporting lost link** and the **lost link procedure being delivered to you**?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

13. What was the impact of having the **longer time interval (4 min)** between **UAS reporting lost link** and the **lost link procedure being delivered to you**?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

14. What was the impact of having the **shorter time interval (2 min)** between **the lost link procedure being delivered** and the **UAS making its first diversion** from the previous route?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

15. What was the impact of having the **longer time interval (4 min)** between **the lost link procedure being delivered** and the **UAS making its first diversion** from the previous route?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

16. What was the impact of the **lost link procedure being delivered via the pilot (comms)?**	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

17. What was the impact of the **lost link procedure being delivered via a handout**?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

18. What was the impact of the **lost link procedure being delivered via automation**?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

____

____

____

19. What was the impact of having to **manually enter the lost link procedure**?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

____

____

____

20. What was the impact of having the **lost link procedure automatically entered**?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

____

____

____

21. What was the impact of **losing pilot communications** during a lost link event?	Extremely Negative	①②③④⑤⑥⑦⑧⑨⑩	Extremely Positive

Comments:

____

____

____

22. What suggestions do you have for effectively integrating UAS in the NAS?

Comments:

____

____

____

23. Do you have any suggestions regarding time intervals between UAS reporting lost link and the lost link procedure being delivered to you?

Comments:

24. Do you have any suggestions regarding time intervals between the lost link procedure being delivered and the UAS making its first diversion from the previous route?

Comments:

25. Do you have any suggestions regarding the method for delivering the lost link procedure?

Comments:

26. Do you have any suggestions regarding the method for entering the lost link procedure?

Comments:

27. Do you have any thoughts on when/how a lost link aircraft might change altitude? For example, shifting 1000 ft to correct for direction of flight, or descending into an alternate airport.

Comments:

APPENDIX B—WHY BAYESIAN STATISTICS AND GENERALIZED LINEAR MODELS

Takeaway points of this appendix:

- Various data collected during the HITL (e.g., survey responses, WAK responses, number of arrivals) are not normally distributed, so they should not be summarized or analyzed under the assumption of a normal distribution
- Generalized linear models, such as implemented in R or other statistical software, properly model non-normal data
- The output from Bayesian analyses better map onto researchers' expectations about significance and confidence values
- The practitioner faces overhead in learning software and better statistical techniques but is rewarded with better, more valid results

Human factors experimenters typically use frequentist statistical methods (t-tests, ANOVAs, etc.) to analyze data. In this report we have used different (but still frequentist) statistical methods, namely generalized linear models, in a Bayesian framework. Both of these choices provide a better description of the data and more intuitive results for decision-makers. Since statistical analysis exists as a way to assist decision making, we believe that this choice is the proper one.

Generalized Linear Models

The general linear model is a statistical model that says that outcome/dependent variable(s), **Y**, can be described as a combination of input/independent variables, **X**, and 'weights', **B**, along with some error term, **E**. Common statistical analyses such as multiple regressions, correlations, ANOVAs, and t-tests are examples of a general linear model where there is a single dependent variable and one or more independent variables. These models are 'linear' because Y is taken to be a linear combination of the Xs: if there are two independent variables, for example, the regression equation would be written as $Y = X0 + B1*X1 + B2*X2 + e$. B1 and B2 are the weights that are statistically tested to see if independent variables X1 and/or X2 affect the value of Y (X0 is an intercept term that is rarely of experimental interest and e is the error that helps describe how well the regression describes the data). These models are 'general' because they assume Y is normally distributed, or roughly shaped like a bell curve. General linear models are perfectly adequate for data like human heights or weights or, in the case of air traffic control experiments, variables like aircraft distance and time flown in sector.

However, not all data are normally distributed. Counts of events, such as airport arrivals or the number of losses of separation, tend to be log distributed: they can only be positive and tend to be positively skewed such that most events are bunched together at some counts but there are fewer events at higher counts. These kinds of variables are not necessarily well-described by a general linear model. Instead we can use generalized linear models, which still assume a linear association between Y and X but pass it through a distribution other than the normal (thus, the general linear model is an example of a generalized linear model where the distribution happens to be normal). By respecting the kind of data being described, generalized linear models provide a better description and more appropriate statistical analysis of the data.

Bayesian Statistics

Frequentist statistics, at a conceptual level, takes the view that experiments can (or hypothetically will be) repeated many/infinite times. The frequentist description of a coin having a 50% probability of coming up heads, for example, means that if you were to flip the coin a high number of times, you would observe equal numbers of heads and tails. When applied to experiment analysis, frequentist statistics are typically summarized with p values and, sometimes, confidence intervals. However, p values and frequentist confidence intervals are not intuitive to interpret and have a number of drawbacks.

The p value for a statistical test, such as the t value in a t-test, is defined as the probability of obtaining a result (t value) of the observed value or more extreme assuming that a null hypothesis is true (in the case of a t-test, usually that the difference in two means is zero). Rewording the definition slightly, the p value is the probability that the experimental data occurred given the null hypothesis. Reading the definition likely makes a couple of observations stick out: first, it relies on potential outcomes that are never actually observed (e.g. the 'more extreme' t values), and second, it assumes a null hypothesis that researchers rarely believe (e.g. that two conditions we are testing, likely because we expect them to differ, are actually the same). P values are involved in the creation of frequentist confidence intervals, which means that confidence intervals inherit their problems. Confidence intervals also have a non-intuitive meaning: a 95% confidence interval, for example, does not mean that there is a 95% chance that the 'real' (i.e., in the entire population) value falls within the interval. Instead, a 95% confidence interval means that if one were to conduct a large number of the exact same experiment and create a confidence interval for each, 95% of those confidence intervals would contain the 'real' value.

Bayesian statistics come from Bayes' rule, which connects the probability of an event to prior knowledge about that event. For example, one would think that the probability that it is going to rain is different depending on if it is cloudy or not. As it relates to statistical analysis Bayesian inference happens by multiplying a statistical model (such as a t-test or generalized linear model) by prior probabilities on the various elements of the model (such as the **B** weights) to provide an outcome. Practically speaking, this multiplication is carried out by software which provides many possible results via Monte Carlo random sampling. By looking at the outcomes of the random samples the analyst can decide how likely a particular outcome is. For example, if all of the random samples have a positive difference in means between two conditions, the analyst can assume that it is unlikely that the two conditions are the same. The results of frequentist analyses are an example of Bayesian analyses that use certain priors, in the same way that the general linear model is an example of a generalized linear model that uses a certain distribution.

The outcome of a Bayesian analysis is essentially the opposite of a p value: instead of the probability of data given a hypothesis, a Bayesian analysis provides the probability of the hypothesis given the data. This is what an analyst actually wants to know: given the data I've observed in my experiment, how likely is it that these two conditions differ? Additionally, the Monte Carlo random samples used in conducting a Bayesian analysis provide probabilities and confidence intervals that can be intuitively interpreted. If a Bayesian analysis results in 95% of the samples for some parameter (such as the difference of means between two conditions) being above zero, then the analysis says that there is a 95% chance that the 'real' difference is above zero. There is no need for considering infinite exact versions of the experiment or convoluted definitions; Bayesian outcomes align with what people typically expect statistical outcomes to mean.

It is important to note that Bayesian analysis is not a panacea. Bayesian methods can be used for simple procedures like correlations or t-tests; if these procedures are poorly chosen, such as when a generalized linear model would be more appropriate, then the outcome will still be less than optimal. Additionally, any statistical analysis is limited by the quality and quantity of the data available. If there is little data or it was collected in a haphazard way, the best statistical analysis could still be a poor, or just noisy, description of the real world. However, we believe that using generalized linear models and Bayesian inference provides the best opportunity for an accurate description of the data collected, and that the results will be more intuitive and suitable for assisting decision making on the study outcomes.

APPENDIX C—MODELING APPROACH

Modeling trial-level outcomes

For outcomes that are measured at the trial level (e.g., the number of departures), we will use a generalized multilevel model. We model the distribution for measure Y for trial i from participant j as:

$$Y_{ij} \sim g^{-1}(\eta_j + \boldsymbol{\beta} \cdot \mathbf{X_{ij}})$$

g is the link function, which will be chosen based on the support of Y; i.e., the type (integer vs. real) and range of values that the outcome can take on. Examples include the *identity* function for metric outcomes, and the *log* function for count outcomes.

$\mathbf{X_{ij}}$ is row i from the model matrix for participant j.

$\boldsymbol{\beta}$ is a vector of coefficients for "fixed" effects.

η_j is the second-level model for participant effects, which are modeled using the distribution:

$$\eta_j = \gamma_0 + \mu_j$$

γ_0 is the overall intercept.

μ_j is the random error component for participant j. I.e., it estimates the extent to which a given participant differs from the "average" participant.

Model matrix

We decided to model the scenario effects with multiple regression, such that each trial (e.g., short T1 – short T2 – manual delivery) had its own fixed effect $\boldsymbol{\beta}$. Each measure had at least 12 fixed effects, one for each experimental trial, and some included fixed effects for the baseline trials. It did not make sense to include fixed effects for the baseline trials for some measures, such as number of amendments since no amendments were made on those trials.

Post-hoc comparisons

The model fitting procedure allowed us to make *post-hoc* comparisons between specific types of trials. Following the study's hypotheses, we always performed the following comparisons:

- *Short T1 versus Long T1*
- *Short T2 versus Long T2*
- *Automatic delivery versus Manual delivery*
- *Automatic delivery versus Pilot delivery*
- *Manual delivery versus Pilot delivery*
- *Communications available versus Communications unavailable*

Link functions

The following link functions were used for the generalized linear model.

Measure	Link Function
Average distance in sector (all aircraft)	Identity
Average distance in sector (UAS)	Identity
Average time in sector (all aircraft)	Identity
Average time in sector (UAS)	Identity
Duration of communications (average)	Identity
Number of communications	Log
Number of sim pilot commands entered	Log
Number of amendments made	Log
Time from lost link route delivery to first amendment entry (or T clicked in automatic condition)	Log
Time from lost link route delivery to last amendment entry (or amendment accepted in automatic condition)	Log
Time between first and last amendment entry	Log

APPENDIX D—SURVEY ANALYSIS DETAILS

The survey data were modeled somewhat differently than the measures listed in Appendix C. This choice was made because survey responses are considered ordinal data and are relatively different from the other measures. Ordinal data means that while the response options are ordered (e.g. 3 is greater than 2 which is greater than 1), there is no reason to think that the difference between 1 and 2 is the same as the difference between 2 and 3, or that options 1 and 2 mean the same thing to different participants. Research shows that, while being common practice, it is a mistake to model ordinal data with methods that assume a normal distribution because those methods lead to increased erroneous results.

We modeled the survey data using a Bayesian ordered probit regression. The ordinal probit model assumes an underlying normally distributed variable that is then binned into the Likert-style response. For example, the model assumes that people have a continuous, normally distributed property called 'happiness', which they have to put into discrete bins when asked to rate how happy they are on a 1 to 7 scale. The model uses a mean response for each question and thresholds for dividing the response into each category (the point at which participants switch from responding '1' to '2' or '3' to '4'). That is to say, each survey question is allowed to have its own mean and standard deviation for the underlying distribution as well as its own thresholds for binning that distribution into a response. However, every question on a survey (either the exit survey or the post-scenario surveys taken as a whole) is fit by the model simultaneously, which allows them to be compared in the Bayesian framework. Participants are grouped into a single set of responses for each question, which is different from the model for the other data where participants were included as a second-level factor. Once the posterior probability has been calculated (the model has been fit to the data), statistical hypotheses are answered in a manner similar to bootstrapping. RStan creates a number of Monte Carlo draws of parameter estimates that fit the posterior probability distribution. If responses change for a question across scenarios, we can pull the Monte Carlo draws for those questions and compare the mean parameters. If, say, 95% of them exclude a 0 difference, then we can conclude with 95% confidence that the response changed.

After the model is fit to the data, comparisons of interest were made by creating an expected response for each question and then comparing the expected responses. We will walk through a result reported in the results section. Question 4 on the post-scenario questionnaire asked participants to rate how successful they were at accomplishing their task goals. In the short T1, short T2, manual delivery scenario, six participants responded with a 10 rating, two participants responded with a 9 rating, and one participant each responded with an 8 and 6 rating. The ordered probit model was fit to those data and provided a range of probabilities for any given rating (1 through 10) if the experiment were to be run again. It predicts virtually no chance of a 1 through 5 rating, a small range of probabilities of a 6 or 7 rating (around 0 to 5%), anywhere from ~5% to ~10% of a 8 rating, and higher probabilities of a 9 or 10 rating. These probabilities are then used to generate plausible response sets, each of which is then averaged to find an expected response. The model predicts that the expected response, were additional data to be collected, would likely be between 7 and 10. This prediction meshes with the actual expected response, which is the average of the responses from our participants (9.2).

In order to compare the three delivery conditions, this process is carried out for each question that contributes to that comparison: each experimental scenario with automatic delivery contributes to the automatic average and similarly for the other two; the three baseline scenarios are not included. The range of expected responses for all versions of the automatic scenarios are then averaged together, as are the expected responses for all versions of the manual and pilot scenarios, and then they are subtracted from each other in pairs (pilot minus automatic, pilot minus manual, and manual minus automatic). We observed that over 95% of the expected responses were higher in the automatic condition than either the manual or pilot conditions. Thus we report a significant difference based on the 95% HDI that participants gave a higher rating in the automatic condition.

This process is carried out for each comparison of interest in both the exit and post scenario questionnaires. One exception is for some questions which are not meant to be compared to another but instead to provide its own direct answer. For example, Question 21 of the exit questionnaire asked participants to rate the impact of losing pilot communications during a lost link. This is a self-contained question where participants provided a rating from 1, meaning very negative, to 10, meaning very positive, with 5 being a neutral response. We are interested in knowing what rating we would expect participants to give in general (i.e., if we were to collect more data) and if they believed lost communications was a negative (i.e., if the response is less than 5 on average). For questions like these, we simply take the various expected responses and see how often they fall above 5. The model predicts an expected response anywhere from around 3 to 5, although likely around 4 (which fits with our observed data, which had an average of 4.3). 88% of the Monte Carlo draws of the expected response are below 4.5 (we would consider any average between 4.5 and 5.5 to essentially be neutral, so we look below 4.5 or above 5.5), so we say that we are 88% confident that participants rated lost communications as negative.

APPENDIX E—WAK ANALYSIS

1. Modeling Workload

We modeled the *workload* construct using a two-level model.

1.1 Level 1

The first level of the model predicts that the unobserved variable 'workload' y at time t on trial i for participant p will be:

$$y_{itp} = \mathbf{x}_{\mathbf{itp}}^{\mathbf{T}} \boldsymbol{\beta}_{\mathbf{p}} + \varphi y_{i,t-1,p}$$

Response-level explanatory variables ($\mathbf{x}_{\mathbf{itp}}$) are described below.

This model captures the lag-1 autoregressive effect of the response variable, (φ). The inclusion of this term corrects for the violation of independence of samples over time, motivated by an analysis of WAK data from a previous study.

1.1.1 Level 1 Predictors

Each run is partitioned into 10 s *epochs* in which various objects and events are counted (i.e., the number of aircraft in the sector, the number of communications made by the participant). E.g., the first epoch will cover the time period $(0,10]s$, the second $(10,20]s$, etc.

The vector of response-level variables $\mathbf{x}_{\mathbf{itp}}$ is composed of the following elements. Some variables were normalized to a mean of 0 and standard deviation of 1; these and are indicated below.

1. Run number: linear component {1, 2, 3, …, 14} (normalized)
2. Run number: quadratic component {1, 4, 9, …, 196} (normalized)
3. Number of communications by the participant during the prompt epoch {0, 1, 2, …} (normalized)
4. Number of communications from pilots to the participant during the prompt epoch {0, 1, 2, …} (normalized)
5. Total number of aircraft in the sector during the prompt epoch (including UAS) {0, 1, 2, …} (normalized)
6. Number of UAS in the sector during the prompt epoch {0, 1, 2, …} (normalized)
7. Whether a LL UAS is present that has not yet notified the ATC of its intent {0, 1}
8. Whether a LL UAS has notified the ATC of its intent but not started its turn {0, 1}
9. Whether a LL UAS is following a LL flight plan that deviates from its original plan {0, 1}
10. Whether LL intent was delivered since last prompt via automated ABRR-like method {0, 1}
11. Whether LL intent was delivered since last prompt via manual method {0, 1}
12. Whether LL intent was delivered since last prompt via pilot communication {0, 1}

1.2 Level 2

The second level of the model predicts effects for each participant, p. Each element n in the vector $\boldsymbol{\beta}_{\mathbf{p}}$ is distributed as:

$$\beta_{pn} \sim \mathcal{N}(\gamma_{0n}, \sigma_{\beta n}^2)$$

This form allows the parameters γ_{0n} to estimate the average coefficient across all controllers. For example, $\gamma_{0,3}$ is the estimates the effect of increasing pilot communications on workload (corresponding to number 3 in the level 1 predictors list).

1.2.1 No Interactions

This model leaves out potentially interesting interactions. There are already quite a few parameters to estimate for the amount of data available, so any interaction coefficients will be too noisy to be meaningful.

2. Responses

We cannot directly observe the "true" workload of a participant, modeled above as y_{itp}; we see the numeric WAK response (and whether the participant made one at all) and can measure the response time.

2.1.1 Numeric Responses

The numeric response π_{itp} is modeled with a cumulative link model:

$$\pi_{itp} = \begin{cases} 1 \ if \ y_{itp} \leq \theta_1 \\ 2 \ if \ \theta_1 < y_{itp} \leq \theta_2 \\ 3 \ if \ \theta_2 < y_{itp} \leq \theta_3 \\ \vdots \\ 10 \ if \ \theta_9 < y_{itp} \end{cases}$$

We will use a probit link function for this model; therefore:

$$P(\pi_{itp} = k) = \Phi(\theta_k - y_{itp}) - \Phi(\theta_{k-1} - y_{itp})$$

Where $\theta_1, \dots, \theta_9$ are estimated from the data.

2.1.2 Missing Responses

The probability that no response was made on a given trial, ε_{itp}, is modeled by:

$$E(\varepsilon_{itp}) = logit^{-1}(\alpha_0 + \alpha_1 y_{itp})$$

Both α_0 and α_1 are estimated from the data. This is motivated by the assumption that higher workload will increase the probability of missing a prompt, although the strength of this effect was measured.

2.1.3 Response Time

We make no predictions about the relationship between workload and response time here; inspection of previous results showed a weak relationship between WAK values and RT.

3. Priors

Priors on the effects of interest ($\boldsymbol{\gamma}, \boldsymbol{\alpha}$) will generally have a Cauchy prior distribution with location 0 and scale 1, e.g.:

$$\alpha \sim Cauchy(0,1)$$

3.1 Autoregression

The prior for the autoregressive effect on workload is a truncated normal in the range [-1, 1]:

$$\varphi \sim \mathcal{N}(0,1), -1 \leq \varphi \leq 1$$

3.2 Thresholds

Thresholds $\theta_1, \theta_2, \ldots, \theta_9$ are constrained to be monotonically increasing. Each has the prior:

$$\theta_k \sim \mathcal{N}(k-5,2)$$

3.3 Variance

Variance for the participant-level model on the elements of $\boldsymbol{\beta}$ has a half-normal prior distribution:

$$\sigma_{\beta n} \sim \mathcal{N}(0, 10^2)$$

4. Data Analysis

This experiment was designed to examine the effect of three factors on performance:

1. T1 time, the time between the aircraft losing link and reporting its intention.
2. T2 time, the time between the aircraft reporting its intention and beginning its maneuver.
3. Delivery mechanism, the method by which the participant became aware of lost link intention.

4.1 Timing Effects

For sequential workload assessments, the effects of T-times are captured by predictors that indicate the state of the aircraft (#7 and #8 in the list of level 1 predictors). If these predict higher workload, then longer T-times will increase workload.

4.2 Delivery Effects

Three predictors serve as indicator variables for the delivery of lost link intent through different mechanisms (#10 - 12 in the list of level 1 predictors). We'll perform pairwise comparisons of their estimated coefficients to determine which results in highest and lowest workload.

4.3 Baseline Trials

This experiment featured two “baseline” conditions. In one, no lost link occurred; the above predictors of interest are 0 throughout these trials, allowing us to obtain better estimates for each of the “nuisance” regressors (e.g., the effects of run order or the number of communications).

In the second baseline condition, a lost link did occur; after a T1 delay the participant was informed that the UAS would continue on its flight plan. This represents a new condition and could be modeled directly. However, since this is of little experimental interest and there are few WAK responses recorded for this condition across the experiment, we’ve excluded these from the analysis. The samples recorded before the lost link intent was delivered are included in the model.

APPENDIX F—POST-SCENARIO QUESTIONNAIRE COMPLETE RESULTS

The post-scenario questionnaire asked a series of questions where the participant self-rated various aspects of the scenario (i.e. their mental workload, safety of operations). Below are mean and median responses for each of the post-scenario questions.

1) Mental Demand

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	5.0	5.0	10
long	long	ABRR	no	3.3	3.0	10
long	long	pilot	yes	6.1	6.5	10
long	short	manual	no	5.2	6.0	10
long	short	ABRR	no	5.1	5.0	10
long	short	pilot	yes	5.3	5.5	10
long	NA	pilot	yes	3.6	3.0	10
short	long	manual	no	4.5	5.0	10
short	long	ABRR	no	3.9	3.0	10
short	long	pilot	yes	4.9	5.0	10
short	short	manual	no	5.3	5.0	10
short	short	ABRR	no	4.0	4.0	10
short	short	pilot	yes	5.2	6.0	10
short	NA	pilot	yes	3.6	3.0	10
NA	NA	NA	NA	3.7	3.5	10

2) Physical Demand

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	3.6	2.5	10
long	long	ABRR	no	2.1	2.0	10
long	long	pilot	yes	3.3	3.0	10
long	short	manual	no	3.6	3.0	10
long	short	ABRR	no	2.6	2.0	10
long	short	pilot	yes	4.3	4.0	10
long	NA	pilot	yes	2.2	2.0	10
short	long	manual	no	3.4	3.0	10
short	long	ABRR	no	2.6	2.0	10
short	long	pilot	yes	3.3	2.5	10

short	short	manual	no	3.5	3.0	10
short	short	ABRR	no	2.8	2.0	10
short	short	pilot	yes	3.8	3.0	10
short	NA	pilot	yes	2.2	2.0	10
NA	NA	NA	NA	2.1	2.0	10

3) Temporal Demand

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	4.9	6.0	10
long	long	ABRR	no	3.2	2.5	10
long	long	pilot	yes	5.1	6.0	10
long	short	manual	no	5.1	6.0	10
long	short	ABRR	no	4.3	4.5	10
long	short	pilot	yes	5.1	5.5	10
long	NA	pilot	yes	3.6	3.0	10
short	long	manual	no	4.3	5.0	10
short	long	ABRR	no	3.9	3.0	10
short	long	pilot	yes	4.6	3.5	10
short	short	manual	no	4.7	5.0	10
short	short	ABRR	no	3.4	3.0	10
short	short	pilot	yes	5.0	5.0	10
short	NA	pilot	yes	3.6	3.0	10
NA	NA	NA	NA	3.2	3.0	10

4) Performance

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	8.3	9.0	10
long	long	ABRR	no	9.6	10.0	10
long	long	pilot	yes	9.0	9.5	10
long	short	manual	no	8.1	9.0	10
long	short	ABRR	no	9.5	10.0	10
long	short	pilot	yes	8.9	9.0	10
long	NA	pilot	yes	9.5	10.0	10
short	long	manual	no	9.2	10.0	10
short	long	ABRR	no	9.3	10.0	10

short	long	pilot	yes	8.4	9.0	10
short	short	manual	no	9.2	10.0	10
short	short	ABRR	no	9.0	10.0	10
short	short	pilot	yes	8.2	10.0	10
short	NA	pilot	yes	9.5	10.0	10
NA	NA	NA	NA	8.8	10.0	10

5) Effort

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	5.3	5.0	10
long	long	ABRR	no	3.1	3.0	10
long	long	pilot	yes	5.4	6.0	10
long	short	manual	no	4.5	5.0	10
long	short	ABRR	no	4.0	4.0	10
long	short	pilot	yes	5.4	5.5	10
long	NA	pilot	yes	3.7	4.0	10
short	long	manual	no	4.1	4.0	10
short	long	ABRR	no	3.2	2.5	10
short	long	pilot	yes	5.4	5.0	10
short	short	manual	no	4.4	3.5	10
short	short	ABRR	no	3.7	3.0	10
short	short	pilot	yes	5.3	5.5	10
short	NA	pilot	yes	3.7	4.0	10
NA	NA	NA	NA	3.1	2.5	10

6) Frustration

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	3.6	2.0	10
long	long	ABRR	no	1.6	1.0	10
long	long	pilot	yes	3.9	3.5	10
long	short	manual	no	3.1	3.0	10
long	short	ABRR	no	2.8	2.0	10
long	short	pilot	yes	4.3	2.0	10
long	NA	pilot	yes	1.9	1.5	10
short	long	manual	no	2.2	2.0	10

short	long	ABRR	no	2.2	1.0	10
short	long	pilot	yes	3.7	2.0	10
short	short	manual	no	2.9	2.0	10
short	short	ABRR	no	1.9	2.0	10
short	short	pilot	yes	4.9	4.0	10
short	NA	pilot	yes	1.9	1.5	10
NA	NA	NA	NA	1.2	1.0	10

7) Rate your performance for safely resolving the lost link events(s) during this scenario.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	8.4	8.5	10
long	long	ABRR	no	9.7	10.0	10
long	long	pilot	yes	9.0	9.5	10
long	short	manual	no	8.0	9.0	10
long	short	ABRR	no	9.5	10.0	10
long	short	pilot	yes	8.8	9.0	10
long	NA	pilot	yes	9.7	10.0	10
short	long	manual	no	9.4	10.0	10
short	long	ABRR	no	9.3	10.0	10
short	long	pilot	yes	8.0	8.5	10
short	short	manual	no	9.4	10.0	10
short	short	ABRR	no	9.4	10.0	10
short	short	pilot	yes	6.9	7.5	10
short	NA	pilot	yes	9.7	10.0	10
NA	NA	NA	NA	5.5	5.5	8

8) Rate your performance for efficiently resolving the lost link events(s) during this scenario.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	8.3	8.5	10
long	long	ABRR	no	9.5	10.0	10
long	long	pilot	yes	8.8	8.5	10
long	short	manual	no	8.2	9.5	10
long	short	ABRR	no	9.3	10.0	10
long	short	pilot	yes	8.2	8.5	10
long	NA	pilot	yes	9.7	10.0	10

short	long	manual	no	9.0	9.5	10
short	long	ABRR	no	9.5	10.0	10
short	long	pilot	yes	7.6	8.0	10
short	short	manual	no	9.3	10.0	10
short	short	ABRR	no	9.1	10.0	10
short	short	pilot	yes	7.5	8.5	10
short	NA	pilot	yes	9.7	10.0	10
NA	NA	NA	NA	5.5	5.5	6

9) Rate the overall safety of operations during this scenario.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	8.3	9.0	10
long	long	ABRR	no	9.5	10.0	10
long	long	pilot	yes	8.5	8.5	10
long	short	manual	no	8.7	9.0	10
long	short	ABRR	no	9.3	9.5	10
long	short	pilot	yes	8.5	9.0	10
long	NA	pilot	yes	9.8	10.0	10
short	long	manual	no	9.0	9.0	10
short	long	ABRR	no	9.0	10.0	10
short	long	pilot	yes	6.5	7.0	10
short	short	manual	no	8.7	9.5	10
short	short	ABRR	no	9.6	10.0	10
short	short	pilot	yes	7.3	8.5	10
short	NA	pilot	yes	9.8	10.0	10
NA	NA	NA	NA	9.8	10.0	10

10) Rate the overall efficiency of operations during this scenario.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	7.8	8.0	10
long	long	ABRR	no	9.2	10.0	10
long	long	pilot	yes	7.8	8.0	10
long	short	manual	no	8.6	8.0	10
long	short	ABRR	no	9.4	9.5	10
long	short	pilot	yes	8.4	8.5	10

long	NA	pilot	yes	9.7	10.0	10
short	long	manual	no	8.7	8.5	10
short	long	ABRR	no	8.6	9.5	10
short	long	pilot	yes	7.1	8.0	10
short	short	manual	no	8.3	8.5	10
short	short	ABRR	no	8.5	10.0	10
short	short	pilot	yes	7.3	8.5	10
short	NA	pilot	yes	9.7	10.0	10
NA	NA	NA	NA	9.8	10.0	10

11) Rate your workload due to the lost link events(s) during this scenario.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	5.8	6.0	10
long	long	ABRR	no	3.0	2.5	10
long	long	pilot	yes	5.7	6.0	10
long	short	manual	no	5.0	5.5	10
long	short	ABRR	no	3.8	3.0	10
long	short	pilot	yes	6.0	6.0	10
long	NA	pilot	yes	2.3	2.0	10
short	long	manual	no	4.4	4.5	10
short	long	ABRR	no	3.2	2.0	10
short	long	pilot	yes	5.5	5.0	10
short	short	manual	no	6.2	5.5	10
short	short	ABRR	no	3.9	3.5	10
short	short	pilot	yes	6.1	5.5	10
short	NA	pilot	yes	2.3	2.0	10
NA	NA	NA	NA	2.6	1.0	7

12) Rate the duration of the time interval between UAS reporting lost link and the lost link procedure being delivered to you.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	6.5	6.0	10
long	long	ABRR	no	6.1	5.5	10
long	long	pilot	yes	6.6	6.0	10
long	short	manual	no	5.6	5.0	10

long	short	ABRR	no	6.3	6.0	10
long	short	pilot	yes	6.9	6.5	10
long	NA	pilot	yes	4.5	5.0	10
short	long	manual	no	5.0	5.0	10
short	long	ABRR	no	5.3	5.0	10
short	long	pilot	yes	5.3	5.0	10
short	short	manual	no	4.8	5.0	10
short	short	ABRR	no	4.6	5.0	9
short	short	pilot	yes	4.2	5.0	10
short	NA	pilot	yes	4.5	5.0	10
NA	NA	NA	NA	2.0	1.0	4

13) Rate the duration of the time interval between the lost link procedure being delivered and the UAS making its first diversion from the previous route.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	6.1	5.0	10
long	long	ABRR	no	6.1	5.0	10
long	long	pilot	yes	4.8	5.0	10
long	short	manual	no	4.2	4.5	10
long	short	ABRR	no	4.9	5.0	10
long	short	pilot	yes	4.8	5.0	10
long	NA	pilot	yes	5.5	5.0	10
short	long	manual	no	4.8	5.0	10
short	long	ABRR	no	4.4	4.0	10
short	long	pilot	yes	4.2	5.0	10
short	short	manual	no	4.8	5.0	10
short	short	ABRR	no	4.2	4.5	10
short	short	pilot	yes	3.7	4.5	10
short	NA	pilot	yes	5.5	5.0	10
NA	NA	NA	NA	2.0	1.0	4

14) Rate the effectiveness of the method for delivering the lost link procedure.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	5.6	6.5	10
long	long	ABRR	no	8.8	9.0	10

long	long	pilot	yes	5.6	7.0	10
long	short	manual	no	4.3	5.0	10
long	short	ABRR	no	9.3	9.5	10
long	short	pilot	yes	4.3	3.0	10
long	NA	pilot	yes	6.9	6.5	10
short	long	manual	no	6.5	7.0	10
short	long	ABRR	no	8.7	10.0	10
short	long	pilot	yes	5.3	5.0	10
short	short	manual	no	5.8	7.0	10
short	short	ABRR	no	8.2	9.5	10
short	short	pilot	yes	4.8	4.0	10
short	NA	pilot	yes	6.9	6.5	10
NA	NA	NA	NA	2.0	1.0	4

15) Rate the effectiveness of the method for entering the lost link procedure.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	5.7	7.0	10
long	long	ABRR	no	9.0	9.5	10
long	long	pilot	yes	5.7	6.5	10
long	short	manual	no	4.6	5.0	10
long	short	ABRR	no	9.7	10.0	10
long	short	pilot	yes	4.2	3.5	10
long	NA	pilot	yes	5.4	5.0	10
short	long	manual	no	6.2	7.5	10
short	long	ABRR	no	9.1	10.0	10
short	long	pilot	yes	5.4	5.0	10
short	short	manual	no	5.5	6.5	10
short	short	ABRR	no	9.0	10.0	10
short	short	pilot	yes	4.8	5.0	10
short	NA	pilot	yes	5.4	5.0	10
NA	NA	NA	NA	2.0	1.0	4

16) Rate your overall level of situation awareness during this scenario.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	7.4	8.0	10

long	long	ABRR	no	9.2	9.5	10
long	long	pilot	yes	8.4	8.0	10
long	short	manual	no	8.8	10.0	10
long	short	ABRR	no	9.6	10.0	10
long	short	pilot	yes	7.9	7.5	10
long	NA	pilot	yes	8.9	10.0	10
short	long	manual	no	8.9	9.0	10
short	long	ABRR	no	9.3	10.0	10
short	long	pilot	yes	7.3	8.5	10
short	short	manual	no	8.5	10.0	10
short	short	ABRR	no	9.6	10.0	10
short	short	pilot	yes	7.3	8.5	10
short	NA	pilot	yes	8.9	10.0	10
NA	NA	NA	NA	9.4	10.0	10

17) Rate your situation awareness for UAS progress along the lost link contingency procedure during this scenario.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	7.3	7.5	10
long	long	ABRR	no	9.0	9.0	10
long	long	pilot	yes	8.3	8.0	10
long	short	manual	no	8.1	8.0	10
long	short	ABRR	no	9.3	9.5	10
long	short	pilot	yes	8.3	9.5	10
long	NA	pilot	yes	9.5	10.0	10
short	long	manual	no	8.9	9.0	10
short	long	ABRR	no	9.0	10.0	10
short	long	pilot	yes	7.1	7.5	10
short	short	manual	no	8.4	10.0	10
short	short	ABRR	no	9.4	10.0	10
short	short	pilot	yes	7.7	9.0	10
short	NA	pilot	yes	9.5	10.0	10
NA	NA	NA	NA	4.6	1.0	5

18) Rate the overall performance of the simulation pilots in terms of responding to control instructions, phraseology, and providing readbacks.

T1	T2	Delivery	Comms Available	Mean Response	Median Response	# Responses
long	long	manual	no	8.5	9.5	10
long	long	ABRR	no	9.2	10.0	10
long	long	pilot	yes	9.0	10.0	10
long	short	manual	no	8.6	9.5	10
long	short	ABRR	no	9.3	10.0	10
long	short	pilot	yes	8.7	9.5	10
long	NA	pilot	yes	9.5	10.0	10
short	long	manual	no	9.2	9.5	10
short	long	ABRR	no	9.2	10.0	10
short	long	pilot	yes	8.3	10.0	10
short	short	manual	no	9.5	10.0	10
short	short	ABRR	no	9.5	10.0	10
short	short	pilot	yes	8.4	9.5	10
short	NA	pilot	yes	9.5	10.0	10
NA	NA	NA	NA	9.7	10.0	10

APPENDIX G— EXIT QUESTIONNAIRE COMPLETE RESULTS

The exit questionnaire asked a series of questions where the participant self-rated various aspects of the study such as the realism of the HITL. Participants also answered questions that focused on important topics such as their preference for UAS lost link timing. Below are mean and median responses for each of the exit questions.

Question 1. Rate the overall realism of the simulation experience compared to actual ATC operations

mean_resp	median_resp	n_resp
8	7.5	10

Question 2. Rate the realism of the simulation hardware compared to actual equipment

mean_resp	median_resp	n_resp
6.7	7	10

Question 3. Rate the realism of the simulation software compared to actual functionality

mean_resp	median_resp	n_resp
7.7	7.5	10

Question 4. Rate the realism of the airspace compared to actual NAS airspace

mean_resp	median_resp	n_resp
8.9	10	10

Question 5. Rate the realism of the simulation traffic scenarios compared to actual NAS traffic

mean_resp	median_resp	n_resp
8.4	8	10

Question 6. To what extent did the WAK online workload rating technique interfere with your ATC performance?

mean_resp	median_resp	n_resp
2.2	2	10

Question 7. How effective was the airspace training?

mean_resp	median_resp	n_resp

8.5	8.5	10

Question 8. How effective was the UAS training?

mean_resp	median_resp	n_resp
8.7	8.5	10

Question 9. How effective were the procedures used to manage UAS operations?

mean_resp	median_resp	n_resp
7.2	7.5	10

Question 10. How effective was the display system for managing UAS operations?

mean_resp	median_resp	n_resp
8.67	10	9

Question 11. How effective was the 7400 squawk in making you aware of the lost link?

mean_resp	median_resp	n_resp
8.9	10	10

Question 12. What was the impact of having the shorter time interval (1 min) between UAS reporting lost link and the lost link procedure being delivered to you?

mean_resp	median_resp	n_resp
8.4	9.5	10

Question 13. What was the impact of having the longer time interval (4 min) between UAS reporting lost link and the lost link procedure being delivered to you?

mean_resp	median_resp	n_resp
3.7	4	10

Question 14. What was the impact of having the shorter time interval (2 min) between the lost link procedure being delivered and the UAS making its first diversion from the previous route?

mean_resp	median_resp	n_resp
3.8	4	10

Question 15. What was the impact of having the longer time interval (4 min) between the lost link procedure being delivered and the UAS making its first diversion from the previous route?

mean_resp	median_resp	n_resp
5.4	6	10

Question 16. What was the impact of the lost link procedure being delivered via the pilot (comms)?

mean_resp	median_resp	n_resp
4	3	10

Question 17. What was the impact of the lost link procedure being delivered via a handout?

mean_resp	median_resp	n_resp
7.2	7.5	10

Question 18. What was the impact of the lost link procedure being delivered via automation?

mean_resp	median_resp	n_resp
9.9	10	10

Question 19. What was the impact of having to manually enter the lost link procedure?

mean_resp	median_resp	n_resp
3.7	3.5	10

Question 20. What was the impact of having the lost link procedure automatically entered?

mean_resp	median_resp	n_resp
9.8	10	10

Question 21. What was the impact of losing pilot communications during a lost link event?

mean_resp	median_resp	n_resp
4.3	4.5	10

APPENDIX H. AMENDMENT TIMES

T1	T2	delivery	Comms Available	Ave Time to First Entry from end of T1 (sec)	# With at least 1 Amend	Ave Time to Last Entry from end of T1 (sec)	# With >1 Amend	Ave Time btwn First and Last Entry (sec)	# Amend	Average # Amend per Scenario
long	long	Auto	no	69.6	10	74.2	10	4.6	20	2.0
long	long	manual	no	37.5	10	90.3	10	52.8	27	2.7
long	long	Voice	yes	89.1	10	223.6	10	134.5	29	2.9
long	short	Auto	no	65.2	10	70.0	10	4.8	20	2.0
long	short	manual	no	41.3	10	101.5	10	60.2	25	2.5
long	short	Voice	yes	89.0	10	209.1	9	117.4	38	3.8
short	long	Auto	no	67.8	10	73.4	10	5.6	20	2.0
short	long	manual	no	50.4	10	83.3	9	29.9	23	2.3
short	long	Voice	yes	129.5	10	187.3	10	57.8	27	2.7
short	short	Auto	no	29.5	9	32.3	9	2.8	18	2.0
short	short	manual	no	42.2	10	150.1	10	107.9	42	4.2
short	short	Voice	yes	114.0	10	267.8	10	153.8	47	4.7

APPENDIX I. SIMULATION PILOT COMMANDS

T1	T2	delivery	Comms Available	Total Commands	# Scenarios	Ave. Commands Per Scenario
long	long	Auto	no	176	10	17.6
long	long	manual	no	212	10	21.2
long	long	pilot	yes	177	10	17.7
long	short	Auto	no	185	10	18.5
long	short	manual	no	170	10	17.0
long	short	pilot	yes	199	10	19.9
long	NA	pilot	yes	94	5	18.8
short	long	Auto	no	164	10	16.4
short	long	manual	no	177	10	17.7
short	long	pilot	yes	176	10	17.6
short	short	Auto	no	168	10	16.8
short	short	manual	no	178	10	17.8
short	short	pilot	yes	174	10	17.4
short	NA	pilot	yes	66	5	13.2
NA	NA	NA	NA	174	10	17.4

APPENDIX J. PUSH TO TALK

T1	T2	Delivery	Comms Available	Total Time	Total #	# Scenarios	Ave Time per Push	Ave # Push
long	long	auto	no	2805.1	693	10	4.05	69.3
long	long	manual	no	2817.7	665	10	4.24	66.5
long	long	pilot	yes	2883.8	713	10	4.04	71.3
long	short	auto	no	2990.3	725	10	4.12	72.5
long	short	manual	no	2765.4	697	10	3.97	69.7
long	short	pilot	yes	3047.5	763	10	3.99	76.3
long	NA	pilot	yes	1424.7	364	5	3.91	72.8
short	long	auto	no	2870.0	694	10	4.13	69.4
short	long	manual	no	2886.8	688	10	4.19	68.8
short	long	pilot	yes	3183.4	763	10	4.17	76.3
short	short	auto	no	2700.9	673	10	4.01	67.3
short	short	manual	no	2901.5	686	10	4.23	68.6
short	short	pilot	yes	3156.9	721	10	4.38	72.1
short	NA	pilot	yes	1195.8	294	5	4.07	58.8
NA	NA	NA	NA	2816.9	690	10	4.08	69.0

APPENDIX K. TIME AND DISTANCE

1) All Aircraft

T1	T2	Delivery	Comms Available	Total Distance (meters)	Total Time (sec)	Total Aircraft	Ave Distance Per Aircraft (meters)	Ave Time Per Aircraft (sec)
long	long	auto	no	35697321	192960	255	139989.5	756.7
long	long	manual	no	36263554	196379	265	136843.6	741.0
long	long	pilot	yes	36217377	196324	265	136669.3	740.8
long	short	auto	no	34932044	189106	255	136988.4	741.6
long	short	manual	no	35437009	191824	255	138968.7	752.2
long	short	pilot	yes	35748672	192875	255	140190.9	756.4
long	NA	pilot	yes	19050652	99933	135	141115.9	740.2
short	long	auto	no	35324564	191323	255	138527.7	750.3
short	long	manual	no	35227700	191184	255	138147.8	749.7
short	long	pilot	yes	35273500	191064	255	138327.5	749.3
short	short	auto	no	35721460	193688	255	140084.2	759.6
short	short	manual	no	35974638	194730	255	141077.0	763.6
short	short	pilot	yes	34844373	189727	252	138271.3	752.9
short	NA	pilot	yes	16688062	93837	120	139067.2	781.9
NA	NA	NA	NA	35889448	194541	255	140742.9	762.9

2) UAS Only

T1	T2	Delivery	Comms Available	Total Distance (meters)	Total Time (sec)	Total Aircraft	Ave Distance Per Aircraft (meters)	Ave Time Per Aircraft (sec)
long	long	auto	no	13981277	96970	85	164485.6	1140.8
long	long	manual	no	14375062	99647	95	151316.4	1048.9
long	long	pilot	yes	14378253	99670	95	151350.0	1049.1
long	short	auto	no	13326228	93551	85	156779.2	1100.6
long	short	manual	no	13701561	95582	85	161194.8	1124.5
long	short	pilot	yes	13934485	96678	85	163935.1	1137.4
long	NA	pilot	yes	6609314	44618	40	165232.8	1115.4
short	long	auto	no	13714859	95725	85	161351.3	1126.2
short	long	manual	no	13796020	96286	85	162306.1	1132.8

short	long	pilot	yes	13776999	96223	85	162082.3	1132.0
short	short	auto	no	14087564	97987	85	165736.0	1152.8
short	short	manual	no	14072803	97917	85	165562.4	1151.9
short	short	pilot	yes	13531340	95366	86	157341.2	1108.9
short	NA	pilot	yes	7491064	53398	45	166468.1	1186.6
NA	NA	NA	NA	14126343	98160	85	166192.3	1154.8

APPENDIX L. POST-SIMULATION DEBRIEF

The following table includes only some of the questions asked during the post simulation debrief. It is not the complete set of questions and responses discussed.

Was the T1 timing (1 minute/4 minutes) long enough?	
1	The shorter the better. Want to know where to turn the surrounding aircraft.
2	Want information as soon as possible. Time needs to coincide with a/c speed.
3	Quicker is better.
4	The quicker we get the new route, the better. Should be known to us before we get the reroute.
5	With 4 minutes, you are having to think.
Was the T2 timing (2 minutes/4minutes) long enough to allow you to prepare and react (i.e enter flight plan in the manual scenario) to the maneuver the UAS was going to execute?	
1	4 minutes is not long enough
2	For this airspace, there aren't any MOAs etc, so time may be long enough
3	Would want at least 2-3 minutes, so you can move traffic out of the way
4	Time depends on how information is delivered (manually takes more time)
5	The longer the better
6	Two minutes is too short
How much total time (T3) should be allowed between when the UAS squawks 7400 and when the UA maneuvers off the original flight path? (short/short = 3 minutes, short/long = 5 minutes, long/short = 6 minutes, long/long = 8 minutes)	
1	8 minutes is too long; 5 minutes is okay
2	5 to 8 minutes; might need a little more time
IF traffic levels were higher, would these times still work?	
1	Yes. If it were my airspace I would know the area
Did you have a preference as to how the route is given to you (i.e. handout/have to enter manually, ABRR, voice)?	
1	Prefer ABRR, manual, voice
2	Prefer for the pilot not to tie up the frequency; also there could be times when the pilot confuses the identifier
3	ABRR, manual, pilot
4	ABRR, manual, pilot
5	ABRR is easiest; I don't mind talking to the pilot, but it takes up your frequency
6	ABRR, handout, pilot

7	ABRR, manual, pilot
Do you feel the instances in which UAS lost comm were more difficult than those in which comm was maintained?	
1	No, not necessarily. As long as I knew what the aircraft was going to do, I don't need to talk to the pilot
2	Not really
3	Communication with UAS can sometimes have a delay and has a lot of interference. If VOIP worked, it would not be an issue.
4	No
5	Didn't really make a difference; would rather have pilot not take up our frequency
Do you have any recommendations for automation that could help mitigate UAS lost link?	
1	If the UAS could sent the flight plan directly to ATC (if possible)
2	Controller Pilot Data Link Communication (CPLC) could be very helpful. Maybe it wouldn't feel like we are talking to them in a tin can
3	T-route (even if you do have comms)
4	Automation reduces workload/head down time
Was the blinking LLNK (to signify lost link) sufficient or is there some indicator you would prefer to see or that would be helpful in identifying the UAS is lost link?	
1	Flashing was distracting; would like to be able to suppress
2	Change data block to a different color (i.e. red) if possible to get your attention really quickly (in EDST, yellow gets your attention)
3	Could be a little distracting (can encode an option to manually suppress the flashing)
4	LLINK doesn't catch your attention; didn't flash some a/c because wanted the UAS to stand out
5	Make the data block for UAS to stand out; not necessarily when it loses link
6	The whole data block may flash; it gets your attention
Did you treat the UAS differently from manned aircraft?	
1	No, not really
2	Put a J-ring around the UAS just to keep track of it; I would have highlighted it (but could not because of sim limitations)
3	No
Is it more appropriate for a UAS to turn off the route at a specific time instead of a fix?	
1	A fix
2	Fix; I don't want to have to think about time. I know exactly where the aircraft is. Feels like you have more control that way.

3	Fix; easier to know aircraft is going to turn at a certain point and not to have to calculate the distance
Based on the potential amount of limited number of lost link events (i.e. UAS won't lose link so frequently) could you live with these procedures the way they are?	
1	Take the pilot out of it; frequency congestion and delays in reading back info to you
2	Wouldn't say they are unacceptable
How did you enter the route?	
1	Usually have a D-side. After you write down your route, it takes times to enter it in.
2	Entered route through amendment by numbers; avoided ACL because it took time away to look at the menu options and buttons.
3	I didn't feel stress to enter the route in, because the aircraft will do what it's going to do anyway. I updated it when I could. I didn't think it was a priority.
4	If I wasn't talking to the PIC, I knew I would receive the info on the URET
5	Since I'm not familiar with the airspace, I looked at maps and how the aircraft is going to exit the airspace. Did route display feature if I wasn't sure, but I usually got it quick enough.