NASA/TM—2015–218839

Measured Response for UAS Integration into the National Airspace System

Kim-Phuong L. Vu
Dan Chiappe
Thomas Z. Strybel
California State University, Long Beach

Lisa Fern
Ohio State University

Conrad Rorie
Vernol Battiste
San Jose State University Foundation

Robert J. Shively
NASA Ames Research Center

April 2015

NASA STI Program…in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include creating custom thesauri, building customized databases, and organizing and publishing research results.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at *http://www.sti.nasa.gov*

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621-0134

- Phone the NASA STI Help Desk at (301) 621-0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076-1320

NASA/TM—2015–218839

Measured Response for UAS Integration into the National Airspace System

Kim-Phuong L. Vu
Dan Chiappe
Thomas Z. Strybel
California State University, Long Beach

Lisa Fern
Ohio State University

Conrad Rorie
Vernol Battiste
San Jose State University Foundation

Robert J. Shively
NASA Ames Research Center

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California

April 2015

Available from:

NASA Center for Aerospace Information
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

This report is also available in electronic form at http://www.sti.nasa.gov
or http://ntrs.nasa.gov/

Contents

Figures

Tables

Acronyms and Definitions

AC	aircraft
ADS-B	Automatic Dependent Surveillance-Broadcast
ADRS	Aeronautical Datalink and Radar Simulator
AFRL/RH	Air Force Research Laboratory Human Effectiveness Directorate
ANOVA	analysis of variance
AP	auto-pilot mode
ASAS	Airborne Separation Assistance System
ATCo	air traffic controller
CHAAT	Center for Human Factors in Advanced Aeronautics Technologies
CONOPS	concept of operations
DSR	display system replacement
FAA	Federal Aviation Administration
ft	foot/feet
GCS	ground control station
HITL	human-in-the-loop
hr	hour
IFR	instrument flight rules
LAX	Los Angeles International Airport
LOS	loss of separation
M	manual mode
MAC	Multi-Aircraft Control System
min	minute
MQ-1	General Atomics MQ-1 Predator
MQ-9	General Atomics MQ-9 Predator B UAS, "MQ-9 Reaper"
MR	measured response
MR1	pilot verbal latencies
MR2	pilot execution initiation latency
MR3	aircraft response latency
MR4	display visibility latency
MUSIM	Multiple UAS Simulation
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
NASA-TLX	NASA Task Load Index
nm	nautical mile
RQ-4	Northrop Grumman RQ-4 Global Hawk UAS
s	second
SART	Situation Awareness Rating Technique
SEM	standard error of the mean
TRACON	Terminal Radar Approach Control Facilities
UAS	unmanned aircraft systems
UHF/VHF	ultra-high frequency/very-high frequency radio
UTC	Coordinated Universal Time
VFR	visual flight rules
VSCS	Vigilant Spirit Control Station
WP	waypoint-to-waypoint mode
ZLA	Los Angeles Air Route Traffic Control Center
ZOA	Oakland Air Route Traffic Control Center

Measured Response for UAS Integraton into the National Airspace System

Kim-Phuong L. Vu[1], Dan Chiappe[1], Thomas Z. Strybel[1], Lisa Fern[2],
Conrad Rorie[3], Vernol Battiste[3], Robert J. Shively[4]

Executive Summary

The measured response (MR) is the response time of aircraft to air traffic controller (ATCo) commands and clearances. The overall MR can be broken up into several components, including the pilot verbal latencies (MR1), the time between the end of an ATCo clearance and the beginning of the pilot's read back, and the execution initiation latency (MR2), the time between the end of the ATCos' clearance and when the pilot begins to execute a maneuver. The MR is a crucial concern for the integration of unmanned aircraft systems (UAS) into the National Airspace System (NAS) due to potentially greater latencies stemming from remote pilot communication and command execution. As a result, it is important to quantify what latencies in verbal responding and command execution are acceptable for safe and efficient operations in the NAS. The present studies begin to address these issues in a series of four simulations supported by the UAS Integration into the NAS program.

- *Simulation A* was designed to develop procedures for extracting the MR components using the Multiple UAS Simulation (MUSIM) infrastructure as a ground control station (GCS) testbed. Controllers issued different clearances serially to UAS pilots flying in a simulation environment. The pilots were required to read back the commands and execute them as quickly as possible. ATCos also rated the acceptability of how long it took pilots to respond. We found: (1) MUSIM and Multi-Aircraft Control Systems (MACs) can be used to extract MR components and (2) UAS pilot latencies to initiate verbal responses to ATCo instructions (MR1) averaged 2.5 s when they did not have tasks to perform other than respond to ATCo instructions. Execution initiation latencies (MR2) varied widely (1.7 s to 7.63 s) across commands and may reflect the fact that immediate execution is generally not seen as critical to ATCos as immediate verbal responding (MR1).
- *Simulation B* examined how short (1.5 s) versus long (5 s) delays in verbal responding and command execution by one UAS in a NAS environment affected the acceptability ratings of ATCos and the performance of their duties. It also examined whether the predictability of the delays has an effect on ATCos. We found: (1) Delays in verbal responding are more important to ATCos than are delays in command execution; (2) verbal response latencies (MR1) of 2.1 s

---

[1] California State University, Long Beach.
[2] Ohio State University.
[3] San Jose State University Foundation.
[4] NASA Ames Research Center.

are generally acceptable to ATCos, though this likely applies to only sectors similar in characteristics to the one examined in the simulation; (3) the safety and efficiency of ATCo duties was neither affected by the delays in verbal responding and command execution, nor by the predictability of the delays; (4) ATCos are more tolerant of delays in responding by UAS pilots than by pilots of conventional aircraft; and (5) ATCo judgments of the acceptability of latencies in verbal responding and initiation of command execution are complex judgments that reflect multiple features of the air traffic management situation. Thus, specifying an upper limit on what is an acceptable MR will require taking multiple contextual factors into account.

- *Simulation C* examined whether short delays in verbal responding and command execution that were acceptable in Simulation B continued to be so when there were multiple UAS in the sector (1, 2, or 4), and when they were flying at different speeds (fast vs. slow). We evaluated the consequences of these manipulations for ATCo performance, communication in the sector, as well as the ATCo acceptability ratings of the latencies in verbal responding and command execution by UAS pilots and pilots of conventional aircraft. We found: (1) Additional verbal delays of 1.5 s were generally acceptable to controllers regardless of the number and speed of UAS; (2) as with Simulation B, latencies in verbal responding were more salient to ATCos than latencies in initiating command executions; (3) additional delays of 1.5 s in initiation of command executions were acceptable to controllers, regardless of number and speed of UAS; and (4) although 1 or 2 UAS appear to be manageable, 4 UAS flying unpredictable missions in a complex traffic sector with arrivals and overflights appears to exceed the limits of what is manageable to ATCos, as this led to greater losses of separation.
- *A Full Mission Simulation* of a UAS operating in the NAS was also conducted by the program. The goal of this simulation was to compare three different control modes (Waypoint, Autopilot, and Manual) for their effects on the MR1 and MR2 components. Results from this simulation showed: (1) The Autopilot condition had fastest MR1 latencies compared to Waypoint and Manual modes; (2) MR1 was generally fast, with 90% of pilots replying verbally to ATCo advisories and clearances in 3 s or less; and (3) Autopilot condition also had faster MR2 latencies compared to the other two conditions. In general, 90% of pilots initiated an edit in response to ATCo clearances in 11 s or less.

Taken together, these simulations provide an important step producing data to help identify acceptable latencies in verbal responding (MR1) and initiation of command executions (MR2) for UAS operations in the NAS.

1.0 Introduction

The U.S. Congress passed the Federal Aviation Administration (FAA) Modernization and Reform Act in 2012, which calls for a plan to integrate unmanned aircraft systems (UAS) into the National Airspace System (NAS). For UAS to be allowed to operate with unrestricted access in the NAS, they will be required to "act and respond as manned aircraft do" in terms of safety and efficiency (ICAO, 2011, p. 5). What standards and regulations should apply, be modified, or be developed is an open question (Askelson et al., 2013; Blickensderfer et al., 2012; Dillingham, 2013). FAA regulations, for example, require that pilots respond promptly to air traffic controller (ATCo) commands. However, they do not quantify what an acceptable latency is, stating only that they must not compromise the safe separation of aircraft. We maintain that research needs to be carried out to identify an acceptable latency in aircraft responding, known as the "measured response" (Askelson et al., 2013). While the measured response can be broken down into several components, the present set of studies examined how UAS pilot latencies in verbal communication and command execution affect ATCo performance, communication, and controller acceptability ratings of their responses when UAS are operating in the NAS environment. It also examined how contextual factors, such as the number of UAS in the sector and their airspeed affect the acceptability of delays in verbal responding and command execution.

2.0 Background on the Measured Response

2.1 Components of the Measured Response

ATCos manage traffic by issuing verbal commands to aircraft. Pilots then read back the clearances and execute the maneuvers. ATCos monitor the read backs and correct the pilots if they make an error (Hopkin, 1995). This confirmation and correction loop serves to minimize the likelihood of incidents and accidents resulting from communication failures (Cushing, 1995). However, this time-consuming procedure can also add significantly to the workload of ATCos. This is especially the case if they have multiple aircraft to communicate with in a short period of time, such as when they are merging and spacing several arrival streams. In such contexts, if the responses of pilots are not prompt, the workload of ATCos can be increased even further, hampering their efforts at effective time management (Yang, Rantanen, & Zhang, 2010; Wickens & Hollands, 2000).

The end-to-end response time of pilots to ATCo verbal commands is known as the "measured response" (MR; Askelson et al., 2013; Shively, Vu, & Buker, 2013; Vu et al., 2013; Ziccardi et al., 2013). It can be broken down into the following components as shown in Figure 1: (1) Pilot verbal latency (MR1); the lag between the end of an ATCo's verbal clearance and the beginning of the pilot's read back; (2) pilot execution initiation latency (MR2); the lag between the end of the ATCo's clearance and when a pilot starts to execute the maneuver; (3) aircraft response latency (MR3); the lag between the pilot entering a command in the control interface and the aircraft acting in response to that command; and (4) display visibility latency (MR4), the time for the maneuver to be available on the ATCo radar screen after the aircraft has started its maneuver.

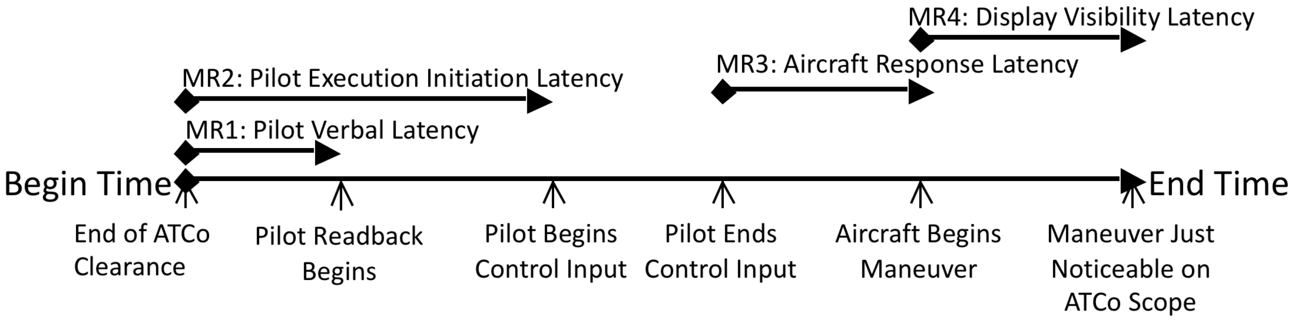


Figure 1. Illustration of the measured response (MR) components in the course of the ATCo-Pilot communication interaction.

In the present set of studies, we focused mainly on pilot verbal latency and execution latency components (MR1 and MR2). We did not examine the aircraft response time latencies (MR3) because these are not manipulated by our simulation infrastructure, that is, it does not simulate aircraft performance parameters. In our simulations, aircraft respond immediately once pilots enter commands. We measured the MR4 component, display visibility latency, but did not manipulate it as it was relatively constant throughout the simulation and was influenced by the update rate set by our simulation and recording software.

2.2 UAS Characteristics and the Measured Response

UAS differ from conventional aircraft in many important respects, ones that may well affect the time required to respond to ATCo commands (Askelson, et al., 2013). For example, UAS tend to be smaller, slower, and have very different control interfaces. Most crucially, UAS pilots are not co-located with their aircraft (Williams, 2007). As a result, the pilots lack many of the sensory cues (e.g., ambient vision, vestibular and acoustic information) available to conventional pilots to make quick assessments about the state of their aircraft. UAS pilots also have difficulties in scanning the environment around their aircraft because the cameras typically have limited spatial resolution and a small field of view (Merlin, 2013). Therefore, at least at present, UAS have poor "sense and avoid" capabilities. This is one of the main challenges to the integration of unmanned aircraft into the NAS (Dillingham, 2013; McCarley & Wickens, 2005; Vu & Chiappe, 2015; Williams, 2012). It can thus take UAS pilots longer to determine whether they can safely carry out a command issued by an ATCo. This can yield longer latencies in verbal responding, that is, a greater MR1 component than for conventional aircraft.

Verbal latencies can also be greater for UAS depending on the communication infrastructure employed. Satellite communication links, as opposed to the UHF/VHF radio communications used by pilots of conventional aircraft, also lengthen the time to verbally respond by increasing audio set up and signal propagation times (Dillingham, 2013; Wourms, Ogden, & Metzler, 2001). Indeed, long latencies associated with satellite technology can negatively affect performance of ATCos. For example, Nadler, Mengert, DiSario, Grossberg, and Spanier (1993) manipulated the audio set up latencies associated with several potential satellite communication systems in a study where ATCos had to manage traffic in simulated en route sectors. They measured the number of "step-ons," that is, communications that were simultaneously transmitted, leading at least to a partial blocking of the messages. They measured these because of their link with serious aviation accidents and incidents. Nadler et al. (1993) found a greater number of step-ons with the longer set up latencies, particularly when traffic density was high. Many of these step-ons led ATCos to reissue clearances, further increasing their workload. These results were replicated in a simulation study by Sollenberger, McAnulty, and Kerns (2003). They found that not only did lengthy audio set up latencies lead to more step-ons, these also led ATCos to issue fewer communications (likely decreasing the optional services provided), and those that they did issue were shorter (possibly because of an increase in speaking rate). In sum, if UAS are going to rely on satellite communication systems, the latencies involved could negatively affect the safety and efficiency with which ATCos manage traffic in the NAS.

Pilot execution latencies can also be greater for UAS than conventional aircraft depending on the type of control interface involved (Askelson et al., 2013). For example, stick, throttle and rudder control inputs, such as those in the Predator-B, can allow for faster response execution than systems that involve supervisory control (i.e., human-on-the-loop), such as the Global Hawk (Williams, 2007). In supervisory control systems, pilots issue high-level commands to aircraft, which then perform them autonomously (Sheridan, 1992). However, pilots enter commands by engaging in often time-consuming operations, such as clicking through computer screens and pull-down menus, and making keyboard entries to carryout maneuvers such as changing altitudes, airspeed, or navigating to a new waypoint (Merlin, 2013). Pilot execution latencies can increase depending on the UAS interface. Thus, it is important to consider system-wide implications when designing control interfaces, as the latter can increase the time required to execute clearances issued by ATCos (Vu & Chiappe, 2015).

Although the measured response is a crucial issue for UAS operations in the NAS, there has been little research conducted on this topic. A notable exception is a study by Askelson et al. (2013). They used a Predator (MQ-9) simulator flying in stick and rudder mode to execute commands issued by ATCos. The commands (heading, altitude, and speed changes) were issued during different phases of flight. Askelson et al. (2013) measured the response time, defined as the lag between the end of the ATCo's clearance and when it became apparent on the ATCo display that the UAS is complying with the command. They also measured the maneuver completion time—the time between the end of the ATCo clearance and when the desired result was achieved by the UAS. They found average response times of 11.85 s across all maneuvers, and completion times that averaged 71.20 s (ranging from 58.70 s to 88.25 s depending on the clearance). Although these findings do provide important information regarding overall response times of one type of UAS, they are limited with respect to the questions addressed in the present studies. Specifically, they did not examine the latency in verbal responding by UAS pilots (MR1) or when the UAS pilots began executing maneuvers (MR2). Moreover, they did not examine the ATCo's acceptability ratings of the latencies that they obtained. Even if they had measured acceptability, it is possible that although an overall response time is acceptable to ATCos, some of the latencies that make it up may not be. That is why it is important to also look at the various components of the measured response.

2.3 The Measured Response Research on Conventional Aircraft

The measured response has received some attention from researchers interested in conventional aircraft. One of the earliest studies was conducted by Cardosi (1993), who examined recordings of verbal interactions between pilots and controllers from three en route sectors. Her goal was to identify the time from the ATCo's initial transmission to the end of the pilot's verbal response. She did not measure the pilot execution latency (MR2). The total communication time was found to be approximately 10 s, across three types of clearances. Cardosi (1993) broke down this overall response time into several components and found a mean pilot verbal latency of 3.31 s (*SD* = 4.80 s) for maneuvers for traffic avoidance. She found a mean verbal latency of 2.68 s (*SD* = 4.60 s) for turns not for traffic. For traffic advisories, Cardosi (1993) found a mean verbal latency of 2.67 s (*SD* = 6.25 s). From these results, however, one cannot precisely determine what an acceptable pilot verbal latency is. This is because Cardosi (1993) did not measure the acceptability to ATCos and the fact that there was a substantial variability associated with each of these means. Thus, it is quite likely that the means are based on some response latencies that were too long to be acceptable to controllers.

More recently, Smith (2008) conducted a similar study to that of Cardosi (1993), using recordings from Terminal Radar Approach Control Facilities (TRACON) centers. Given that these are generally busier sectors it is not surprising that Smith (2008) found shorter average communication times between ATCos and pilots. These ranged from 5.18 s to 8.62 s, depending on the type of clearance. She also found that the verbal latencies between pilot and controller responses were shorter (an average of about 1 s) than those identified by Cardosi (1993) for en route sectors. However, the 1 s latency reported by Smith (2008) included both pilot responses (the ATCo initiated the communication) and ATCo responses (the pilot initiated the communication). Consequently, on its own, this result cannot be used to identify what is an acceptable pilot verbal latency. This is also the case because Smith (2008), like Cardosi (1993), did not have ATCos rate acceptability of the latencies in verbal responding by pilots.

Other studies involving conventional aircraft have examined the effect of different kinds of latencies on ATCo performance in simulation environments. Rantanen, McCarley and Xu (2004), for example, tested the effects of varied audio system latencies (i.e., set up and propagation times) and pilot latencies on ATCo performance. They examined these delays in part-task simulations that differed in terms of the number of communication exchanges that were required between pilots and ATCos to carry out the actions. The tasks either required one communication exchange (i.e., ATCo issues clearance and pilot verbally acknowledges) or multiple communication exchanges (i.e., the task requires ATCos and pilots take more than one conversational turn). They found that both audio system and pilot latencies negatively affected ATCo performance, but only when multiple exchanges were required. In these conditions, greater audio latencies led to decreased lateral separation between aircraft and increased communication time. The presence of pilot latencies was also associated with lower lateral separation between aircraft, greater communication time, and a higher number of step-ons. However, in the conditions where they varied pilot latencies, the latencies were randomly selected from a distribution whose mean was based on those observed by Cardosi (1993). It is therefore not possible to determine whether all of the pilot latencies were equally problematic, or whether some were acceptable to controllers.

In summary, one cannot make precise claims about acceptable pilot verbal latencies based on the current research conducted with conventional aircraft. Furthermore, with respect to the second MR component there is very little known about how long it takes pilots to begin executing commands, let alone what is acceptable in this respect. One exception is research carried out by Consiglio, Hoadley, Wing, Baxley and Allen (2008). They conducted a modeling study examining interactions between pilots and an Airborne Separation Assistance System (ASAS) flight deck tool. ASAS alerts pilots to the presence of conflicts, as well as consisting of a conflict resolution tool that calculates route modifications upon a pilot's request. Their simulation used pilot models and varied their latencies in responding to conflict alerts. Their pilot execution latencies ranged from 5 s to 240 s, and they also manipulated whether the traffic density was 5x, 8x or 12x current day en route levels. Using losses of separation (LOS) as their dependent variable, they found that the ASAS system was able to maintain high levels of safety, except when pilot execution latencies were greater than 90 s and traffic densities were at 8x and 12x current levels. Although the results from this simulation are useful for guiding future research, the values identified do not allow us to determine acceptable pilot execution latencies. This is because the simulated pilots were not responding to ATCo communications, but rather to a flight deck automation tool.

3.0 The Measured Response Simulations

3.1 Overview

To begin addressing important questions not answered by the current literature on the measured response, we carried out three simulations. The first of these, Simulation A, is a feasibility study designed to develop procedures for extracting the MR components using the Multiple UAS Simulation (MUSIM) infrastructure as a ground control station (GCS) testbed. Two controllers issued different clearances to UAS pilots flying in a simulation environment. The pilots were required to read back the commands and execute them as quickly as possible. In addition to recording the MR components, we had ATCos rate the acceptability of how long it took pilots to respond.

The second simulation, Simulation B, manipulated the pilot verbal latencies (MR1) and execution latencies (MR2). Specifically, it manipulated whether the latencies were long or short by adding either 1.5 s or 5 s to the initiation of the UAS pilot verbal responses and adding either 1.5 s or 5 s to the onset of their execution of ATCo commands. The study also manipulated whether or not the delays were predictable (i.e., whether they were constant in a scenario, or varied). In the context of UAS operations in the NAS, the predictability of the response latencies is important to investigate because UAS may have different communication latencies within the same sector, depending on whether pilots are communicating with ATCos via VHF/UHF radio or through a satellite communication link. Execution latencies may also differ, depending on whether the UAS pilots are controlling their aircraft via line of sight or satellite links, the latter yielding longer latencies (Dillinger, 2013; Merlin, 2009). We measured the consequences of these manipulations for the performance of ATCo duties, communications in the sector, and ATCo ratings of the acceptability of the latencies of both UAS pilots and conventional aircraft pilots. A central issue is whether the imposed delays in UAS pilot responding affect interactions between ATCos and pilots of conventional aircraft. Indeed, from a human systems integration perspective, it is crucial to examine the system-wide effects of UAS characteristics (Vu & Chiappe, 2015).

The third simulation, Simulation C, sought to further examine the contextual factors that may affect the acceptability of latencies in verbal responding and command execution by UAS pilots. It used the short additional delays in verbal responding and command execution by UAS pilots from Simulation B and held these constant for all UAS. However, unlike Simulation B, which only had 1 UAS in the sector at any given time, Simulation C manipulated the number of UAS (1, 2, or 4) as well as manipulating the airspeed of those UAS (fast vs. slow). It examined the consequences of these manipulations for ATCo performance, communication in the sector, as well as the ATCo acceptability ratings of the latencies in verbal responding and command execution by UAS pilots and pilots of conventional aircraft. Although the short latencies in verbal responding and initiation of command execution were generally found to be acceptable in Simulation B, the fact that ATCo acceptability ratings appeared to reflect broad contextual features of the scenarios they were managing led us to examine whether the acceptability of these short delays are fairly robust, despite changes in the number of UAS and the speed with which they fly.

Finally, a Full Mission simulation conducted at NASA Ames compared three different control modes for their effects on the MR1 and MR2 components. This last simulation demonstrates how the MR components can be captured and analyzed in higher fidelity simulations to examine issues relating to UAS integration into the NAS.

3.2 Summary of Simulation Goals

Simulation A

- Demonstrate our ability to capture the measured response components with the MUSIM GCS architecture.
- Identify pilot verbal latencies and latencies in onset of command execution of standard ATCo commands and clearances and measure acceptability of these latencies to ATCos.

Simulation B

- Identify latencies in UAS verbal responding and command execution that are acceptable to ATCos when multiple aircraft are in their sector.
- Determine whether predictability of latencies is an important factor in determining acceptability of response latencies.
- Examine how context influences ATCo acceptability ratings of MR1 and MR2 latencies.

Simulation C

- Examine robustness of ATCo acceptability of short delays in MR1 and MR2 by manipulating number of UAS in the sector, as well as UAS speed.
- Determine consequences of number of UAS and UAS speed for performance of ATCo duties as well as communication in the sector.

Full Mission Simulation

- Compare different control modes and their effect on both MR1 and MR2.
- Provide a demonstration for the benefit of capturing MR components in simulation studies examining UAS integration into the NAS.

3.3 Simulation A

Simulation A was designed to establish a methodology for capturing some of the measured response components of UAS in a human-in-the-loop (HITL) simulated environment. Measured response is a measurement of UAS response times to ATCo commands, and includes the following four components described in detail above: (MR1) Pilot verbal response latency; (MR2) Pilot initiation of command execution latency; (MR3) Aircraft response latency; and (MR4) Display visibility latency. A methodology to accurately capture these intervals is integral to future experimental research investigating the communication dynamic between ATCos and UAS, and to determine critical factors that differ from that of manned aircraft. The intention was that, once established, this method could be used with more elaborate experimental manipulations to further elucidate the communications relationship that ATCos and UAS pilots will have in the national airspace.

The present study captures the measured response by simulating the presence of UAS in an air traffic environment and recording the response time of the UAS to standard air traffic commands. Out of the four components of MR described earlier, we mainly focused on the measurement of the first two components: the interval between the ATCo's instruction and the UAS pilot's verbal response and the time for the UAS pilot to begin execution of ATCo command. In addition, Simulation A measured pilot workload and ATCo acceptability of the latencies in verbal responding and initiation of commands. This study is summarized in Ziccardi et al., (2013) and Shively, Vu, and Buker (2013).

3.3.1 Method

3.3.1.1 Participants

Two retired ATCos participated in the study, in addition to 14 Instrument Flight Rules (IFR) rated pilots. The ATCos had experience in both civilian towers and civilian TRACONs. Both indicated that they were experienced with the Multi-Aircraft Control System (MACS; Prevot, 2002) software that was used to simulate the radar, air traffic control environment. The 14 instrument flight rules (IFR) pilots did not have any UAS flying experience, but all had experience with the MUSIM ground control station used in the study. The pilots all had flight training simulator experience, and an average of 1,171 actual flight hrs.

3.3.1.2 Simulation Environment

The data were collected in the Flight Deck Display Research Laboratory at National Aeronautics and Space Administration (NASA) Ames Research Center. The MUSIM-UAS ground control station was used to simulate the actions of UAS pilots (for a detailed description of MUSIM, see Fern and Shively, 2009). The MUSIM display consisted of a north-up map with ownship and the currently filed route of the UAS (Figure 2). The aircraft's path was controlled by dragging and dropping waypoints along the planned route, while altitude was controlled by editing planned altitude levels at waypoints along the planned route. During the simulation the MUSIM screen did not display any other traffic. The ATCo MACS-DSR (Display System Replacement) display simulated the sector used in the simulation Los Angeles Air Route Traffic Control Center (ZLA) 20. Targets on the ATCos' DSR were depicted as a chevron icon on the display. The icon's shape and position allowed the ATCo to determine aircraft's heading, while the altitude of the aircraft was displayed in its data tag. Unlike the MUSIM display, air traffic in addition to the UAS was displayed. However, the ATCo was instructed to only pay attention to the traffic when instructed to issue a traffic call to the UAS. In other words, the ATCos were not managing the traffic in the sector. The pilot, ATCo, and experimenters that were facilitating the data collection spoke to each other using push-to-talk headsets over a voice system.

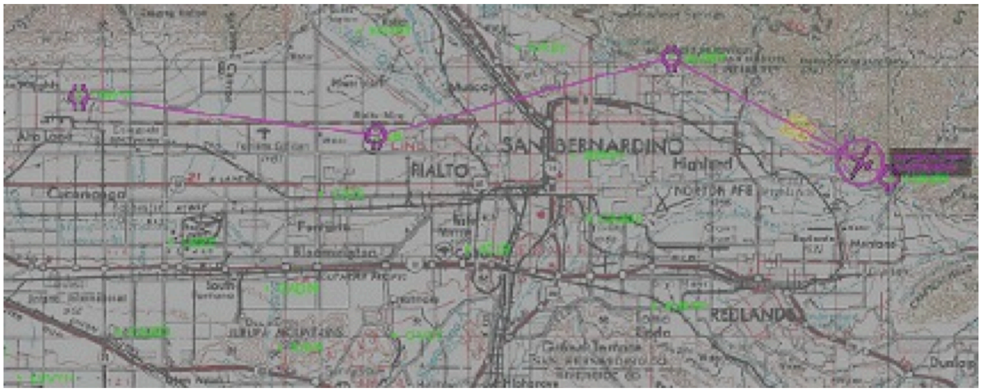


Figure 2. MUSIM-UAS map display with a waypoint path for the UAS specified.

3.3.1.3 Procedure

Pilots participated in this data collection effort following an unrelated study that used MUSIM, so pilots only received a 5-min training presentation to familiarize them with this simulation and reviewed important MUSIM procedures. Following the training presentation and prior to the experimental trials, each pilot completed a 5-min training scenario during which they received five different commands that included two frequency changes, one altitude change, one heading change, and one command that involved editing a waypoint. These commands were pulled from a list of 15 base commands, which were presented in a partially counterbalanced order. The commands consisted of crossing restrictions, traffic calls, radio frequency changes, route amendments and direct to waypoint clearances. The following are examples of each of the 15 command types used, presented here in an arbitrary order:

> "PD-1, cross GAREY at one-four-thousand."
>
> "PD-1, turn left heading two-four-zero, proceed direct ZIGGY, then resume own navigation."
>
> "PD-1, descend and maintain one-two-thousand."
>
> "PD-1, turn right, fly heading one-two-zero."
>
> "PD-1, climb and maintain one-five-thousand, traffic twelve o'clock, five-zero miles, westbound one-three-thousand."
>
> "PD-1, contact center 133.25."
>
> "PD-1, traffic nine o'clock, five-zero miles, eastbound one-four-thousand. Turn right immediately, fly heading one-five-zero."
>
> "PD-1, contact tower 129.00."
>
> "PD-1, traffic one o'clock, five-zero miles, westbound one-five-thousand. Turn left immediately, fly heading two-six-zero."
>
> "PD-1, contact approach 126.30."
>
> "PD-1, descend and maintain one-three-thousand. Traffic eleven o'clock, five-zero miles, eastbound one-four-thousand."
>
> "PD-1, turn left, fly heading one-eight-zero."
>
> "PD-1, climb and maintain one-six-thousand."
>
> "PD-1, turn right three-one-zero, direct KRAIL, then resume own navigation."
>
> "PD-1, cross CAJON at one-five-thousand."

As noted earlier, the commands were partially counterbalanced between participants, and each participant completed two identical blocks of trials of each of the 15 commands. The ATCo was provided with a list of commands for each trial. However, in their list, some commands were missing specific information (which had to be changed on a trial-to-trial basis depending on the current UAS location), such as the waypoint to direct the UAS. The missing information was provided during the trial by an experimenter seated adjacent to the ATCo. Each command was issued after the MUSIM observer confirmed the MUSIM pilot was ready. The ATCo issued each command individually and the pilot verbally acknowledged the command and began its execution. As the aircraft started the action on the ATCo scope, the ATCo confirmed the command was being executed on his MACS-DSR screen, and this signaled the end of the trial. The MUSIM pilot was instructed to focus only on completing ATCo commands as quickly and accurately as possible, and s/he was not performing any other UAS relevant tasks.

Timestamps were collected for verbal exchanges between ATCos and UAS pilot communication, UAS pilot actions on the MUSIM station, and observable indications of aircraft changes on video recordings of both the pilot's MUSIM display and ATCo's MACS radar displays. For each of the 15 commands, these data allowed a detailed and accurate timeline of ATCo and UAS verbalizations and actions. In addition to the timing data, pilot workload and ATCo acceptability ratings were also collected using seven-point rating scales with higher numbers indicating greater workload and acceptability, respectively.

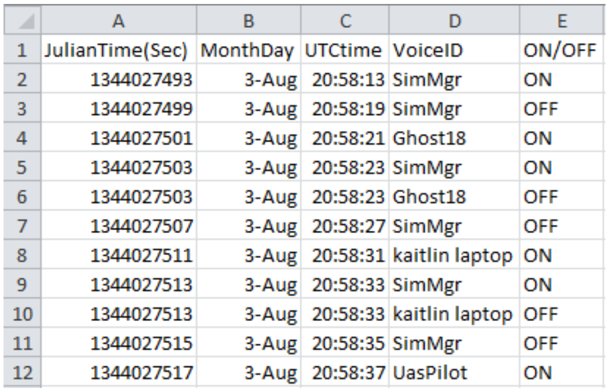
3.3.1.4 Verbal Exchange Timing Data

The verbal exchange data that were collected included the time when the ATCos and pilots depressed and released the push-to-talk button on their headset, as well as the time that they actually started and stopped speaking. The timing data for the push-to-talk button were collected from a voice logger program that provided exact time in Coordinated Universal Time (UTC) of button presses that was written to a text output file (Figure 3). These UTC times were synched with a voice recording of the trial and coded as different aspects of the exchange between pilot and ATCo (ATCo beginning verbal clearance, ATCo ending verbal clearance, UAS pilot beginning verbal response, etc.). Timing data for actual speech were determined by listening to a video file that displayed UTC synchronous with an audio recording of the exchanges.

Using this method, the following time points were collected for each command:

- Time ATCo depressed/released push-to-talk button in order to deliver clearance, and the time the ATCo verbally began speaking and when the ATCo ended issuing the clearance.
- Time pilot depressed/released push-to-talk button for acknowledgement of the ATCo command was collected in addition to the time the pilot verbally began speaking and the pilot's verbal response ended.
- Time that the ATCo depressed/released push-to-talk button, which signaled the detection of the executed command on the ATCo screen.

In the event that the command was a frequency change, timing data related to the pilot checking out of the original frequency and checking in to the new frequency, in addition to the response of the ATCo on the new frequency, were collected. All of these data were used to determine the interval for the first component of the measured response (MR1)—the time for the pilot in command of the UAS to verbally respond to ATCo instruction.

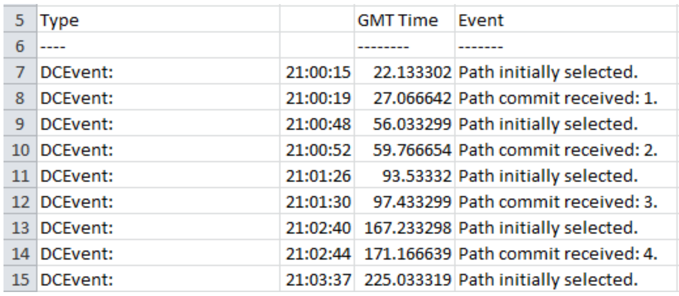


	A	B	C	D	E
1	JulianTime(Sec)	MonthDay	UTCtime	VoiceID	ON/OFF
2	1344027493	3-Aug	20:58:13	SimMgr	ON
3	1344027499	3-Aug	20:58:19	SimMgr	OFF
4	1344027501	3-Aug	20:58:21	Ghost18	ON
5	1344027503	3-Aug	20:58:23	SimMgr	ON
6	1344027503	3-Aug	20:58:23	Ghost18	OFF
7	1344027507	3-Aug	20:58:27	SimMgr	OFF
8	1344027511	3-Aug	20:58:31	kaitlin laptop	ON
9	1344027513	3-Aug	20:58:33	SimMgr	ON
10	1344027513	3-Aug	20:58:33	kaitlin laptop	OFF
11	1344027515	3-Aug	20:58:35	SimMgr	OFF
12	1344027517	3-Aug	20:58:37	UasPilot	ON

Figure 3. Output from logger program that recorded exact UTC time of button presses used to initiate and end voice transmission.

3.3.1.5 MUSIM Pilot Actions Timing Data

The actions that the pilots executed on the MUSIM station in response to the ATCo commands were coded, time stamped and recorded in a text output file. Using this output (Figure 4), the exact time the pilots began and ended executions were recorded. For frequency changes, a video recording that displayed UTC time in addition to the frequency change panel was used to visually determine timing. These data were used to calculate the interval for the second component of measured response (MR2): the time for the UAS pilot to begin an action after completion of the ATCo's instruction.



5	Type		GMT Time	Event
6	----		--------	-------
7	DCEvent:	21:00:15	22.133302	Path initially selected.
8	DCEvent:	21:00:19	27.066642	Path commit received: 1.
9	DCEvent:	21:00:48	56.033299	Path initially selected.
10	DCEvent:	21:00:52	59.766654	Path commit received: 2.
11	DCEvent:	21:01:26	93.53332	Path initially selected.
12	DCEvent:	21:01:30	97.433299	Path commit received: 3.
13	DCEvent:	21:02:40	167.233298	Path initially selected.
14	DCEvent:	21:02:44	171.166639	Path commit received: 4.
15	DCEvent:	21:03:37	225.033319	Path initially selected.

Figure 4. Output from MUSIM station detailing pilot actions and execution times.

3.3.2 Results

The mean time (i.e., averaged across the two blocks of trials) for each of the MR components was obtained for each participant. In several cases, there were negative times for the MR components. The negative times were due to pilots carrying out the MR components in parallel. Because we wanted to get a measure of the pure components, we "zeroed" all negative values when computing the means below. Descriptive analyses were performed for each of the MR components measured. We also used paired *t*-tests to determine if the time for each MR component, for each clearance type, differed from the first to second block of trials. This was done to determine whether practice effects were evident. Finally, we used correlation analyses to determine whether the time for the MR components, by clearance type, was correlated with the pilot workload rating and/or ATCo acceptability rating.

3.3.2.1 Descriptive Analyses

We provide the overall means and standard deviations for the MR components as a function of the clearance issued separately for each MR component in Table 1. Pilot workload and acceptability ratings were also provided.

Table 1. Means (and Standard Deviations) for each MR Component, Pilot Workload, and ATC Acceptability Ratings as a Function of Clearance Type

	Clearance type						
Measures	*Crossing Restriction*	*Direct To*	*Frequency*	*Route Amend– Altitude + Traffic*	*Route Amend– Heading*	*Route Amend– Altitude*	*Traffic Alert + Immediate Turn*
MR1 time (s)	2.75 (0.70)	2.64 (0.82)	2.24 (0.62)	2.41 (0.57)	2.77 (2.0)	2.13 (0.34)	2.76 (1.18)
MR2 time (s)	7.63 (5.66)	7.32 (7.14)	N/A	1.88 (2.17)	4.77 (3.11)	2.63 (1.91)	1.70 (0.93)
MR3 time (s)	Not captured because event occurs instantaneously in MUSIM						
MR4 time (s)	4.38 (2.75)	2.84 (1.36)	N/A	4.00 (1.99)	3.07 (1.15)	4.11 (1.43)	2.61 (1.70)
Pilot work-load rating	2.25 (1.04)	2.20 (0.88)	1.32 (0.44)	1.61 (0.71)	1.63 (0.71)	1.45 (0.56)	1.79 (0.75)
ATCo acceptability rating	6.10 (0.51)	6.15 (0.71)	6.87 (0.35)	6.55 (0.51)	6.39 (0.49)	6.38 (0.56)	6.51 (0.59)

Note: Workload ratings were on a scale of 1 = very low to 7 = very high. Acceptability ratings were on a scale of 1 = not acceptable to 7 = highly acceptable.

3.3.2.2 Practice Effects

Comparison of the response times for each MR component as a function of Clearance Type and Block (first or second) only showed two significant effects. The verbal response latency (MR1) component for the crossing restriction clearance (e.g., "PD-1, cross CAJON at one-five-thousand") was faster during the second block ($M = 4.25$ s) than during the first block ($M = 11.11$ s), $t(13) = 3.70, p = .003$. In addition, for the traffic alert + immediate turn clearance (e.g., "PD-1, traffic nine o'clock, five-zero miles, eastbound one-four-thousand. Turn right immediately, fly heading one-five-zero."), the display visibility latency (MR4) was also faster during the second block ($M = 2.29$ s) than during the first block ($M = 3.04$ s), $t(13), = 2.27, p = .041$. No other effects were significant, indicating that practice had very little effect on the data.

3.3.2.3 Correlational Analyses

We performed correlational analyses to determine whether the MR components for each Clearance Type were correlated with either pilot workload ratings or ATCo acceptability ratings. For workload ratings, none of the MR components for any Clearance Type were significantly correlated with pilot workload. For ATCo acceptability ratings, only a handful of MR components by Clearance Type were significantly correlated with ATCo acceptability. For Direct-to clearances, the execution initiation latency (MR2) was negatively correlated with the ATCo acceptability rating, $r(12) = -.61$, $p = .02$. For the Route Amendment – Altitude + Traffic clearance (e.g., "PD-1 climb and maintain one-five-thousand, traffic twelve o'clock, five-zero miles, westbound one-three-thousand."), the pilot verbal latency (MR1) was positively correlated with the ATCo acceptability rating, $r(12) = .58$, $p = .03$, but the execution lag (MR2) was negatively correlated with the ATCo acceptability rating, $r(12) = -.83$, $p < .001$. Finally, for the Route amendment – Altitude clearance (e.g., "PD-1, climb and maintain one-six-thousand."), the execution initiation latency (MR2) was negatively correlated with the ATCo acceptability rating, $r(12) = -.661$, $p = .01$.

3.3.3 Discussion

This simulation showed that the MR components could be extracted successfully from the MUSIM GCS and MACS ATCo configuration, making MUSIM a feasible testbed for studying UAS operations in the NAS. The average pilot verbal latency (MR1) was approximately 2.5 s. This is faster than the mean pilot verbal latencies observed by Cardosi (1993), which ranged from 2.65 s to 3.31 s in actual en route sectors. Although the pilot verbal latency did differ somewhat across clearance types, the differences were small. These findings indicate that pilots promptly acknowledged ATCo clearances regardless of the exact clearance being issued to them. However, because the UAS pilots were not doing any other tasks, it may be the case that the pilots were faster than they would have been in actual flying conditions, when they have to prioritize the different tasks they are being asked to perform.

The execution initiation latency (MR2) varied widely between clearance types as well as within a clearance type, as indicated by the fact that the standard deviation is oftentimes as large as the mean values. The wide variation in time between clearances can be attributed to the fact that pilots are able to start executing certain ATCo commands before the ATCo clearance is completely issued. We did observe negative execution lags in the data for the MR2 component, which supports the notion that pilots do not wait for the entire clearance to be read before they start executing their commands. As such, the MR2 component can begin before the MR1 component.

The display latency (MR4), though not a primary focus in our studies, showed some variance between clearance types, but the differences were much smaller than that for MR2. MR4 is influenced by the update rate on the ATCo's scope, which was set to 1 s in the current simulation to reflect Automatic Dependent Surveillance-Broadcast (ADS-B) update rates. However, because the component was extracted using post-simulation video of the ATCo scope, the actual update rate could have been as much as 2 s (the 1 s update rate for the scope and an additional 1 s screen recording rate of the software). This difference in time may not be as important to the ATCos because once the pilot acknowledges the clearance (MR1), the ATCos know that the pilot received the clearance instruction and simply need to check the aircraft state to determine if the pilot is executing the clearance correctly. In certain environments, such as en route, ATCos can move on to other tasks and check back about the status of the aircraft state when convenient, rather than continuously monitoring the aircraft state. This ability to check back later makes the promptness of the display latency less important. This is especially the case when the update rate of the radar sweep is long. In other environments and situations, such as closer to the airport and during emergencies, ATCos will need to check the aircraft status more often and thus accurate latencies and measurements of these times is critical for future studies.

Although the participants were given each clearance twice, once in block 1 and again in block 2, there was little evidence for practice effects. The lack of a practice effect is likely due to the pilots already being familiar with the MUSIM interface through their participation in prior simulations using MUSIM. In future studies, researchers need to make sure that the pilots are highly trained before the experimental runs to rule out practice effects.

The MR components did not correlate with pilot workload. However, the lack of significant correlations was probably a result of the pilot workload ratings being relatively low in most situations. More interesting was the fact that command execution initiation latency (MR2) was shown to be negatively correlated with ATCo acceptability ratings for a few of the clearances. The negative direction of the correlation indicates that shorter execution times are associated with higher ATCo acceptability ratings. There was only one positive correlation for pilot verbal latency (MR1) in which the longer latency was associated with higher ATCo acceptability. There was no obvious reason for the positive correlation. However, because the specific clearance included a traffic alert, it could be the case that the controller was expecting the pilot to locate the traffic before responding. This explanation, though, is only speculative.

The goal of the study was to develop the methodology to capture the measured response components of UAS pilots using MUSIM as a GCS testbed. However, the simulation environment used in the study was simple. There was no other traffic in the sector displayed on the MUSIM station displays other than the single UAS, and the UAS pilot was not doing any other tasks associated with typical UAS operations during this simulation. The pilot response times may therefore be longer under conditions of higher workload, which may in turn affect ATCo acceptability ratings. In addition, the data presented in this study reflect a small sample of pilots, who are professional pilots but not actual UAS pilots. As such, the numerical values may not reflect those obtained using actual UAS pilots.

3.3.3.1 Lessons Learned from Simulation A

- We can use MUSIM and MACS to extract MR components for UAS operating in a simulated NAS environment.
- Acceptability rating of UAS pilot latencies in verbally responding to ATCo instructions (MR1) and in initiating execution of commands (MR2) can be readily incorporated into simulation procedures.
- UAS pilot latencies to initiate verbal responses to ATCo instructions (MR1) averaged 2.5 s when they do not have tasks to perform other than responding to ATCo instructions.
- Execution initiation latencies (MR2) varied widely (1.7 s to 7.63 s) across commands and may reflect the fact that this MR component is generally not as important to ATCos as MR1, which was generally shorter and featured much less variability.

3.4 Simulation B

In Simulation B, we manipulated the pilot verbal and execution latencies and examined the consequences for the performance of ATCo duties, and on ATCo ratings of the acceptability of the latencies, the ratings being collected after each scenario. Specifically, we manipulated whether UAS pilot verbal and execution responses were “short” or “long.” This was done by either adding a delay of 1.5 s or 5 s to their verbal responses and by adding 1.5 s or 5 s prior to the beginning of their command executions. These values were chosen based on the acceptability ratings collected in

Simulation A, which found that execution initiation latencies greater than 5 s were generally rated unacceptable by the controllers. We also examined the effects of the latencies on the number of step-ons in communication exchanges between pilots and controllers. These have been linked with many air traffic incidents and accidents, including the 1977 Tenerife crash involving two 747s (Nadler et al., 1993).

Simulation B also investigated whether it is the latencies per se or their predictability that is relevant to ATCo performance and acceptability ratings. Much research has shown that adding signal delays can negatively affect performance (e.g., Clark, 2003; Ferrell, 1965; Held, Efstathiou, & Greene, 1972; Sheridan, 1992; 1993). For example, in tele-robotic operations, lengthy signal latencies can result in system instability and control difficulties (Sheridan, 1993). However, researchers have generally held delays constant within trials. It is possible that negative effects of delays are more pronounced when they are unpredictable to operators. When a lag is constant, they may be able to develop a strategy for managing it; an ATCo, for example, may learn to wait for a specific amount of time before repeating a command, but this strategy could not be employed if the delay is unpredictable. Delay unpredictability has been found to be important to coordination efforts, as the uncertainty of aircraft arrival times is a major source of operational costs for the airline industry (Agbolosu, Millner, Baden, Coville, & Mondolini, 2012). In the context of UAS operations in the NAS, the predictability of the response latencies is important to investigate because UAS may have different communication latencies within the same sector, depending on whether pilots are communicating with ATCos via VHF/UHF radio or through a satellite communication link. Execution latencies may also differ, depending on whether the UAS pilots are controlling their aircraft via line of sight or satellite links, the latter yielding longer latencies (Dillingham, 2013; Merlin, 2009).

Finally, Simulation B examined the factors that influence ATCo acceptability ratings of pilot verbal and execution initiation latencies. The ATCos were asked to rate these for both UAS pilots and conventional aircraft pilots after each scenario, even though we only manipulated the latencies of the UAS pilots. A crucial question we asked is on what basis are these judgments made? Do these judgments reflect the latencies on their own, or do they reflect how these latencies affect broader sector performance characteristics such as the number of step-ons, the number of LOS, or measures of efficiency, such as the average distance of aircraft through the sector? The results of this simulation are discussed in Vu et al. (2013) and Vu et al. (2015).

3.4.1 Method

3.4.1.1 Participants

Eight current or previously certified radar ATCos participated in the simulation. Six of the participants were retired and two were off-duty volunteers. The controllers reported an average of 28 years of air traffic management experience in military or civilian facilities. All controllers had prior experience with ZLA airspace that was used in this simulation. In addition, all of the controllers had participated in prior, unrelated studies using our simulation facility. The present study was scheduled over two consecutive days. All participants completed the study, but the data from one participant was excluded for non-compliance with some of the experimental procedures.

3.4.1.2 Design

The present study manipulated the UAS-pseudopilot Verbal Delay (short or long), UAS-pseudopilot Execution Delay (short or long), and the Delay Predictability (constant or variable). For both Verbal Delay and Execution Delay variables, a delay of either 1.5 s (short delay) or 5 s (long delay) was inserted into the UAS-pseudopilot communications or prior to the UAS pilot initiating the execution of the clearance. For the Constant conditions, combining the two communication and execution initiation delays resulted in four conditions: voice delay-short/execution delay-short, voice delay-short/execution delay-long, voice delay-long/execution delay-short, and voice delay-long/execution delay-long, tested in separate scenarios. In terms of the Delay Predictability variable, we compared findings from the constant delay scenarios that employed one of the four delay combinations throughout the scenario with the variable delay trials that employed two repetitions of each delay combination within each scenario (Table 2). We measured two of the MR components for both UAS and conventional aircraft: Verbal Latency (MR1) and Execution Initiation Latency (MR2).

Table 2. Number of Communications at Each Combination of Voice Delay and Execution Initiation Delay per Condition

Condition/Scenario	*Number of Communications (occurs in pairs: controller initiated; pilot initiated)*
Constant: VS-ES	VS-ES (n = 8)
Constant: VS-ES	VS-EL (n = 8)
Constant: VS-ES	VL-ES (n = 8)
Constant: VS-ES	VL-EL (n = 8)
Variable delays (4 replications, randomly presented)	VS-ES (n = 2) VS-EL (n = 2) VL-ES (n = 2) VL-EL (n = 2)

VS = Voice Delay-Short; VL = Voice Delay-Long; ES = Execution Delay-Short; EL = Execution Delay-Long.

3.4.1.3 Simulation Environment

To ensure that communications between operators were only made via radio, the simulation was conducted in three rooms of the Center for Human Factors in Advanced Aeronautics Technologies (CHAAT) at California State University, Long Beach. Each room contained workstations for one of the three operator roles: (1) participant ATCos; (2) UAS pseudopilots; and (3) conventional-aircraft pseudopilots. For ATCos and conventional-aircraft pseudopilots, the simulation used MACS, a medium-fidelity simulation architecture. MACS simulated a DSR of sector ZLA 20 for ATCos, and a pseudopilot flight-deck display for conventional-aircraft pseudopilots. The MUSIM GCS was used to fly the UAS. The UAS pseudopilots controlled their aircraft by altering the altitude and waypoints (point-to-point navigation) in MUSIM. The UAS callsign was “PD-1” in all scenarios. Confederate UAS and conventional-aircraft pseudopilots were students in CHAAT with extensive training on each role.

Two parallel worlds were run simultaneously. The controllers' stations included two computers, one DSR and one 'Mission Control' display used to instruct the controller on the desired heading of PD-1. The ATCo, conventional-aircraft pseudopilots, and UAS pseudopilots communicated with each other through a VoiceIP system using push-to-talk headsets. The voice client for the UAS station was modified so that a delay could be inserted in the UAS-pilot transmissions. For all operator roles, the voice system blocked transmissions if any operator speaking stepped-on the transmission of another operator. In this case, both transmissions would be unintelligible for everyone listening on the frequency until only one operator was speaking. The voice system also logged the number and duration of step-ons. To implement the delays in command execution, the UAS pseudopilot activated a countdown timer of 1.5 s or 5 s after acknowledging the clearance. The UAS pseudopilot was instructed to not begin executing the clearance until the timer counted down. During the variable delay conditions, the voice and execution delays were counterbalanced within the scenario and each delay combination was presented twice.

3.4.1.4 Procedure

The simulation was run over two consecutive days. After giving informed consent, participants completed demographic questionnaires and were briefed on the operational environment for the simulation and specific simulation procedures. Following the briefing, the ATCos engaged in a training session that consisted of four, 30-min practice scenarios. The first three scenarios did not include the UAS, and were designed to familiarize the ATCos with the airspace and traffic flows used in the simulation. The UAS was introduced in the final practice trial, but it did not produce any delays in communication or execution. After the training session, participants were allowed to ask questions about any of the simulation procedures.

The experimental scenarios were run in the afternoon of the first day and morning of the second day. Each experimental scenario was 40-min in duration, with about 50 aircraft entering the sector during the entire scenario and eight aircraft in the sector at any given time. During the experimental trials, the ATCos were instructed to give priority to arrivals into Los Angeles International Airport (LAX), and to ensure that arrival aircraft left the sector at an altitude of 12,000 ft or below and a speed of 250 kts. The controllers were told that a Letter of Agreement with the Center was in effect, requesting accommodation of requests from either Mission Control or the UAS ground station. The UAS flew in the sector for the entire scenario in a triangular pattern at an average speed of 120 kts, and altitudes between 10,000 and 16,000 ft. These flight plans ensured that the UAS crossed the arrival streams several times during a scenario.

ATCos received audio alerts and text messages on the Mission Control station indicating a new waypoint for the UAS (e.g., "PD-1 proceed direct EDITS"). The ATCo then issued the clearance when appropriate. UAS pseudopilots requested altitude changes (e.g., "LA Center, PD-1, request descent 1-4 thousand"). Eight planned requests for PD-1 were completed in each scenario, four initiated by the controller and four by the UAS pseudopilot. ATCo- and UAS-initiated communications occurred every 4-5 min in alternating orders, beginning 2-3 min into the scenario. Note that the UAS station produced only transmission delays, and no receiving delays. The ATCo and conventional-aircraft-pseudopilot stations produced neither receiving nor transmission delays.

At the end of each trial, controllers rated the acceptability of the verbal and execution response latencies (MR1 and MR2) for both UAS and conventional aircraft. In addition, situation awareness (SART; Taylor, 1990), workload (NASA-TLX; Hart & Staveland, 1987), and other subjective assessments were obtained. After finishing all eight experimental trials, the ATCos completed a post-simulation questionnaire and participated in a debriefing session. For all UAS-ATCo communications, MR1 and MR2 were collected using the procedures outlined in Simulation A. For conventional-aircraft-ATCo communications, MR1 and MR2 were also measured, using a sample of communications for each scenario and participant ATCo. The communications were equal in number to those of UAS pilots and ATCos, and were taken from similar points in the scenarios to reflect equivalent ATCo workload. In addition, we obtained several measures of sector performance: number of LOS, mean distance travelled, and the number of step-ons.

3.4.2 Results

In the following, we begin by describing the effect of the inserted delays of UAS verbal communications and initiation of command executions on the MR components of UAS and conventional aircraft. Then we describe ATCos' subjective ratings of the acceptability of the resulting latencies. As the acceptability ratings were made post scenario, the individual combinations of pilot verbal and execution initiation delays could not be separated in the variable conditions because the delays were intermixed within each scenario. We therefore present two separate sets of analyses for the acceptability ratings. One set examines the acceptability of the pilot verbal and execution latencies for the constant conditions taking into account Pilot Role (UAS vs. conventional aircraft pilot). The other set examines the effect of the predictability of the verbal and execution latencies by comparing the constant and the variable delay scenarios, collapsing across the different pilot verbal and execution delay combinations. Finally, we examine the basis of ATCos' acceptability ratings by correlating these ratings with MR1, MR2, and sector performance outcomes.

3.4.2.1 The MR Components

The means for the MR1 component (Verbal Latency) and MR2 component (Execution Latency) are presented in Table 3 for each Pilot Role (UAS vs. conventional aircraft), along with the mean ATCo acceptability ratings for these response times (see Vu et al., 2013, for a detailed analysis of the MR1 and MR2 components as function of the delay type and predictability). For the conventional aircraft, response times were calculated from a sample of pilot-ATCo communications, as outlined above. Not surprisingly, the UAS pilot verbal and execution response times are longer than those of conventional aircraft pilots, $Fs(1,12) > 23.4$, $ps < .001$, because verbal and execution initiation delays were added only to the former.

Table 3. Means (and Standard Deviations) for Pilot Verbal and Execution Latencies and for ATCo Acceptability Ratings for Constant and Varied Conditions as a Function of Pilot Role

Condition	*Pilot Role*	*Pilot Verbal Delay (s; MR1)*	*ATCo Acceptability Rating (MR1)*	*Pilot Execution Delay (s; MR2)*	*ATCo Acceptability Rating (MR2)*
Constant: VS-ES	UAS	2.07 (0.50)	5.14 (0.69)	6.27 (2.52)	5.14 (0.69)
	Conventional*	0.81 (0.23)	5.00 (1.41)	3.24 (2.21)	5.00 (1.83)
Constant: VS-EL	UAS	2.12 (0.27)	4.86 (1.57)	10.57 (1.32)	5.00 (1.00)
	Conventional*	0.80 (0.35)	5.43 (1.13)	4.79 (1.48)	5.43 (0.98)
Constant: VL-ES	UAS	5.52 (0.32)	4.00 (1.41)	7.17 (1.43)	4.43 (1.90)
	Conventional*	0.86 (0.34)	4.43 (1.51)	5.56 (3.48)	5.29 (0.95)
Constant: VL-EL	UAS	5.43 (0.25)	4.43 (1.13)	9.98 (2.35)	4.43 (1.13)
	Conventional*	1.15 (0.51)	4.43 (1.51)	4.97 (2.46)	5.29 (1.38)
Variable (averaged across the 4 scenarios)	UAS	3.72 (1.82)	5.07 (1.25)	8.98 (1.90)	5.18 (1.12)
	Conventional *	0.94 (0.54)	5.43 (1.07)	5.58 (2.49)	5.71 (1.06)

* For conventional aircraft pseudopilots, only a small, matched sample of pilot-ATCo interactions were examined. Acceptability ratings were on a scale of 1 = not at all acceptable to 7 = very acceptable.

3.4.2.2 Verbal Response Latencies (MR1)

ATCo acceptability ratings for the pilot verbal latencies were obtained after each scenario for each type of pilot. The distribution of ATCo acceptability ratings for the verbal response latencies of UAS and conventional aircraft pilots in each of the UAS Verbal Delay conditions is shown in Figure 5. These are separated by long and short verbal delays, collapsing across execution response times in the constant conditions and averaged across the four scenarios for the variable delay condition. Figure 5A shows ATCos' ratings of the UAS verbal latencies and Figure 5B shows ATCos' ratings of the conventional aircraft verbal latencies.

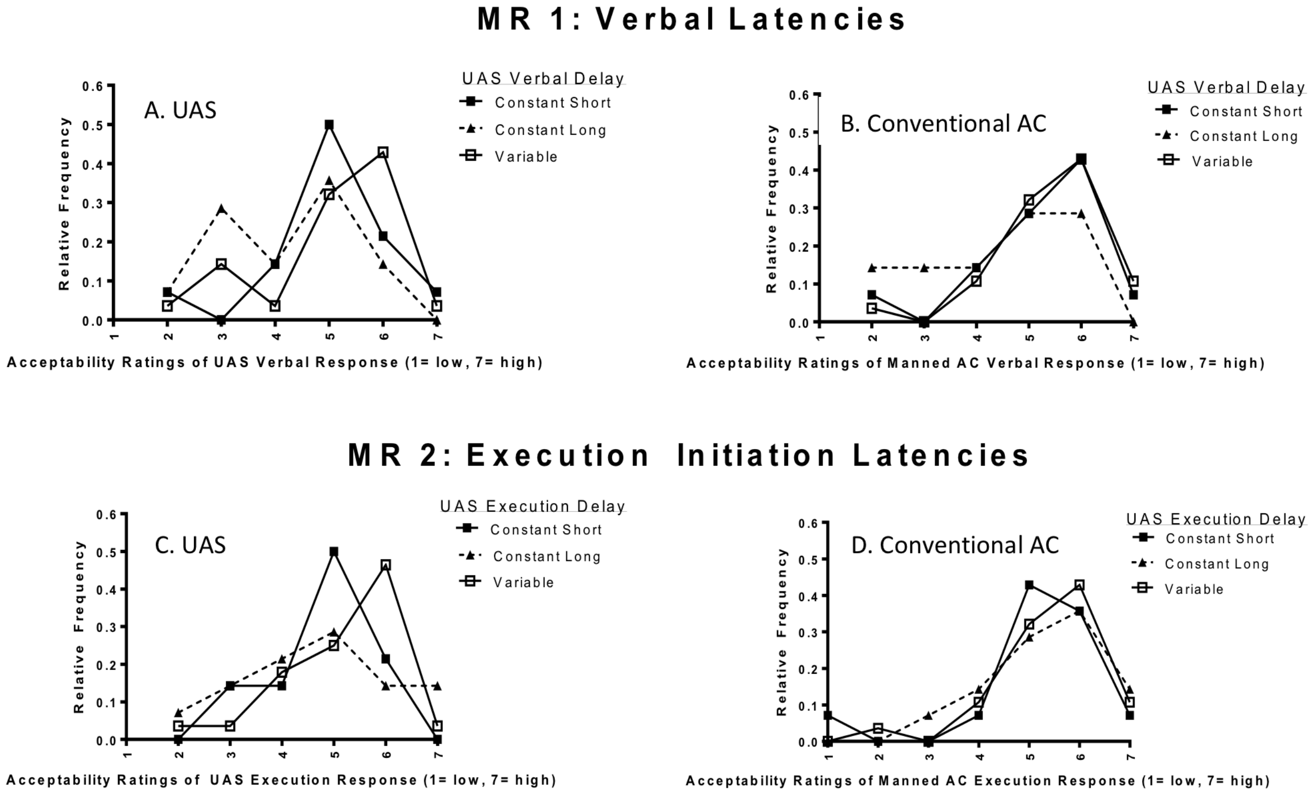


Figure 5. Relative frequency distribution of ATCo acceptability ratings (1 = low; 7 = high) as a function of UAS Verbal Delay condition for MR1 of UAS (A) and conventional aircraft (B), and MR2 of UAS (C) and conventional aircraft (D).

3.4.2.3 Acceptability Ratings

Overall, the majority of the ratings fell within an acceptable range of 5 or higher, showing that our controllers for the most part found these latencies manageable. For the constant-short verbal delay condition, which led to an average verbal latency of 2.10 s (*SEM* =.09 s) for UAS pilots and .80 s (*SEM* =.09 s) for conventional pilots, 78% of the ratings were in the acceptable range for UAS pilots (Figure 5A) and 79% for conventional pilots (Figure 5B). The fewest number of acceptable ratings occurred for constant-long verbal latencies, where the average latency was 5.48 s (*SEM* = .09 s) for UAS pilots and 1.01 s (*SEM* = .09 s) for conventional pilots. In this condition, only 49% of the ratings were acceptable for the UAS pilots and 57% for the conventional pilots. The drop in proportion of acceptable responses for conventional aircraft pilot responses in the constant-long condition is noteworthy because the delays were only added to the UAS pilot responses. Furthermore, the average latency for the conventional pilots is lower (1.01 s) than that for UAS pilots in the constant-short condition (2.10 s), where the acceptability of the UAS pilot delays were much higher than those of conventional pilots in the constant-long condition. These findings suggest that acceptability ratings by ATCos may reflect broader sector characteristics, a topic that will be further examined below.

To analyze the effects of our manipulations of pilot verbal delays on the acceptability ratings of controllers, the ATCo acceptability ratings of verbal latencies were analyzed with a three-factor repeated-measures analysis of variance (ANOVA) with the following factors: UAS-Verbal Delay (long vs. short), UAS-Execution Delay (long vs. short) and Pilot Role (UAS vs. conventional aircraft). We found a significant main effect of UAS-Verbal Delay, $F(1,12) = 7.49$; $p = .018$. On average, the mean acceptability rating for the short verbal delay conditions ($M = 5.1$ s; $SEM = 0.24$ s) was higher than for the long verbal delay conditions ($M = 4.3$ s; $SEM = 0.32$ s). There was no significant main effect of Pilot Role, even though delays were inserted only in the UAS verbal communications and execution initiations. This finding indicates that longer latencies in UAS verbal responses affected ATCos' acceptability ratings of all communication latencies in a scenario, not just those of UAS pilots.

Next we examine the effects of the predictability of the verbal delays on the acceptability ratings of ATCos. As shown in Figures 5A and 5B, the acceptability ratings for the variable delay condition were similar to those of the constant-short verbal delay condition. For variable delay scenarios, 78% of the ratings of UAS and conventional aircraft verbal response delays (which produced an average MR1 of 3.72 s and 0.92 s, respectively) were in the acceptable range of 5 or greater. If predictability of verbal responses had negatively affected our ATCos, we would have found the variable delay condition to produce the lowest proportion of acceptable responses. Moreover, the fact that the percent of acceptable responses in the variable delay condition was comparable to the constant-short verbal delay condition, and higher than for the constant-long verbal delay condition suggests that the acceptability of verbal response latencies depended not only on the length of the latencies but also on the number of verbal responses having longer latencies: In the variable delay scenarios, half of the verbal latencies were long, yet the distributions of acceptability ratings were consistent with the constant-short delay condition.

To evaluate the effect of predictability on ATCo acceptability ratings, a two-factor repeated-measures ANOVA was run on verbal latency acceptability ratings, with the factors of Verbal Delay condition (constant-short, constant-long and variable delay) and Pilot Role (UAS vs. conventional aircraft). A main effect of Verbal Delay condition was found, $F(2,24) = 6.21$; $p < .01$. Acceptability ratings were lower for the constant-long ($M = 4.3$; $SEM = .32$) than constant-short ($M = 5.1$; $SEM = .24$), $p = .05$, and variable delay ($M = 5.3$; $SEM = .16$), $p = .04$, conditions. The mean acceptability rating for the variable delay condition was not significantly different from the mean acceptability rating for the constant-short delay condition.

Taken together, our results indicate that short UAS verbal latencies averaging approximately 2.10 s were mostly acceptable to ATCo, but long UAS verbal latencies averaging approximately 5.48 s were not. The longer verbal latencies were only acceptable on approximately half of the scenarios. Thus, what is acceptable to ATCos likely depends on additional scenario factors. This is also illustrated by the fact that ATCos rated most scenarios with variable latencies to be as acceptable to scenarios with constant-short latencies, yet half of the UAS communications in the variable delay scenarios were delayed by 5 s. Apparently, the acceptability of a verbal response latency depended not only on the length of the response, but also on the number of longer responses that occurred. It is also important to point out that UAS communication latencies affected ATCos' acceptability ratings of conventional aircraft, because the distributions in Figures 5A and 5B are quite similar, despite the fact that no delays were inserted into the communications of the conventional aircraft.

3.4.2.4 Execution Latencies (MR2)

The distribution of ATCo acceptability ratings of the UAS-execution initiation latencies is shown for UAS and conventional aircraft in Figures 5C and 5D, respectively. These are separated by long and short delays in the constant conditions and averaged across the four scenarios for the variable delay condition. As illustrated in Figure 5C, 71% of the UAS ratings were in the acceptable range for constant-short execution delays, which had a mean MR2 of 6.72 s (*SEM* = .63 s), and 56% of the ratings were in the acceptable range for long-execution delays, which had a mean MR2 of 10.28 s (*SEM* = .53 s). For conventional aircraft (Figure 5D), acceptable responses were given for 92% of the scenarios in the constant-short execution delay condition, which had a mean MR2 of 4.4 s (*SEM* = .63 s). In the constant-long execution delay condition, which had a mean MR2 of 4.8 s (*SEM* = .54 s), acceptable responses were given for 77% of the scenarios.

Similar to the analysis performed for the verbal latencies, a three-factor repeated-measures ANOVA on the acceptability ratings of execution-initiation latencies (MR2) for constant scenarios was performed. This analysis showed no significant main effects of Verbal Delay, Execution Delay, or Pilot Role (*p*'s>.14). These factors also produced no significant interactions.

To examine the effect of predictability of the execution initiation latencies, a two-factor repeated-measures ANOVA was run on acceptability ratings of the Execution Delays, with three conditions (constant-short, constant-long and variable) and Pilot Role (UAS vs. conventional aircraft). The main effect of Pilot Role was not significant, indicating that ATCos tended to rate the acceptability of conventional aircraft execution initiation similarly to UAS execution initiation latencies, even though delays were added to the latter and not to the former. The distribution of the acceptability ratings of execution latencies (MR 2) followed a similar pattern to those of the verbal latencies (MR1). For constant-long delays, roughly half of the ratings were unacceptable; for constant-short and variable delays, over 70% of the ratings were acceptable. However, our ANOVA found no significant effect of Execution Delay on mean ATCo acceptability ratings, as well as no significant differences in acceptability between constant and variable scenarios.

Failing to find a difference between the long and short execution delay conditions could have happened for several reasons. First, it is possible that latencies in the execution of maneuvers are less relevant to ATCos than those associated with verbal responding. This would make sense given that controllers typically assume that the commands are executed promptly once clearances are acknowledged (Cushing, 1995). They typically wait a little while to go back and make sure that they have been carried out. Second, it is possible that although the execution latencies are relevant to ATCos, they were unable to discriminate between long and short latencies because they could not directly know when pilots began executing a clearance. ATCos would have to infer this from changes in the aircraft state on their radar scope (i.e., MR4). In our simulation, the mean MR4 was approximately 5 s and reflected the 1 s sweep rate of the simulated radar screen, update rates of the simulation software and the screen-recording software used to determine MR4. ATCos' ratings of the acceptability of execution latency most likely included this time and this could have made judgments of acceptability difficult.

Third, it is possible, indeed likely, that acceptability ratings of verbal responses and execution initiation were based on more than just the response latencies on their own. ATCos' ratings may have included the impact of the response latency on other scenario and performance factors. Suggestive of this idea was the fact that a strong, positive Pearson correlation was obtained between acceptability of verbal latency and acceptability of execution initiation latency, across all delay conditions and Pilot Roles, ($r = +.64$; $p < .001$). High acceptability ratings of verbal latencies were associated with high ratings of execution latencies, despite the fact that verbal and execution delays were crossed (i.e., long verbal delays occurred equally often with short and long execution delays). Consequently, we examined performance and other scenario factors that may have affected the acceptability ratings and determined the effect of UAS latencies on these outcomes in the next section.

3.4.2.5 Sector Performance

For each scenario and ATCo, we measured the number of LOS (a measure of sector safety), the average time of aircraft through the sector (a measure of efficiency), and the number of communication step-ons (which could affect both safety and efficiency). Table 4 shows the means and standard errors (in parentheses) for each measure of sector performance for constant and variable delay scenarios.

To examine whether our manipulations affected sector performance of our ATCos in the constant scenarios, we first conducted a series of two-factor ANOVAs with the following factors: UAS Verbal Delay (short vs. long) and UAS Execution Initiation Delay (short vs. long). These analyses were conducted on the number of LOS, the mean distance travelled by aircraft, and the number of step-ons. There were no significant main effects or interactions for any of these variables, showing that our controllers were able to effectively manage traffic regardless of the delays in verbal responding and command execution that we imposed on the UAS (all *Fs*<1).

Table 4. Mean Sector Performance Measures (and SEM) for
Constant and Variable Delay Scenarios

Performance Metric	*Constant Delay*	*Variable Delay*
Number of LOS	1.5 (.70)	2.1 (.70)
Distance traveled (nm)	27.2 (.54)	27.1 (.42)
Number of communication step-ons	47.4 (8.5)	47.5 (7.2)

Furthermore, the predictability of the delays did not seem to affect our ATCos. This was determined by conducting a separate set of repeated measures ANOVAs, comparing the three Verbal Delay conditions (constant-short, constant-long, and variable delay). We found no differences between them in the number of LOS, in the time of aircraft through the sector, or in the number of step-ons (all *Fs*<1). The same results were obtained in a set of repeated measures ANOVAs comparing our three execution initiation delay conditions, i.e., constant-short, constant-long, and variable delay (all *Fs*<1). Our controllers were able to develop strategies for safely and efficiently managing traffic regardless of whether the delays in UAS responding were predictable or not.

Nonetheless, it is important to point out that although the sector metrics were not affected by our manipulations, they were correlated with each other: Number of LOS was significantly correlated with mean distance through the sector ($r = +.45$; $p < .001$) and with the number of step-ons ($r = +.47$; $p < .001$). The mean distance travelled by aircraft was also correlated with number of step-ons ($r = +.45$; $p < .001$).

As noted above, how ATCos rate the acceptability of specific latencies in responding may depend on broader sector characteristics rather than just the length of the delays themselves. This is reflected, for example, by the fact that the acceptability of the verbal and execution latencies of conventional pilots was highly related to that of the UAS pilots, even though only the latter delays were manipulated. It is also reflected in the fact that the acceptability of the long verbal latencies depended on the number of long verbal latencies that were present in the scenario. It is therefore likely that ATCo ratings of the acceptability of specific latencies in responding are complex judgments that involve integrating various contextual factors. To determine if ATCos used sector outcomes to determine the acceptability of verbal and execution latencies, correlations were computed separately for constant and variable delay conditions. Table 5 shows the relationships between acceptability ratings and performance for the constant delay conditions.

Beginning with the ratings of the UAS response latencies, the correlation between Verbal Delay acceptability and MR1 showed a trend, such that shorter latencies tended to be associated with higher acceptability ratings. The acceptability rating of Verbal Delay was also significantly correlated with the number of step-ons and average distance through the sector; as the number of step-ons and distance traveled decreased, the acceptability ratings increased. There was no significant correlation between Execution Delay acceptability and MR2. This is surprising because Verbal Delays and Execution Delays were crossed in our experiment. However, it may have been difficult for controllers to detect differences in the latency of command execution during the scenarios and their rating of the acceptability of execution delays may have been based instead on delays in verbal responding. Indeed, neither rating of acceptability was correlated with MR2.

Table 5. Correlations (and *p* Values) between Acceptability Ratings, Response Latencies, and Sector Metrics for Constant Delay Conditions

ATCo Ratings	*Verbal Response Latency (MR1)*	*Execution Response Latency (MR2)*	*Number Step-Ons*	*Number LOS*	*Distance Travelled*
UAS					
Acceptability rating of MR1	-.35 (.07)	-.04 (.83)	-.46 (.01)	-.09 (.67)	-.37 (.05)
Acceptability rating of MR2	-.34 (.08)	.20 (.32)	-.32 (.10)	-.007 (.97)	-.21 (.29)
Conventional Aircraft					
Acceptability rating of MR1	-.30 (.13)	-.02 (.93)	-.54 (.003)	-.40 (.04)	-.61 (.001)
Acceptability rating of MR2	-.18 (.36)	-.03 (.87)	-.48 (.009)	-.27 (.17)	-.42 (.03)

With respect to the acceptability ratings of verbal delays and execution initiation delays of conventional aircraft, neither was correlated with MR1 or MR2, the actual latencies in verbal responding and command execution observed in our experiment. However, the acceptability ratings of verbal response latencies were negatively correlated with the number of step-ons, number of LOS, and average distance aircraft traveled through the sector. As these metrics decreased (performance increased), acceptability ratings increased. In terms of the acceptability rating of conventional aircraft execution delays, these were negatively correlated with the number of step-ons and with the average distance traveled by aircraft through the sector. As the number of step-ons and distance traveled decreased, the acceptability ratings increased. The fact that this was obtained with conventional aircraft and not with the UAS may reflect the fact that none of the LOS involved the UAS.

For the Variable Delay conditions, UAS acceptability ratings of verbal response latencies and execution initiation latencies were uncorrelated with MR1, MR2 and all sector metrics as shown in Table 6. For conventional aircraft, the acceptability ratings of verbal delays were not related to either response component, but the acceptability ratings of execution initiation latencies were related to MR1 and MR2: as the MR 1 and MR 2 latencies decreased, acceptability ratings increased. Acceptability of execution initiation latencies were also negatively correlated with the number of step-ons. As mentioned above, finding correlations between the acceptability ratings and performance metrics for conventional aircraft but not the UAS is most likely due to the fact that there were many more conventional aircraft than UAS in the sector at any given time.

Table 6. Correlations (and p Values) between Acceptability Ratings and Response Latency for Variable Delay Conditions

ATCo Ratings	*Verbal Response Latency (MR1)*	*Execution Response Latency (MR2)*	*Number Step-Ons*	*Number LOS*	*Distance Travelled*
UAS					
Acceptability rating of MR1	-.18 (.36)	.21 (.28)	.04 (.86)	.10 (.61)	-.34 (.15)
Acceptability rating of MR2	-.05 (.81)	.29 (.14)	-.29 (.14)	-.01 (.96)	-.12 (.56)
Conventional Aircraft					
Acceptability rating of MR1	-.05 (.81)	-.29 (.13)	-.26 (.18)	.004 (.98)	-.08 (.67)
Acceptability rating of MR2	-.40 (.04)	-.50 (.007)	-.43 (.02)	-.17 (.40)	-.18 (.35)

3.4.3 Discussion

The present study examined the influence of additional 1.5 s or 5 s UAS pilot verbal and execution delays on ATCo acceptability ratings of responses by UAS pilots and conventional aircraft pilots when the UAS is operating in the NAS environment. We found that the short delays added to the UAS verbal response latencies, which led to an average UAS MR1 of 2.1 s, were generally acceptable to ATCos. They were also more acceptable than the long delays that led to an average UAS MR1 of 5.48 s, and which were only acceptable in approximately half of the scenarios. Although these response times were longer than those of the conventional aircraft (.81 s and 1.01 s, respectively) the acceptability ratings for the two types of pilots did not differ. We also found that the manipulation of UAS verbal delays did not affect the performance of ATCos in terms of LOS, mean distance traveled by aircraft, or the number of communication step-ons. Thus, in terms of our safety and efficiency metrics, our ATCos were able to adapt to the imposed verbal delays in responding by a single UAS.

The UAS MR1 latency that we observed in the present study under the long verbal delay condition was greater than that observed by Cardosi (1993) and Smith (2008), which is not surprising given that we added 5 s delays to the UAS responses. Under the Short Verbal Delay condition, however, our MR1 for UAS pilots was shorter than what Cardosi (1993) observed. For conventional aircraft verbal communication latencies in our simulation averaged between 1 s and .81 s. These response times are lower than those of Cardosi (1993) and similar to those observed by Smith (2008) in a TRACON sector. The relatively fast communication time is likely due to the fact that the pseudopilots in our simulation had fewer tasks to complete than the pilots actually flying in the Cardosi (1993) and Smith (2008) studies.

Interestingly, the ATCos in our study appeared to be more tolerant of delays in responding by UAS pilots than conventional aircraft pilots. This is evidenced by the fact that the acceptability ratings for both types of pilots were very similar, even though the responses of conventional aircraft pilots were two to five times faster than UAS pilots. If they had been equally stringent on both pilot types, the acceptability ratings for UAS pilots should have been much lower. One possible reason for this difference is that our ATCos may have assumed that the UAS pilots were less familiar with NAS operations, so may have expected them to take longer in responding, much like controllers often assume that new pilots or foreign pilots will take longer to respond to their verbal commands, or will make more mistakes in doing so (Cushing, 1995). It is also possible that the ATCos assumed that the differences in response times reflect characteristics of the UAS control and communication interface that pilots cannot do anything about.

Regarding latencies in execution initiation, we also found that our manipulations did not affect any of our ATCo performance metrics. However, unlike the verbal delay manipulation, adding a short execution initiation delay to the UAS, which led to an MR2 of 6.72 s, did not yield higher acceptability ratings than adding a long execution delay, which led to an MR2 of 10.28 s. Thus, UAS execution delays had less of an impact on ATCo acceptability ratings than verbal delays. In the debriefing session, ATCos indicated that they were not as concerned with the speed of the UAS execution because it was traveling at a slow speed of 110 kts. Due to the fact that most of the conventional aircraft were at speeds of 250 kts or higher, the ATCos moved the faster moving traffic in the sector when necessary, and created a “buffer” zone around the UAS. The ATCos did note that they were able to create this buffer zone because there was only 1 UAS in the sector. They indicated that the execution delay would likely have played a bigger role if the UAS was much faster. Moreover, if there were multiple UAS operating in the sector, the ATCos would have been less able to protect all UAS in the sector.

Another purpose of our study was to examine whether it is delays per se, or their predictability that affects ATCo performance and communication. To answer this question, we manipulated whether UAS verbal and command execution delays were constant or variable within scenarios. We did not, however, find that variable verbal and execution delays affected either the safety or efficiency with which controllers managed aircraft more than constant delays. One potential reason for this is that our simulation did not require multiple exchanges between the pilots and ATCos, unless step-ons occurred. This is because our pseudopilots did not engage in negotiations with ATCo, but rather were required to be compliant with their commands. Indeed, Rantanen et al. (2004) found that pilot verbal delays negatively affected ATCo performance only when multiple rounds of communication were required. In sum, although it appears as though the expert controllers in our study were able to compensate when verbal delays and execution delays were unpredictable, it remains to be seen whether this would continue to be the case if multiple communication exchanges were required, or if the sector workload increased.

Our study also revealed that ATCo acceptability ratings of MR1 and MR2 latencies for UAS flying in the NAS are complex judgments that reflect broader sector characteristics and depend on a variety of factors. First, in the present study we showed that the overall acceptability rating depended on the number of longer versus shorter latencies, with fewer longer latencies leading to higher ratings. This is because our variable latency conditions were rated equally acceptable as the scenarios that have constant short latencies, and the former have fewer of the long latencies than the constant-long scenarios. Second, ATCos determine acceptability of a delay within a scenario from the standpoint of all aircraft, as we found very little difference in the ratings of UAS versus conventional aircraft, even though delays were added only to the UAS responses. Although the ATCos judged the acceptability ratings separately for the UAS and conventional aircraft, it was done post-scenario, and this could have led to the highly correlated ratings. Third, the acceptability ratings were based on several factors in addition to the actual latencies produced by the UAS. These include number of LOS, number of step-ons, and mean distance travelled through the sector. When studying the measured responses of UAS in the NAS it is therefore critical that a system-wide perspective be adopted, examining how ATCos interact with other aircraft as well (Vu & Chiappe, 2015).

One important implication of these results is that an attempt to assign an upper limit on the latencies will be difficult. At the lower end, it was clear that a UAS MR1 of approximately 2 s was generally acceptable to ATCos. Indeed, this mean verbal latency is lower than that observed by Cardosi (1993) in en route sectors, so it is not surprising that it was acceptable to ATCos. However, we also want to know how long a verbal delay can be and still be acceptable to controllers. Our findings suggest that in finding an upper limit it may be difficult to identify a precise value due to the variability in responding. For example, we found that under long verbal delay conditions, MR1 was 5.48 s, but this was acceptable half of the time and unacceptable half of the time. Thus, rather than identifying a specific MR value for an upper limit, it may be more fruitful to require that time *intervals* be determined based on other factors influencing sector and ATCo performance.

3.4.3.1 Lessons Learned from Simulation B

- Delays in verbal responding are more important to ATCos than are delays in command execution, at least when UAS speeds are slow.
- Verbal response latencies of 2.1 s are generally acceptable to ATCos, though this likely applies to only sectors similar in characteristics to the one examined in the present simulation.
- Performance of ATCo duties in terms of safety and efficiency was neither affected by the manipulated delays in verbal responding and command execution, nor by the predictability of those delays.
- ATCos are more tolerant of delays in responding by UAS pilots than by pilots of conventional aircraft.

- ATCo judgments of the acceptability of latencies in verbal responding and initiation of command execution are complex judgments that reflect multiple features of the air traffic management situation, including measures of ATCo performance of their duties.
- Specifying an upper limit on what is an acceptable measured response will require taking multiple contextual factors into account.

3.5 Simulation C

The previous simulation had a single UAS flying in the sector at any given point in time, moving at a relatively slow speed of 120 kts. We found that ATCos were able to compensate for the presence of this vehicle in their airspace without compromising the safety and efficiency with which they managed traffic in the sector. This was the case despite variations in the length of the pilot verbal latency and in how long it took UAS pilots to begin executing maneuvers. Nonetheless, in a fully integrated airspace there will be multiple UAS in a given sector at any given time, each of which may vary in key response characteristics, including size, speed and maneuverability. This may make it more difficult for ATCos to simply create a buffer around the UAS where they move other aircraft around it, as they did in our study when there was only a single UAS in the sector.

Simulation C examined the effect of having multiple UAS in a simulated NAS environment. A short delay of 1.5 s was added to UAS pilot verbal communications. As we saw in Simulation B, this led to an MR1 of 2.1 s, which was generally acceptable to ATCos. In addition, we imposed a 1.5 s delay in initiating command executions on the UAS pilots, which had also led to an acceptable MR2 rating in Simulation B. In this simulation, we test the robustness of these ratings of the delays by examining whether they remain acceptable when the sector is more complex and challenging to ATCos. To this end, we manipulated the number of UAS in the airspace, ranging from 1, 2, or 4 such aircraft. These flew surveillance missions throughout the sector, which was also occupied by conventional aircraft. The speed of the UAS was also manipulated. They flew at either a “slow” speed of about 120 kts (e.g., representing the characteristics of a Predator) or a “fast” speed of about 240 kts (e.g., representing the characteristics of a Global Hawk). We assessed ATCo performance, communication in the sector, and acceptability ratings of UAS pilots’ and conventional pilots’ verbal response latencies (MR1) and command executions (MR2) as a function of the number of UAS and the UAS speed in the sector.

3.5.1 Method

3.5.1.1 Participants

Eight radar-certified ATCos volunteered to participate in this study. The ATCos averaged over 25 years of military and civilian experience in air traffic management. All had prior experience with ZLA airspace during their active air traffic management period. The simulation lasted two days.

3.5.1.2 Design

This simulation manipulated the Number of UAS (1, 2, or 4) and their Speed (slow, mixed, or fast) using a repeated measure design (Table 7).

Table 7. Illustration of the Design

	Speed		
Number UAS	*All Slow*	*Mixed*	*All Fast*
1	1 slow	1 (speed change during scenario)	1 fast
2	2 slow	1 slow 1 fast	2 fast
4	4 slow	2 slow 2 fast	4 fast

3.5.1.3 Simulation Environment

The study was conducted in three separate rooms: one room to manage the simulation and run the ATCo participants, a second room for conventional pseudopilots, and the third for UAS pseudopilots. All pseudopilots were trained experimental confederates.

Unlike Simulation B, which used MUSIM and MACS, this simulation was run entirely using the latter. MACS provided a display of the controller's radar scope for sector ZLA 20 (Figure 6). The multi-aircraft mode was used by pseudopilots to control all conventional aircraft. The single pilot mode was used for each of the UAS in the sector. The ATCo, conventional pseudopilots, and UAS pseudopilots spoke to each other using push-to-talk headsets over the voice server. The voice system was modified for the UAS stations. A fixed delay of 1.5 s was inserted before transmitting the UAS pseudopilots' message to the ATCo and conventional pseudopilots.

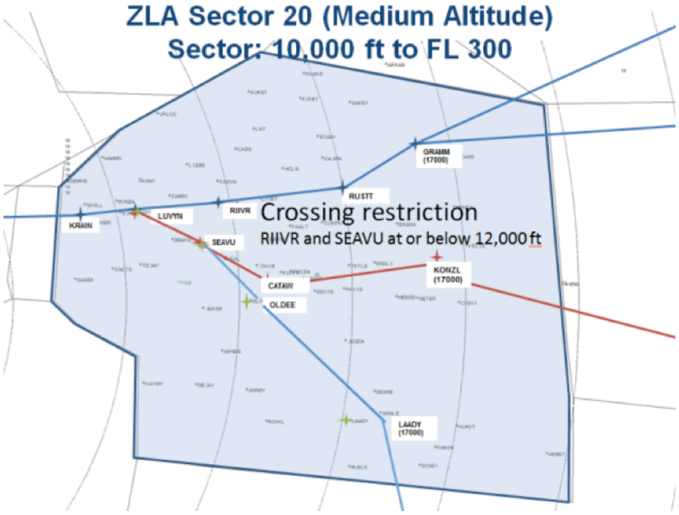
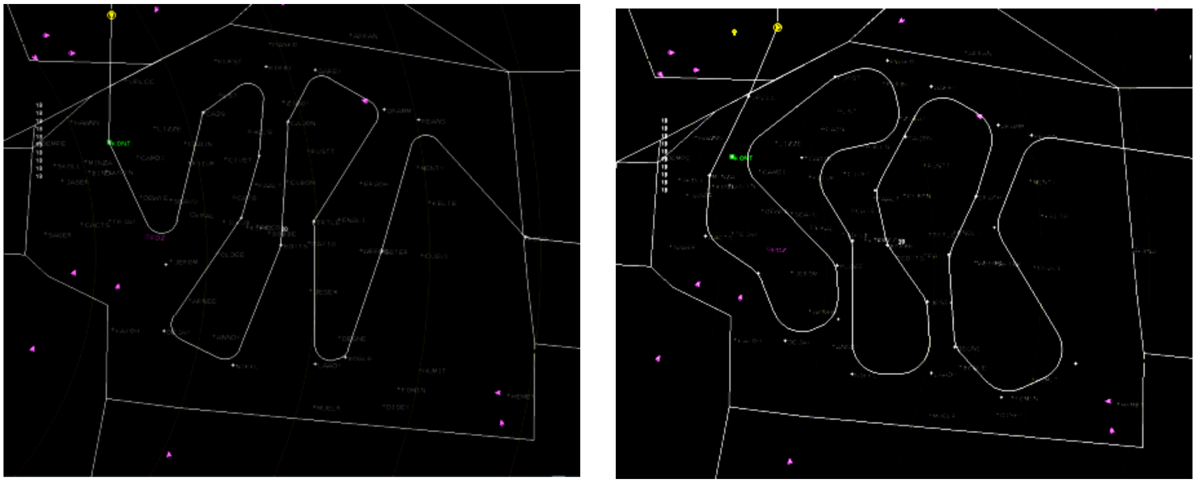


Figure 6. Illustration of sector ZLA 20 with three arrival paths.

ATCos were told that UAS pilots were flying “surveillance” missions, and that their flight paths would oscillate throughout the sector. Examples of flight paths for UAS are shown in Figure 7.



A *B*

Figure 7. Sample flight paths for 1 slow UAS (A) and 1 fast UAS (B).

Two parallel worlds were run at a time. Nine to 12 computers were required for each simulated world, depending on the number of UAS in the sector. The ATCo station included two computers, one to simulate the radar scope and a second small display acting as “Mission Control” to provide instructions to the controller regarding altitude changes to the UAS. Two stations were used for each conventional aircraft pseudopilot: one station was a “ghost” controller station that allowed them to check the traffic coming into their sector and the second to control all conventional aircraft within sector ZLA 20, with the exception of the UAS. One to four stations were used to control the UAS aircraft, depending on the number of UAS. Additional computers were used to manage the simulation software, record communication data between the controller and pseudopilots over the voice service, and for the Aeronautical Datalink and Radar Simulator (ADRS) simulation hub. As in Simulation B, the voice software was modified to simulate stepped-on transmissions so that when simultaneous transmissions occurred, they were unintelligible for everyone listening on the frequency.

The number of conventional aircraft in the scenarios was altered, depending on the number of UAS in the sector. In scenarios where there were 2 UAS, for example, there was one less conventional aircraft than in scenarios where there was only 1 UAS. This was done so that there would be the same number of aircraft in each scenario, the only difference being the number of which were UAS. This allowed us to examine the effects of UAS numerosity without the confounding variable of more or fewer aircraft overall in the sector.

3.5.1.4 Procedure

The simulation was conducted over two days for each participant. On the morning of the first day, participants completed consent forms, demographic questionnaires, and were briefed on simulation procedures. Following the briefing the controllers worked three practice trials and nine experimental trials for the remaining time. Each experimental trial lasted 40 min. During the trials, ATCos managed all air traffic coming into their sector. They were instructed that arrival traffic on approach to LAX had priority and that these aircraft had to leave the sector at an altitude no greater than

11,000 ft, and at airspeed of 250 kts. The ATCos were told that their sector had a Letter of Agreement to accommodate requests regarding the UAS in order to fulfill the UAS flight objectives while maintaining safe operations for all air traffic. All UAS were given a Predator call sign (e.g., PD-1) for the "slow" UAS and a Global Hawk call sign (e.g., GH-1) for the "fast" UAS.

Beginning approximately 1–2 min into each experimental trial and occurring every 4–5 min afterwards, either Mission Control or the UAS pseudopilot initiated a request to sector ATCos. Requests from Mission Control, simulated by the screen to the right of the ATCo radar scope, alerted the ATCo through his or her headset and display of a new, optimal altitude that the UAS should be flying to accomplish the mission objective (e.g., "PD-1 climb and maintain 14,000"). Mission Control only requested altitude changes for the UAS. During the next period of time, the UAS pseudopilot requested speed changes. A total of eight requests for all UAS were completed each trial, four initiated by the ATCo and four initiated by the UAS pseudopilot. When multiple UAS were in the sector, the total number of planned communications was divided evenly between each UAS. They did not increase overall with more UAS in the sector.

Voice delays were controlled at the UAS pseudopilot station. Voice software at this station automatically held the UAS's audio transmission for 1.5 s before broadcasting to the ATCo and conventional pseudopilots. The UAS station only included a transmission delay, not a receiving delay. The ATCo and conventional pseudopilot stations had no transmission delays.

After each trial, ATCos rated their situation awareness, workload, and quality of experience interacting with the conventional and UAS pseudopilots. They did so using an electronic version of the Situation Awareness Rating Technique (SART), NASA Task Load Index (NASA-TLX), and a 22-item questionnaire that included acceptability ratings of the verbal and execution initiation delays. Once all nine experimental trials were completed on the second day, ATCos answered post-simulation and debriefing questionnaires. Finally, the ATCos were interviewed during a debriefing session covering the same topics as the debriefing survey.

3.5.2 Results

We begin by analyzing the effects of our manipulations on the efficiency and safety with which ATCos performed their duties. Then we examine their effects on ATCo workload, situation awareness, and communication. Finally, we examine the ATCo acceptability ratings of UAS and conventional pilot verbal response and execution latencies.

3.5.2.1 ATCo Performance

A series of repeated-measures ANOVAs manipulating Number of UAS (1, 2, or 4) and Speed (slow, mixed, fast) were carried out on performance measures that include number of LOS and the time that conventional aircraft traveled through the sector. With respect to the LOS, a mean of 2.6 (SE = .28) occurred each scenario. Although the number of LOS did not differ significantly as a function of the variables examined (all ps > .20), the descriptive data do show that LOS increased with the number of UAS in the sector, and this linear trend was significant, $F(1,7) = 5.42$; $p = .05$ (Figure 8). A closer examination of the LOS data looked at the percentage of the incidents that involved a UAS. On average, 56% of the LOS involved these aircraft. Although not statistically significant (all ps > .16), we found that as the number of UAS increased, the percentage of LOS involving UAS increased as well, with approximately two-thirds of the LOS involving a UAS when there was 4 UAS in the sector (Figure 9).

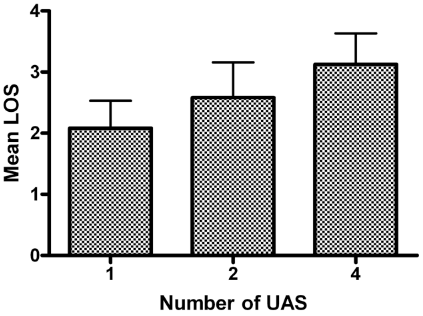


Figure 8. LOS frequency as a function of number of UAS in sector.

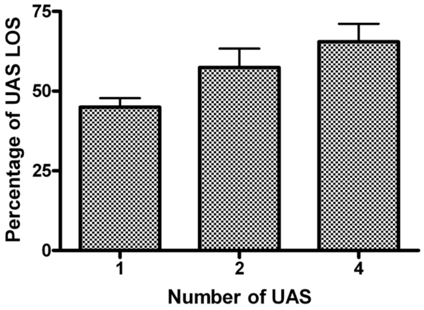


Figure 9. Percentage of LOS involving UAS as a function of number of UAS.

Turning now to a measure of the efficiency with which ATCos managed traffic, the amount of time that conventional aircraft traveled through the sector, we found a significant main effect of Number of UAS, $F(2,14) = 4.675$, $p = .028$ (Figure 10). Time through sector for conventional aircraft increased as the number of UAS in the sector increased, with 4 UAS having the greatest time through sector. This is noteworthy because when there were more UAS in the sector, there were also fewer conventional aircraft present. The main effect of Number of UAS is qualified by a significant Number x UAS speed interaction, $F(4,28) = 8.326$, $p < .001$ (Figure 11). When there was only 1 UAS in the sector, the time aircraft traveled in the sector did not differ depending on the UAS speed. However, when there were 2 UAS in the sector, slow UAS increased the average time that conventional aircraft traveled through the sector. In contrast, when there were 4 UAS in the sector, the slow UAS decreased the average time that aircraft traveled through the sector relative to mixed and fast UAS. In short, more efficient traffic flows occurred with faster UAS when there were only 2 UAS present, but with slower UAS when there were 4 UAS in the sector. It is possible that ATCos had to change their strategy for managing the conventional aircraft depending on the combination of the number of UAS present in the sector and how fast they were going.

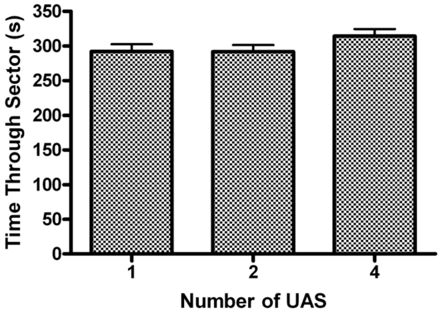


Figure 10. Conventional aircraft time through sector as a function of number of UAS.

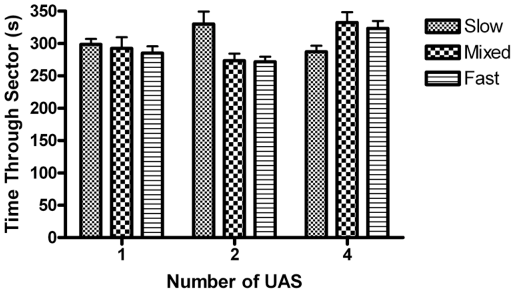


Figure 11. Time conventional aircraft traveled through sector by number and speed of UAS.

3.5.2.2 Workload

Workload was examined through the hand-off accept time (shorter delays reflecting lower workload), and by the NASA-TLX. The two are very different measures of workload, the former being objective, and based on measures collected throughout the trial. The latter is subjective and collected at the end of a trial. Both were analyzed with a repeated measures ANOVA that manipulated Number of UAS (1, 2, or 4) and Speed (slow, mixed, fast). For hand-off accept time, there was a main effect of Number of UAS, $F(2,14) = 10.536, p = .002$, where the hand-off accept time decreased as the number of UAS increased (Figure 12). This finding likely reflects sector characteristics that vary with number of UAS. In particular, since the number of aircraft in the sector was held constant at roughly eight, with 4 UAS there were fewer conventional aircraft entering the sector, thereby lowering the workload. The main effect was qualified by a significant Number x UAS Speed interaction, $F(4,28) = 3.16, p = .029$ (Figure 13). With 1 UAS in the sector, the hand-off accept time was shorter when the UAS was slow. With 2 UAS, faster speeds led to shorter hand-off accept times. With 4 UAS, the mixed condition was faster than the fast UAS condition.

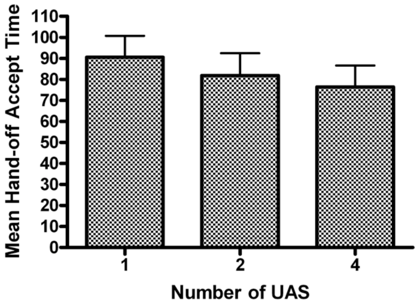


Figure 12. Hand-off accept time and number of UAS.

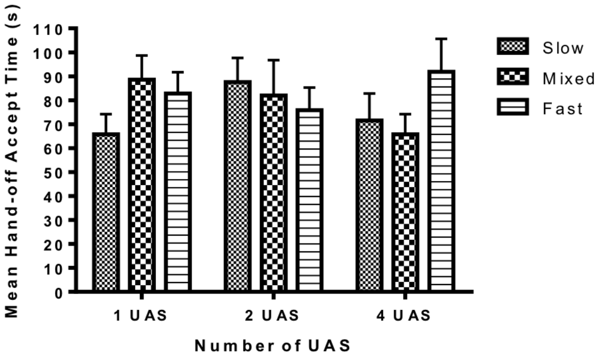


Figure 13. Hand-off accept time by speed and number of UAS.

An analysis of the subjective workload assessed by the NASA-TLX post-trial revealed a significant effect of Number of UAS, *F*(2,14) = 2.730, *p* = .025. However, in contrast to the hand-off accept times, subjective workload was rated the lowest when there were 2 UAS in the sector (*M* = 49) compared with 1 UAS (*M* = 56) or 4 UAS (*M* = 55) UAS. This likely reflects key contextual factors of the scenarios. When there was only 1 UAS in the sector, there were many more aircraft entering the sector than in the other conditions, thereby increasing the perceived workload. On the other hand, when there were 4 UAS, although there were fewer aircraft entering the sector than when there were 2 UAS, there was a greater number of LOS, as we saw above. Increasing the number of such critical incidents in the 4 UAS condition therefore could have affected the subjective assessment of the workload posed by this condition as ATCos more likely had to struggle to maintain safe separation minima. Thus, 2 UAS appeared to be the optimal number in the sector studied in this simulation.

3.5.2.3 Situation Awareness

Situation Awareness, assessed by combined scores on the SART, was analyzed using a repeated measures ANOVA that manipulated Number of UAS (1, 2, or 4) and Speed (slow, mixed, fast). We found a main effect of UAS Number, *F*(2,14) = 4.881, *p* = .025 (Figure 14). Post hoc tests revealed higher SA ratings in the 2 UAS condition compared to the 1 UAS condition. This may be due to the fact that the 2 UAS condition had the optimal level of workload compared to the other conditions, as revealed by the NASA-TLX ratings. SA may have been harder to maintain when there was 1 UAS in the sector because of the greater number of aircraft arriving in the sector, and in the 4 UAS condition, when the airspace was made very complex and ATCos had to work hard to resolve and prevent conflicts.

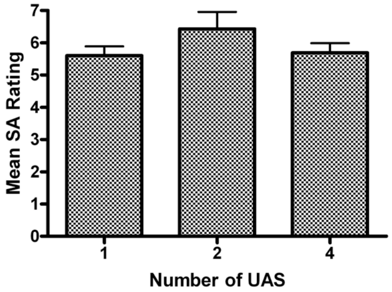


Figure 14. SART ratings as a function of number of UAS.

3.5.2.4 Communication

To analyze the number of stepped-on communications (which include communications of conventional aircraft pseudopilots, UAS pilots and ATCos), we examined the number of step-ons relative to the number of conventional aircraft in the sector. This was done because we expected that since most communications took place between ATCos and conventional aircraft pseudopilots, the more such aircraft are in the sector, the more potential for stepped on communications there were. A repeated measures ANOVA manipulating Number of UAS (1, 2, or 4) and Speed (slow, mixed, fast) revealed no main effects or interactions (all *p*s > .185). Thus, our manipulations did not affect the frequency with which step-ons occurred.

As verbal communication plays a central role in how ATCos manage traffic, we also set out to examine how UAS number and speed affected the number of communications, both between ATCos and UAS pilots, and between ATCos and conventional aircraft pilots. For UAS communications, we found a main effect of Number, with UAS communications increasing with a greater number of UAS in the sector, $F(2,14) = 5.96, p = .013$ (Figure 15). Post hoc tests revealed that there were more communications with 4 UAS in the sector than with 1 UAS or 2 UAS. No other effects were significant.

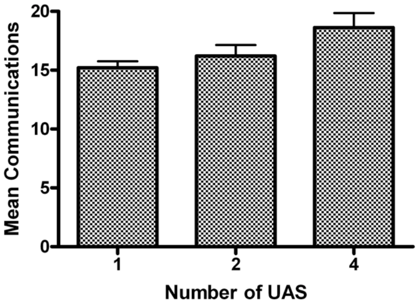


Figure 15. Number of UAS communications as a function of number of UAS.

With respect to ATCo-conventional aircraft pilot communications, a repeated measures ANOVA manipulating Number of UAS (1, 2, or 4) and Speed (slow, mixed, fast) revealed a main effect of Number of UAS, $F(2,14) = 59.09, p < .001$ (Figure 16). As the UAS increased, so too did the mean number of communications with conventional aircraft pilots, even though the number of these aircraft actually decreased. In particular, there were more conventional pilot-ATCo communications when there were 4 UAS in the sector than when there was 1 UAS or 2 UAS present. This implies that adding UAS to the sector increased the amount of verbal communications required with other pilots, likely because the presence of the former complicated the airspace for the conventional aircraft. There was also a significant effect of UAS Speed, with post hoc tests revealing that there were fewer communications with conventional aircraft when the UAS were slow, compared to the mixed and the fast UAS conditions, $F(2,14) = 46.06, p < .001$ (Figure 17). This effect was qualified by a significant interaction between Number of UAS and Speed, $F(4,28) = 12.16, p < .001$ (Figure 18). Post hoc tests revealed that the communication differences across different speeds were limited to when there were 4 UAS present in the sector.

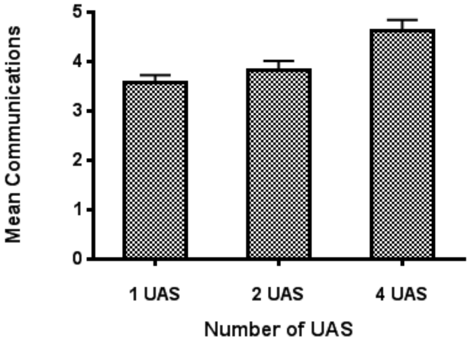


Figure 16. Communications with conventional aircraft as a function of number of UAS present.

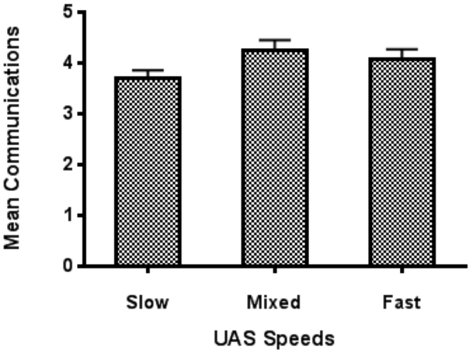


Figure 17. Communications with conventional aircraft as a function of UAS speed.

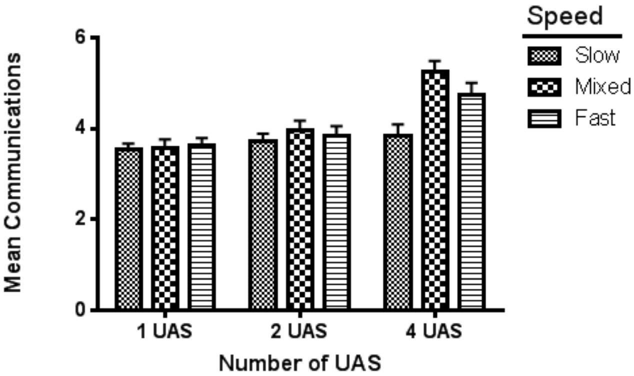


Figure 18. Communications with conventional aircraft as a function of UAS number and speed.

3.5.2.5 MR components

A verbal delay of 1.5 s and an execution initiation delay of 1.5 s were added to the responses of UAS pilots. Table 8 lists the resulting MR1 means for each type of pilot across the nine conditions, and Table 9 lists the MR2 components.

3.5.2.6 ATCo Acceptability Ratings

ATCos were asked to rate the acceptability of the delays in verbal responding (MR1) and command execution (MR2) by both, UAS pilots and conventional aircraft pilots, though a delay of 1.5 s was added only to the verbal responses and command executions of the former. Acceptability ratings for MR1 are listed in Table 8. On average, the verbal latencies were rated acceptable because all the means are greater than four on the seven-point scale. A Number of UAS (1, 2, or 4) X Speed (slow, mixed, fast) X Pilot Type (UAS pilot vs. conventional aircraft pilot) repeated measures ANOVA revealed a trend toward a significant effect of Pilot Type, with conventional aircraft pilots' verbal latencies being rated more acceptable than UAS pilots' verbal latencies (*Ms* = 5.8 and 5.2, respectively), *F*(1,7) = 4.169, *p* = .08. No other main effects or interactions were significant.

Table 8. Mean (and Standard Deviations) of Verbal Delays (MR1) in Seconds and Acceptability Ratings for MR1 as a Function of Experimental Condition

Condition	*UAS Verbal Delays (s) MR1*	*UAS MR1 Acceptability Rating*	*Conventional Verbal Delays (s) MR1*	*Conventional MR1 Acceptability Rating*
1 slow	2.00 (.19)	5.25 (1.04)	0.91 (0.27)	5.38 (1.60)
1 mixed	2.34 (.34)	4.88 (1.25)	1.10 (0.59)	5.25 (1.75)
1 fast	2.28 (.39)	4.88 (0.99)	1.17 (0.34)	6.25 (0.46)
2 slow	2.58 (.47)	5.50 (0.93)	0.85 (.50)	5.63 (1.69)
2 mixed	2.13 (.23)	5.13 (1.13)	1.17 (1.19)	5.63 (1.92)
2 fast	2.21 (.22)	5.25 (1.04)	0.91 (0.34)	6.38 (0.74)
4 slow	2.55 (.27)	5.13 (1.73)	1.40 (0.52)	5.75 (1.28)
4 mixed	3.03 (.96)	4.75 (1.83)	1.26 (0.72)	6.25 (0.71)
4 fast	2.31 (.25)	5.13 (1.46)	1.00 (0.39)	5.75 (1.04)
Overall	2.38	5.10	1.08	5.80

Note: Acceptability was rated on a scale of 1 = not at all acceptable to 7 = very acceptable.

Acceptability ratings for MR2 are listed in Table 9. Generally, the latencies were rated acceptable across all conditions, as the means are greater than four in all cases. A repeated measures ANOVA manipulating Number of UAS (1, 2, or 4), Speed (slow, mixed, fast) and Pilot Type (UAS, conventional aircraft) on MR2 revealed a significant interaction between Pilot Type and Number of UAS, $F(2,6) = 3.444$, $p = .065$ (Figure 19). Post hoc tests showed that conventional aircraft pilots were rated more acceptable, but only when there were 4 UAS present in the sector.

Table 9. Means (and Standard Deviations) for Execution Initiation Delays and Acceptability Ratings for MR2 as a Function of Experimental Condition

Condition	*UAS MR2*	*UAS MR2 Acceptability Rating*	*Conventional MR2*	*Conventional MR2 Acceptability Rating*
1 slow	7.36 (2.17)	5.25(0.46)	3.07 (1.83)	5.63(0.92)
1 mixed	7.70 (0.96)	5.13(1.55)	2.83 (1.25)	5.38(1.60)
1 fast	7.84 (1.37)	5.25(1.67)	3.04 (0.91)	5.88(0.99)
2 slow	9.11 (1.48)	5.25(1.28)	4.13 (4.24)	5.63(1.60)
2 mixed	7.78 (2.01)	5.50(1.07)	3.66 (1.56)	5.38(1.92)
2 fast	8.68 (1.42)	5.88(0.99)	3.39 (1.17)	6.38(0.74)
4 slow	10.01 (1.85)	4.00(2.07)	4.72 (5.81)	5.63(2.00)
4 mixed	9.46 (1.14)	4.88(1.64)	4.69 (5.25)	6.38(0.52)
4 fast	10.01 (1.82)	5.38(1.51)	3.98 (0.89)	6.00(0.93)
Overall	8.66	5.17	3.72	5.80

Note: Acceptability was rated on a scale of 1 = not at all acceptable to 7 = very acceptable.

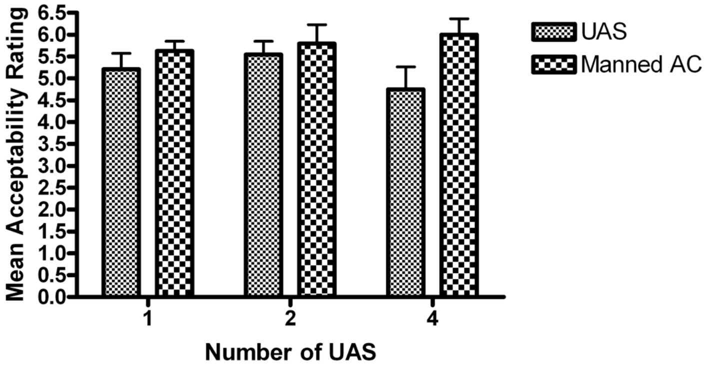


Figure 19. Acceptability ratings of MR2 as a function of number of UAS.

3.5.2.7 Post Scenario Measures

After each scenario, ATCos were asked a number of questions regarding their ability to manage traffic with the UAS in the sector. These were analyzed with a series of repeated measures ANOVAs manipulating Number of UAS (1, 2, or 4) and Speed (slow, mixed, fast). One question asked them to rate “How difficult was it to meet flow and separation requirements with a UAS in your sector compared to normal operations?” on a scale ranging from 1 (not at all) to 7 (very difficult). We found a marginally significant effect of Number of UAS, $F(2,14) = 3.241$, $p = .070$ (Figure 20). It was rated more difficult when there was 4 UAS in the sector compared to 1 UAS or 2 UAS.

Another question asked ATCos to rate "How often did you have to make special accommodations for the UAS to prevent conflicts?" using a scale ranging from 1 (never) to 7 (very often). We found a main effect of Number of UAS, *F*(2,14) = 4.661, *p* = .028 (Figure 21). Post hoc tests showed that this was more common when there was 4 UAS in the sector, compared to 1 UAS or 2 UAS.

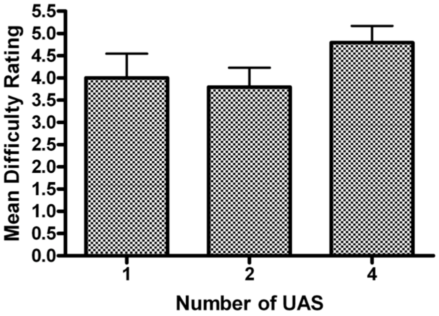


Figure 20. Difficulty of managing traffic in the sector as a function of Number of UAS.

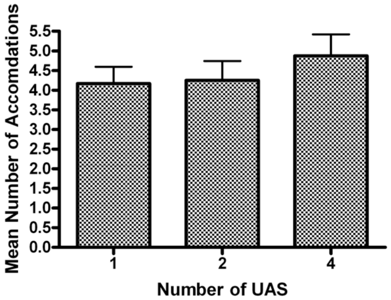


Figure 21. How often accommodations had to be made as a function of Number of UAS.

Finally, ATCos rated how many more UAS aircraft they could have managed in their sector at the same time if the same delays in UAS verbal responding were experienced, as well as how many more UAS they could have managed if the same UAS delays in response execution were experienced. For the question asking about the number of additional UAS if the same UAS verbal delays were in place, we found a main effect of Number of UAS, *F*(2,14) = 7.107, *p* = .007 (Figure 22). The number was lower when they had managed a scenario with 4 UAS than with either 1 UAS or 2 UAS. For the question regarding how many more UAS could be managed with the same execution latencies, we also found a significant effect of Number, *F*(2,14) = 4.789, *p* = .026 (Figure 23), with lower numbers given in the 4 UAS condition.

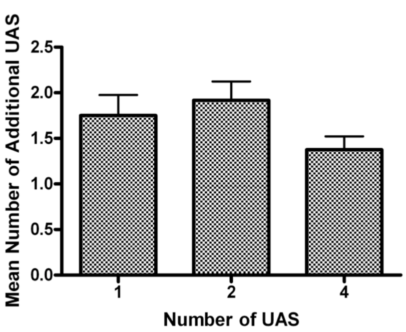


Figure 22. Number of additional UAS given same UAS verbal latencies.

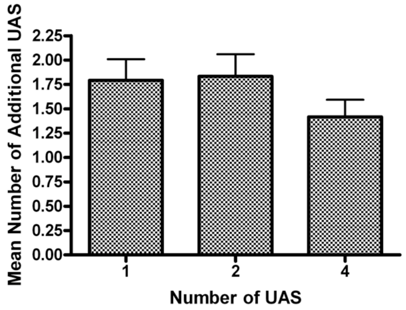


Figure 23. Number of additional UAS given same UAS execution latencies.

3.5.3 Discussion

This study examined the effects of adding 1.5 s delays in verbal responding (MR1) and in the initiation of command executions (MR2) to UAS responses on ATCos. We manipulated the complexity of the airspace by varying both, the number of UAS in the sector and also their speed. Simulation B used the same imposed delays and found these to be generally acceptable to ATCos, although lengthier delays, particularly in verbal responding, were not. However, because that study only had a single UAS in the sector, it is not possible to determine whether the obtained latencies in MR1 and MR2 continue to be acceptable in an airspace that is more complex.

In terms of performance of ATCo duties, our metric of safety was the number of LOS. We found that more than half of these involved a UAS, and that these were the greatest when there was 4 UAS in the sector. We note that although these effects were not statistically significant, given the seriousness of such incidents, the descriptive data urge caution regarding the safety of having more than 2 UAS in a sector with characteristics like the ones used in the present simulation. We did not find that the speed of the UAS had an effect on the number of LOS. Regarding ATCo efficiency, measured by the time taken for conventional aircraft to travel through the sector, we also found that 4 UAS led to the highest amount of time through the sector. Speed of the UAS mattered as well, at least when there was 2 or 4 UAS present. When there was 1 UAS in the sector, its speed did not affect the efficiency with which ATCos managed traffic.

ATCo workload was also affected by our manipulations. Although we found that hand-off accept times decreased the more UAS that were present in the sector, this is likely due to the fact that there were fewer conventional aircraft entering the sector under those conditions. Indeed, the subjective workload measures revealed that the optimal number of UAS in the sector was two. In the 1 UAS condition, there was a greater number of conventional aircraft entering the sector, which contributed to an increased ATCo workload. In the 4 UAS condition, the disruptive presence of these aircraft oscillating through the sector led to increased task demands, as ATCos had to work hard to prevent and resolve conflicts between these aircraft and the surrounding traffic. Interestingly, ATCos also rated their situation awareness as being the highest in the 2 UAS condition. Given our workload findings, this is likely because they had more resources available for acquiring and maintaining their understanding of the situation when there were 2 UAS present.

The communication data also suggests that having 4 UAS in the sector contributed significantly to the workload of ATCos. We found that when 4 UAS were present, ATCos had to communicate more often with pilots of conventional aircraft, even though there were fewer of them in the sector than in the 1 UAS and 2 UAS conditions. This increase in communications likely reflects the fact that the airspace was more complicated in this condition, with the flight plans of the UAS interfering with both arrivals and overflights in the sector. We did not find, however, that the number of step-ons was affected by either the speed or number of UAS. This is not surprising because overlapping transmissions are most likely due to the delays in transmissions (Nadler et al., 1993). That is, these are more likely to happen when there are lengthy transmission delays by some pilots. But, since our UAS verbal delays were short, and were held constant across conditions, we did not expect them to produce differences as a result of the manipulations.

A key goal of the present investigation was to determine whether the imposed delays of 1.5 s in verbal responding and initiation of command executions would be acceptable, regardless of the number and speed of UAS flying in the sector. Our results support this claim. ATCos rated the MR1 and MR2 latencies on average as being on the acceptable end of the scale (greater than 4, the middle of the scale), though they did tend to rate the MR1 of conventional aircraft pilots to be more acceptable than that of UAS pseudopilots. This is not surprising, given that the 1.5 s delay was only added to the latter. For the execution initiation delays, this effect of Pilot Type was only evident when there were 4 UAS present. This is consistent with findings from Simulation B, which showed that in general, delays in verbal responding are more salient to ATCos than delays in response execution.

Taken together, the results of the present simulation suggest that, for sectors with characteristics similar to those in place in the present study, while 1 UAS or 2 UAS may be acceptable, 4 UAS appears to be too many. This conclusion is also supported by the post-scenario questionnaires, which showed that under the latter conditions, it was the most difficult for ATCos to meet flow and separation requirements, and also the condition where the most special accommodations had to be made by them to prevent conflicts.

Although a fully integrated airspace will likely have more than one or 2 UAS in a sector at any given point in time, the present results suggest that ATCos may need extra support in order to handle those aircraft. This could come in the form of another person working with them on the sector, or possibly adding automation tools that aid in maintaining safe separation requirements. Of course, our conclusions need to be tempered by keeping in mind the fact that the UAS were flying a unique type of mission in our sector—a surveillance mission that involved flight plans that oscillated throughout the airspace. It is quite possible that ATCos would be able to manage more UAS in their sector if these aircraft were behaving more like other traffic.

3.5.3.1 Lessons Learned from Simulation C

- Additional verbal delays of 1.5 s were acceptable to ATCos regardless of the number and speed of UAS.
- Latencies in verbal responding are more salient to ATCos than latencies in initiating command executions.
- Additional delays of 1.5 s in initiation of command executions were acceptable to ATCos, regardless of number and speed of UAS.
- Although 1 or 2 UAS appears to be manageable, 4 UAS flying unpredictable missions in a complex traffic sector with arrivals and overflights appears to exceed the limits of what is manageable to ATCos.

3.6 Full Mission Simulation

The purpose of the Full Mission Simulation was to examine the effects of different command and control interfaces on UAS pilots' ability to respond to ATCo commands (see Rorie, Fern, & Shively, 2014). This allows us to examine, for example, what happens when a pilot who is operating "on-the-loop" (i.e., waypoint-to-waypoint/flight plan mode) needs to quickly get "in-the-loop" to respond to ATCo clearances. Although the main goal of the Full Mission Simulation was to examine the influence of different control modes, MR components 1 and 2 were captured and analyzed to examine the impact of the mode control on pilot verbal response latencies and execution initiation latencies. Moreover, the Full Mission Simulation allowed for other MR components to be captured as well. However, for consistency across the various simulations in this report, we will only focus on MR 1 and 2 components.

The three types of control modes that were contrasted include: (1) Waypoint-to-waypoint only; (2) Autopilot; and (3) Manual. In the Waypoint-to-Waypoint Mode (WP) the pilot could only change heading by modifying existing waypoints, and could use override to change altitudes. In the Autopilot Mode (AP), the aircraft retained WP functionality, but in addition could change heading and altitude using a new graphical interface. Finally, in the Manual Mode (M), the aircraft retained WP functionality, but in addition the pilot could change heading and altitude using stick and throttle inputs. Importantly, the pilots were able to use any input method available to them to implement a maneuver. For example, in Autopilot mode, the pilot could perform a vertical maneuver via waypoint edits or edits to the auto-pilot interface.

3.6.1 Method

3.6.1.1 Participants

The participants for the study were 15 RQ-4 Global Hawk pilots. Of these, six were qualified through RQ-4 Basic Training (Air Force Specialty Code 18X) without requiring them to have been previously qualified to fly manned aircraft (AC). Furthermore, nine were qualified through Undergraduate Pilot Training and were previously qualified to fly a manned AC. Finally, nine had previous experience flying UAS in civil airspace, with an average of 98 hrs. All of the participants had military combat and/or non-combat experience, with an average of 323 combined hrs. The study also used one retired ATCo with experience in Oakland Center airspace, who served as an experimental confederate.

3.6.1.2 Simulation Environment

This simulation was run using the Vigilant Spirit Control Station (VSCS) provided by the Air Force Research Laboratory Human Effectiveness Directorate (AFRL/RH) to simulate the UAS GCS (Figure 24), as well as MACS to simulate the airspace and air traffic environment, both for the pseudopilots and for the ATCo.

The scenarios simulated Oakland Center sector ZOA 40/41 (Figure 25) with current day IFR and Visual Flight Rules (VFR) traffic flows. The UAS mission scenarios were derived from FAA concept of operations (CONOPS) scenarios (combination of “Loiter for Surveillance” and “Grid Pattern”). The UAS started at FL190, and descended to 6,000 ft to conduct a stepped grid pattern search, after which it climbed back to FL190. ATCo clearances for traffic and weather were generated to force pilots to make quick control inputs.

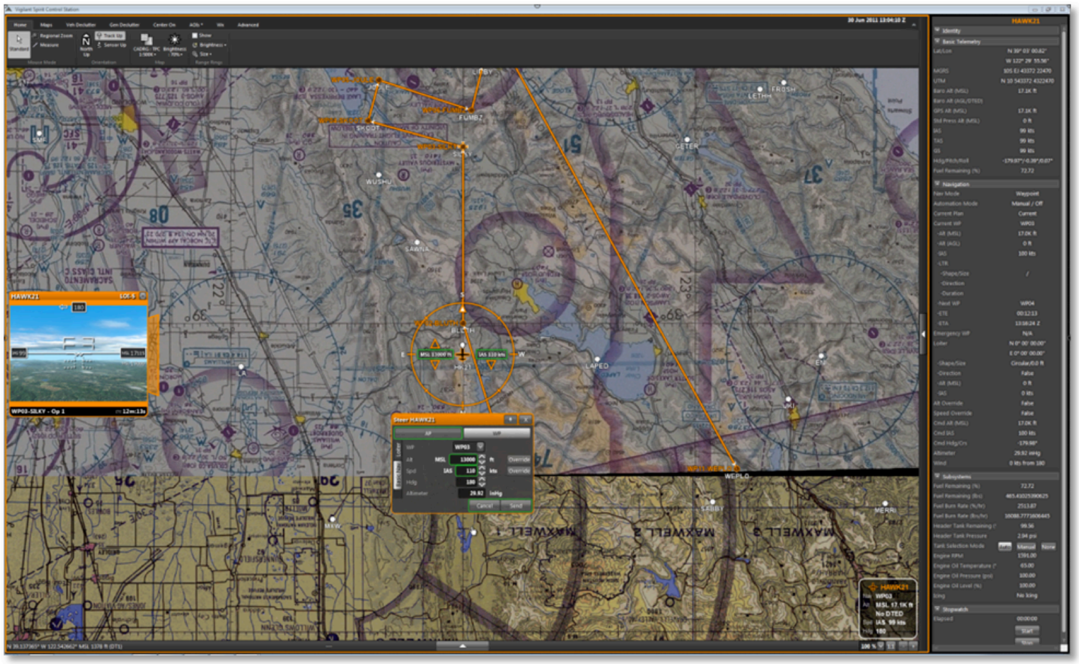


Figure 24. Vigilant Spirit Control Station (AFRL/RH). Distribution A: Approved for public release; distribution unlimited. 88ABW Cleared 3/18/2013; 88ABW-2013-1303.

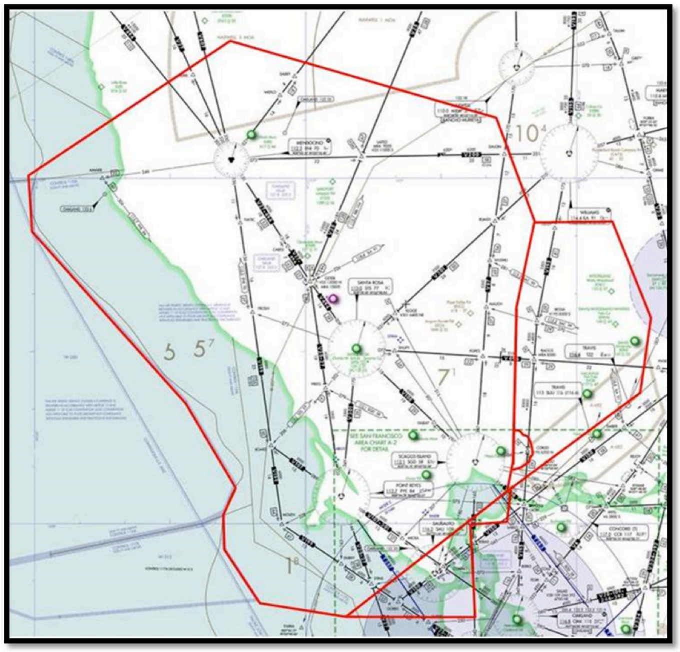


Figure 25. Oakland Center ZOA 40/41.

3.6.1.3 Procedure

The pilots were required to operate a simulated MQ-1 (HAWK21) along a pre-filed flight plan under Instrument Flight Rules. They were responsible only for air vehicle navigation (no sensor operation or system status monitoring tasks). While navigating their aircraft, they had to respond to ATCo traffic advisories (voice only), and respond to and comply with ATCo clearances for traffic and/or weather as necessary. As noted above, the pilots either responded with WP, AP, or M. In WP they could only edit headings by modifying existing waypoints and could use overrides to change altitudes. This was done through the Waypoint window. In AP they retained WP functions but could also edit headings and altitudes using the graphical interface. Edits were made through the steering window or compass rose. In M they retained WP functions but in addition could edit headings and altitude using stick-and-throttle controls.

Pilots were issued a total of 273 Traffic Advisories, with an average of six advisories per trial. There were also 767 Traffic Clearances, an average of 17 clearances per trial. The latter can be broken down by type as follows: (1) Altitude Clearances: 229; (2) Lateral Clearances: 300; and (3) ‘Direct To’ and ‘Resume Own Nav’ Clearances: 463. For the traffic advisories, no actions were required other than verbal responses. For the traffic clearances, pilots were required to immediately comply with the instructions.

3.6.2 Results

Pilots were allowed to use any method of input available to them in a given control mode condition. In the WP condition, there were 270 total edits. All edits were, by definition, made via waypoint edits or override functions. In the AP condition, there were 253 total edits. Of these, 109 edits (43%) were made via waypoint edits or override, and 144 edits (57%) were made via the autopilot interface (Compass Rose). In the M condition, there were 244 total edits, and 98 edits (40%) were made via waypoint edits or override, while 146 edits (60%) were made via stick and throttle.

In what follows, we report the results for both MR1 and MR2 as a function of Control Mode. As noted earlier, the full mission simulation did collect other MR components, but the focus of the current paper is on the pilot verbal response latency (MR1) and execution initiation latency (MR2).

3.6.2.1 Results for MR1

Control Mode was found to have a significant impact on the speed with which pilots replied to ATCo clearances and advisories, $F(2,28) = 6.87$, $p < .05$. Pilots had significantly shorter MR1 values in the AP mode than in the WP mode ($p < .05$; see Figure 26). No other differences were significant.

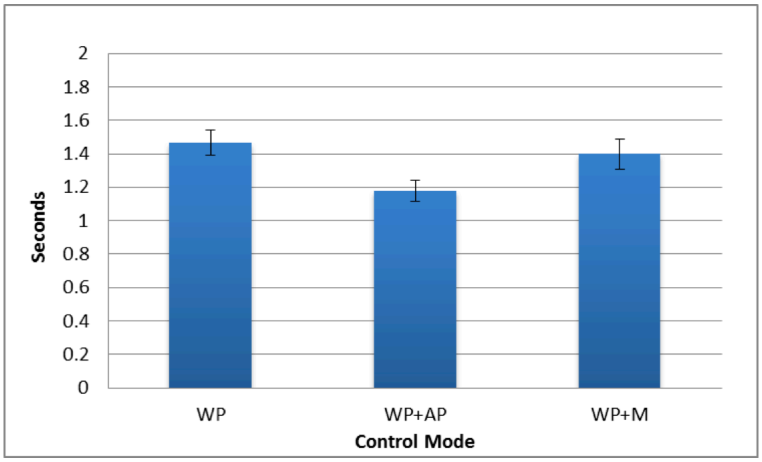


Figure 26. MR1 for the three control modes.

We also examined the distribution of verbal responses (Figure 27). Pilots replied to a total of 1,009 advisories and clearances. Of these, 50% of participants replied 2 s or sooner following the ATCo clearance, and 90% of participants replied 3 s or sooner following the clearance. These times do not include any events that were part of a 'combination maneuver'—that is, the controller issued a lateral and vertical maneuver within the same clearance.

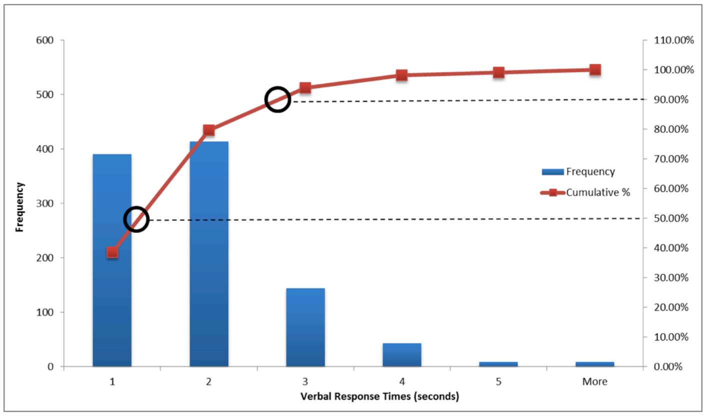


Figure 27. Distribution of verbal responses.

3.6.2.2 Results for MR2

Control Mode was also seen to have a significant effect on MR2 values, $F(2,28) = 14.05, p < .001$. AP mode was found to result in significantly shorter MR2 latencies than both the WP and M mode conditions ($p < .05$; see Figure 28). No other differences were significant.

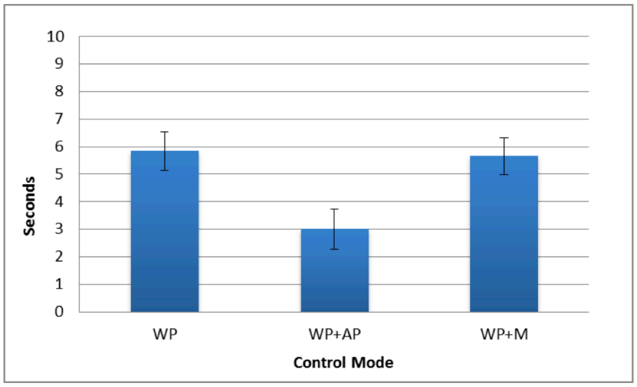


Figure 28. MR2 as a function of control mode.

We also examined the distribution of the MR2 data for the 549 edits initiated by pilots due to ATCo clearances (Figure 29). Of these, 50% of participants started their edit 5 s or sooner following the ATCos clearance, and 90% of participants started their edit at 11 s or sooner following the ATCos clearance.

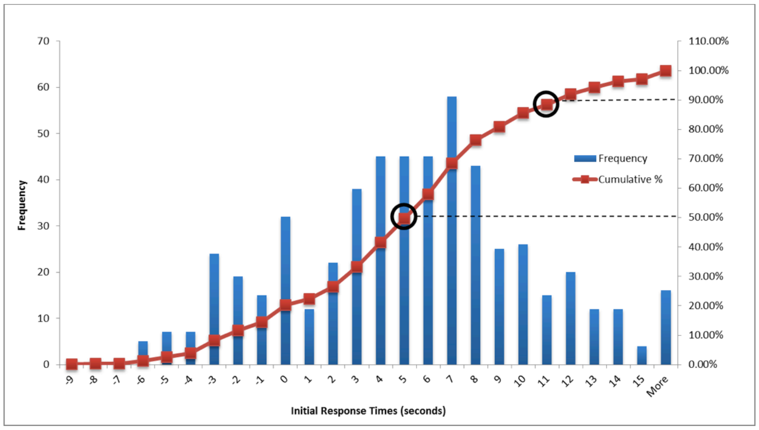


Figure 29. Distribution of response execution initiation latencies (MR2).

3.6.3 Discussion

This study examined the effect of three control modes on pilots' ability to comply with ATCo clearances. We found that mode of control did affect the timeliness of pilot responses. In particular, although the verbal latencies were generally fast, the AP condition led to the fastest MR1 responses. In this condition they could edit headings and altitudes using the graphical interface, in addition to using waypoint control functions. The AP condition also had significantly shorter response execution initiation latencies (MR2) than the other two conditions. Indeed, the MR2 latencies were almost twice as long for WP and M conditions relative to the AP condition. Evidently, the type of control interface is a crucial factor to consider when establishing the measured response of UAS to ATCo communications and clearances.

We also note that the MR1 component obtained in the present simulation is generally shorter than that obtained in our previous simulations. Specifically, with a mean MR1 of 1.35 s, this is faster than what we observed in the short delay (2.1 s) and the long delay (5.48 s) conditions of Simulation B discussed above. The MR2 component, however, at 9.1 s on average, is comparable to what we observed in the long delay condition of Simulation B (which had a latency of 10.28 s). We do note that this really depended on the type of control interface. For the WP mode, MR2 was longer than we observed, at 11.81 s. In contrast, the AP and M modes were both faster, especially the M mode at 6.9 s.

3.6.3.1 Lessons Learned from Full Mission Simulation

- AP condition had fastest MR1 latencies compared to WP and M modes.
- 90% of pilots replied verbally to ATCo advisories and clearances in 3 s or less.
- AP condition also had faster MR2 latencies compared to other two conditions.
- 90% of pilots initiated an edit in response to ATCo clearances in 11s or less.

4.0 General Conclusions

To conclude, these studies present an important step in identifying acceptable latencies in verbal responding (MR1) and initiation of command executions (MR2) for UAS operations in the NAS. Simulation A showed that MUSIM and MACS architectures can be used to extract the measured response components in a simulated NAS environment. These were further used in Simulation B, which found that short delays in verbal responding of approximately 2.1 s were generally acceptable to ATCos, and that delays in verbal responding are more salient to ATCos than delays in initiating command executions. This simulation also found that the predictability of the delays did not affect the safety and performance with which ATCos performed their duties. Furthermore, the study also found that ATCo ratings of acceptability of delays in pilot responding are complex judgments that involve weighing several sector characteristics, and do not merely reflect the duration of the delays. Simulation C extended these findings by examining whether the acceptable short latencies identified in Simulation B would continue to be acceptable in a more complex airspace that varied in the number and speed of the UAS. It was found that although ATCo performance was generally robust regardless of our manipulations, having 4 UAS flying complex routes did compromise the safety of the airspace. Finally, the Full Mission simulation showed that the type of control interface employed by UAS pilots affected the MR1 and MR2 latencies observed, with Waypoint control mode, for example, yielding the longest MR2 latencies. Thus, future work on the MR needs to take into account aircraft model characteristics before making general claims about how UAS interact with ATCos.

5.0 References

Agbolosu-Amison, S., Millner, D., Baden Jr., B., Coville, G., & Mondoloni, S. (2012). *Examining arrival time predictability in the NAS using monte carlo methods*. The MITRE Corporation Technical Report.

Askelson, M. A., Dreschsel, P., Nordlie, J., Theisen, C., Carlson, C., Woods, T., Forsyth, R., & Heitman, R. (2013). MQ-9 unmanned aircraft responsiveness to air traffic controller commanded maneuvers: Implications for integration into the national airspace system. *Air Traffic Control Quarterly*, *21*, 79-92.

Blickensderfer, B., Buker, T. J., Luxion, S. P., Lyall, B., Neville, K., & Williams, K. W. (2012). The design of the UAS ground control station: Challenges and solutions for ensuring safe flight in civilian skies. In *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting* (pp.51-55). Santa Monica, CA: Human Factors and Ergonomics Society.

Cardosi, K. M. (1993). Time required for transmission of time-critical air traffic control messages in an en route environment. *The International Journal of Aviation Psychology*, 3(4), 303–314.

Clark, A. (2003). *Natural born cyborgs: Minds, technologies and the future of human intelligence*. New York: Oxford University Press.

Consiglio, M., Hoadley, S., Wing, D., Baxley, B., & Allen, D. (2008). Impact of pilot delay and non-responsiveness on the safety performance of airborne separation. *Proceedings of the 8th AIAA ATIO Conference*, September 2008.

Cushing, S. (1995). Pilot-air traffic control communications: It's not (only) what you say, it's how you say it. *Flight Safety Foundation*, *14*, 25-31.

Dillingham, G. L. (2013). *Unmanned aircraft systems: Continued coordination, operational data, and performance standards needed to guide research and development* (Report no. GAO-13-346T). Washington, DC: Government Accountability Office.

Fern, L., & Shively, R. J. (2009). A comparison of varying levels of automation on the supervisory control of multiple UASs. In *Proceedings of AUVSI's Unmanned Systems North America 2009*, Washington, D.C.

Ferrell, W. R. (1965). Remote manipulation with transmission delay. *IEEE Transactions on Human Factors in Electronics*, *6*, 24–32.

Hart, S. G., & Staveland, L. E. (1987). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, The Netherlands: Elsevier.

Held, R., Efstathiou, A., & Greene, M. (1972). Adaptation to displaced and delayed visual feedback from the hand. *Journal of Experimental Psychology*, *72*, 887-891.

Hopkin, V. D. (1995). *Human Factors in Air Traffic Control*. New York: Taylor & Francis.

ICAO. (2011). *Unmanned Aircraft Systems (UAS)* (ICAO Cir 328). Montreal, Canada: International Civil Aviation Organization.

McCarley, J. S., & Wickens, C. D. (2005). *Human Factors Implications of UAVs in the National Airspace* (Technical Report AHFD-05-05/FAA-05-01). Savoy, IL: University of Illinois, Institute of Aviation, Aviation Human Factors Division.

Merlin, P. W. (2009). *Ikhana unmanned aircraft system Western states fire missions*. Washington, DC: National Aeronautics and Space Administration.

Merlin, P. W. (2013). *Crash course: Lessons learned from accidents involving remotely piloted and autonomous aircraft*. Washington, DC: National Aeronautics and Space Administration.

Nadler, E., Mengert, P., DiSario, R., Grossberg, M., & Spanier, G. (1993). Effects of satellite- and voice-switching-equipment transmission delays on air traffic control communications. *The International Journal of Aviation Psychology, 3*, 315-325.

Prevot, T. (2002). Exploring the many perspectives of distributed air traffic management: The Multi Aircraft Control System: MACS. *International Conference on Human-Computer Interaction in Aeronautics, HCI-Aero 2002*. Cambridge, MA: MIT Press.

Rantanen, E. M., McCarley, J. S., & Xu, X. (2004) Time delays in air traffic control communication loop: Effect on controller performance and workload. *The International Journal of Aviation Psychology, 14*(4), 369-394.

Rorie, R.C., Fern, L., & Shively, R.J. (2014). UAS measured response: The effect of GCS control mode interfaces on pilot ability to comply with ATC clearances. In *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting*. Chicago, IL.

Sheridan, T. (1992). *Telerobotics, automation, and human supervisory control*. Cambridge, MA: MIT Press.

Sheridan, T. (1993). Space teleoperation through time delay: Review and prognosis. *IEEE Transactions on Robotics and Automation*, *9*, 592-606.

Shively, R. J., Vu, K.-P. L., & Buker, T. J. (2013). Unmanned aircraft system response to air traffic control clearances: Measured response. In *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting* (pp.31-35). Santa Monica, CA: Human Factors and Ergonomics Society.

Smith, E. C. (2008). Analysis of controller-pilot communication performance in Area Navigation (RNAV) and conventional arrival operations. *Digital Avionics Systems Conference,* IEEE/AIAA.

Sollenberger, R. L., McAnulty, D. M., & Kerns, K. (2003). *The effect of voice communications latency in high density, communications-intensive airspace* (Report no. DOT/FAA/CT-TN03/04). Atlantic City International Airport: Federal Aviation Administration William J. Hughes Technical Center.

Taylor, R. M. (1990). Situational Awareness Rating Technique (SART): The development for a tool for aircrew systems design (AGARD-CP-478). In *Situation awareness in aerospace operations* (pp. 3/1–3/17). Neuilly Sur Seine, France: NATO-AGARD.

Vu, K.-P. L., & Chiappe, D. (2015). Situation awareness in human systems integration. In D. Boehm-Davis, F. Durso, and J. Lee (Eds), *APA Handbook of Human Systems Integration* (pp. 293-308). Washington, D.C.: American Psychological Association.

Vu, K.-P. L., Chiappe, D., Morales, G., Strybel, T. Z., Battiste, V. Shively, J., & Buker, T. J. (2015). Impact of UAS Pilot Communication and Execution Latencies on Air Traffic Controller's Acceptance of UAS Operations. *Air Traffic Control Quarterly*, *22*, 49-80.

Vu, K.-P. L., Morales, G., Chiappe, D., Strybel, T. Z., Battiste, V. Shively, J., & Buker, T. J. (2013). Influence of UAS pilot communication and execution delay on controller's acceptability ratings of UAS-ATC interactions. *Proceedings of the 32nd Digital Avionics Systems Conference* (pp. 6D4-1- 4.D.4-13). Syracuse, NY: IEEE.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Williams, K. (2007). *An assessment of pilot control interfaces for unmanned aircraft*. (Report no. DOT/FAA/AM-07/8). Washington, DC: Federal Aviation Administration.

Williams, K. W. (2012). *An Investigation of sensory information, levels of automation, and piloting experience on unmanned aircraft pilot performance* (Report no. DOT/FAA/AM-12/4). Washington, DC: Federal Aviation Administration.

Wourms, D. F., Ogden, J., & Metzler, T. R. (2001). *NEXCOM auditory latency, Volume I: Noncopyrighted literature search results* (Report No. HSIAC–SS–2001–005). Wright Patterson Air Force Base, OH: The Human Systems Information Analysis Center.

Yang, J., Rantanen, E. M., & Zhang, K. (2010). The impact of time efficacy on air traffic controller situation awareness and mental workload. *International Journal of Aviation Psychology*, *20*, 74-91.

Ziccardi, J., Roberts, Z., O'Connor, R., Rorie, C., Morales, G., Battiste, V., Strybel, T. Z., Chiappe, D., Vu, K.-P. L., & Shively, J. (2013). Measuring UAS pilot responses to common air traffic clearances. In S. Yamamoto (Ed.): HIMI/HCII 2013, Part II, *Lecture Notes in Computer Science*, *8017*, 606-612.