

Evaluation of Early Ground Control Station Configurations for Interacting with a UAS Traffic Management (UTM) System

Arik-Quang V. Dao¹, Lynne Martin¹, Christoph Mohlenbrink², Nancy Bienert²,
Cynthia Wolter², Ashley Gomez², Lauren Claudatos², and Joey Mercer¹

¹ NASA, Moffett Field, CA, USA

{Quang.V.Dao, Lynne.Martin, Joey.Mercer}@NASA.gov

² San Jose State University, San Jose, CA, USA

{Christoph.P.Mohlenbrink, Nancy.Bienert, Cynthia.Wolter, Ashley.N.Gomez,
Lauren.E.Claudatos}@NASA.gov

Abstract. The purpose of this paper is to report on a human factors evaluation of ground control station design concepts for interacting with an unmanned traffic management system. The data collected for this paper comes from recent field tests for NASA's Unmanned Aircraft System (UAS) Traffic Management (UTM) project, and covers the following topics; workload, situation awareness, as well as flight crew communication, coordination, and procedures. The goal of this evaluation was to determine if the various software implementations for interacting with the UTM system can be described and classified into design concepts to provide guidance for the development of future UTM interfaces. We begin with a brief description of NASA's UTM project, followed by a description of the test range configuration related to a second development phase. We identified (post hoc) two classes in which the ground control stations could be grouped. This grouping was based on level of display integration. The analysis was exploratory and informal. It was conducted to compare ground stations across those two classes and against the aforementioned topics. Overall, subjective ratings showed no differences with respect to workload and communication, but ratings for situation awareness and effectiveness of the procedures favored integration of displays.

Keywords: Unmanned Traffic Management · Human-systems Integration · Systems Engineering · Human Factors

1 Introduction

Unmanned aircraft, commonly referred to as "drones", are aircraft designed to operate with the absence of an onboard pilot. These aircraft may be part of an unmanned aircraft system (UAS) that allows a pilot to manually control the vehicle remotely, or

provide strategic guidance when the aircraft is autonomous. To name a few, UAS have applications in search and rescue, infrastructure inspections, goods delivery, recreation, and media. Due in part to the diversity of such applications and their relative affordability, the market for UAS is expected to grow significantly. The Federal Aviation Administration (FAA) has projected that sales of UAS of all sizes and types in the United States will grow from 2.5 million in 2016 to 7 million units in 2020 [1]. If these aircraft are eventually deployed, the total number of concurrent UAS operations in the U.S. is expected to be up to 250,000 by 2035 - approximately 175,000 of those will be for commerce alone [2].

Early experience from manned aviation has made the case for an appropriate level of organization, if safety is to be achieved with increasingly congested air traffic [3]. This meant that the air traffic had to be managed. For manned air-traffic, this is done through a system where human controllers are responsible for maintaining separation between aircraft. This system has proven to be very safe. The disadvantage is that the capacity of the airspace will be limited by human cognitive resources - mainly workload. NASA studies, e.g. [4], conducted to examine solutions for meeting growing demand for air transportation services found that sustained high capacity is achievable if automation assisted controllers in separating aircraft. We expect these findings to remain relevant to UAS operations. However, a system for managing UAS operations will need to consider accommodating an unmanned aircraft fleet size that is forecast to be 35 times that of manned aircraft in 2020. Given these considerations, NASA's vision for UAS Traffic Management (UTM) will not mirror the traditional manned air traffic management system. Instead, UTM will be designed to allocate active management of aircraft from human controllers to automation. Humans in this system will then serve as supervisors providing strategic level input [3]. Beyond the U.S., the European aviation community has also recognized the importance of UTM, and acknowledged that rising demand for UAS operations could be addressed by UTM and related technologies. NASA's UTM concept of operations offers initial guidance for testing such technologies.

According to the NASA UTM concept of operations, the UTM will be designed to safely enable large-scale small UAS (i.e., unmanned aircraft less than 50 pounds) operations in low altitude (i.e., below 500 feet) Class G airspace [3]. This will be achieved by providing technical capabilities to UAS operators and stakeholders. Technical capabilities will take the form of information products and services. These products and services will be tested as part of the UTM concept in NASA's research platform [5, 6]. The research platform supports research with both live and simulated aircraft. The simulation capabilities have enabled the testing of off-nominal interactions between virtual and live aircraft in field tests, and it will be the primary tool for feasibly evaluating large scale UAS operations as the tests grow in complexity. This increasing complexity, as well as scope, is reflected in NASA's plan for testing the UTM concept - distinguished by four technical capability levels (TCL). TCL 1 and 2 have already been tested.

In TCL 1, operations were conducted over remotely populated areas (e.g., rural operations) [7]. The traffic density was very low, with 4 aircraft available in early field tests

and at most 2 concurrent operations; each in its own volume that reserved all airspace above and below. The aircraft remained within visual line of sight throughout the operation. The UTM services provided only vetting of operations against conflicts between operations and other constraints, such as existing airports; the information that it provided conveyed whether an operation was accepted into the system and if it was rejected, why it was rejected.

In TCL 2, the technical capabilities from TCL 1 were carried over and the concept was extended to the evaluation of industrial applications of UAS operations over sparsely populated areas. It included a mixture of visual line of sight and beyond visual line of sight operations. TCL 2 included other enhancements such as: alerting for airspace intrusion, alerts to contingency management, and segmented flight planning that allowed stratification of operational volumes among other things. TCL 3 and 4 have not been conducted yet. They will build on the capabilities of TCL 1 and 2 to include: operations over increasingly populated areas, moderate and then high UAS traffic densities, interactions between manned and unmanned operations, as well as large-scale contingency management. The reader is encouraged to see Johnson et al. [7] and Kopardekar et al. [3] for a more detailed discussion of NASA's concept of operation and test plans. The findings reported in this paper will focus on the most recent test - TCL 2.

In TCL 2 we advanced the concept by introducing the ability for operators to plan and schedule beyond visual line of sight (BVLOS) operations. To accomplish BVLOS, a number of capabilities had to be in place for the UAS operators. They needed displays that provided information about where their aircraft were relative to operational boundaries, and other air traffic. This was accomplished either on a few integrated displays or across separate displays. The configuration of these displays as part of the floor plan for the UAS ground stations influenced the size of the flight crews and their roles. In this paper, we describe how our human factors measures varied as a function of these ground station configurations, and offer some interpretation on how those differences reflected their effectiveness in the field tests. The analysis we presented here was informal and intended to be exploratory. In the next section we consider the field test configuration and the UTM architecture before discussing the ground control stations and the human factors measures used to evaluate them.

2 Test Range and Scenarios

TCL 2 flight tests were bound to uncontrolled airspace 2 miles north of the active runway at Reno-Stead Airport (RTS). The range was a flat, dry, desert basin surrounded by steep mountains. UAS flight crews were positioned between 4 of 5 total ground control station (GCS) locations in the area (Figure 1); simulated aircraft were injected into the test from GCS 1 when the test scenarios required it. The flight test director (FTD) coordinated flights on the field, and was located immediately east of GCS 3 on the south end of the test range.

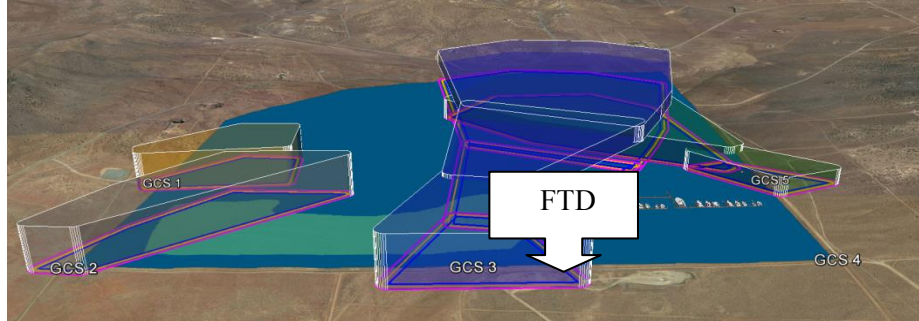


Fig. 1. UAS test range north of Reno-Stead Airport. Three-dimensional extrusions above the map represent operational boundaries for each of the ground control stations.

Four different scenarios, each with a different back story, prompting different combinations of events, were performed. Scenarios shared at least one event (Table 1) and were designed to represent interactions likely to occur if UTM was to be implemented in the future. These interactions included those that involved intruder aircraft and contingency management operations. Scenarios were 30 minutes long with up to 5 concurrent flights over the test range. Actual flight durations were between 6 and 23 minutes. The flights took place over the course of 9 days, and daily between 8:30 am to 12:30 pm to take advantage of favorable weather and wind conditions.

Table 1. Flight Test Scenarios

	Scenario 1 Agriculture	Scenario 2 Lost Hiker	Scenario 3 Ocean	Scenario 4 Earthquake
BVLOS	X	X	X	X
Multiple BVLOS	X		X	
Altitude Stratified VLOS	X	X		X
Altitude Stratified BVLOS			X	
Intruder Aircraft Tracking	X		X	
Intruder Aircraft Conflict Alert	X		X	
Rogue Aircraft Conflict Alerts	X			
Dynamic Re- routing		X		X
Contingency Management Alerts			X	X
Public Safety Operation		X		
Simulated Aircraft		X		X

3 Flight Crew Roles and Responsibilities

A total of eleven flight crews participated over the duration of the TCL 2 flight tests. Two of those were NASA crews. The remaining 9 came from different industry partners. Crew-members were composed of individuals who operated together regularly. The size of regular crews varied between 2 and 5 members, where the most common crew size was four. Larger size crews of 5 were able to assign just one role to each member. Individuals in smaller crews served more than one role. The available roles were; pilot-in-control (PIC), client operator, ground control station operator (GCSO), on-site software engineer, and launch technician (Table 2). On-site software engineers were not required, but were present in some flight crews. Launch technicians were a practical requirement of the type of aircraft platform, i.e., fixed-wing aircraft. Auxiliary members expanded the regular crew to provide support specific to the TCL 2 flight test, but on occasion assist in pre-flight and post-flight procedures. These roles were; UTM representative, visual observer, observer controller, and human factors observer. Table 3 shows how the aforementioned roles were distributed across crew members.

Table 2. An exhaustive list of TCL 2 flight crew roles.

Regular roles	Auxiliary roles
a) Pilot-in-control (PIC)	f) UTM representative
b) Client operator	g) Visual observer
c) Ground control station operator (GCSO)	h) Observer controller
d) On-site software engineer	i) Human factors observer
e) Launch technician [fixed-wing only]	

Table 3. Distribution of roles across flight crews.

	FC1	FC2	FC3	FC4	FC5	FC6	FC7	FC8	FC9	FC10	FC11
1	a	ac	a	a	a	a	abc	a	a	a	ac
2	b	e	c	b	c	c	c	bc	bc	bcd	b
3	h	bf	bf	c	bd	bf	f	f	f	e	f
4	g	g	g	f	f	d	g	g	d	f	e
5	i	h	h	g	g	g	h	d	d	g	e
6		i	i	h	h	h	i	h	g	h	d
7				i	i	i		i	h	i	g
8									i		h
9											i

*Columns headers are for flight crew. Rows represent individual members. The letters within cells are assigned roles as defined in Table 2. Some members supported more than one role.

The PIC was responsible for the operation of the aircraft, and had ultimate authority over the operation of the aircraft. For two flight crews, the PIC doubled as a GCSO, where the aircraft flew autonomously on a pre-defined flight plan. The client operator created, and interfaced with the UTM system to submit flight geometries, as well as send and receive notifications. Examples of the geometries submitted by client operators are illustrated in Figure 1 as 3-dimensional extrusions above the map. The GCSOs coordinated with the PIC to run pre-flight and post-flight checklists; they generated the flight plans and monitored the aircraft from their GCS display during the operations. In some cases, on-site software engineers also served as GCSOs and

client operators on top of providing expertise in the troubleshooting and tweaking of proprietary clients. Launch technicians assembled and directed aircraft launch hardware for catapult fixed-wing platforms, and assisted with retrieval of aircraft upon their return and landing.

UTM representatives doubled as client operators by relaying UTM system information to the flight crew as necessary; their permanent role was to coordinate with the flight test group for adherence to research protocols. The visual observer visually tracked the aircraft to assist with separation between other aircraft, fauna, and terrain. The observer controller maintained radio contact with the test site authorities and the flight test director to coordinate and acquire authorization for take-off at their specific location. In most cases human factors observers watched passively and interacted with the flight crew for data collection purposes, but in a few instances they also provided assistance to the flight crew for miscellaneous tasks, such as equipment setup and packing.

4 Ground Control Station Configurations

The crews operated a mixture of fixed-wing and rotor-wing UAVs. Aircraft flew autonomously on routes defined through a mission planner or manually for take-off and landing by the pilot-in-control. UAS ground stations interacted with the UTM system via a client. All of the clients deployed by the flight crews were in their early development states. At minimum, the clients were required to be able to submit a flight volume to the UTM system and then receive and display messages from the system. For NASA flight crews, the UTM representative submitted volumes on their behalf and verbally relayed the messages. Mission volumes (Figure 1) were sometimes submitted separately in the client and then redrawn on a moving map display to monitor their aircraft conformance to the boundaries. For some flight crews this was done by a single individual on a single display or by two individuals across multiple displays.

For flight crews where a single individual managed both the client and the ground control station, interacting with the UTM system through the client became an integrated part of their pre-flight and post-flight procedures. Interactions with UTM system provided GCSOs with information about the success or failure of their volume submission; if the volume was rejected, they had access to notifications that explained why, and were able to make adjustments to their missions accordingly. When the client operator and GCSO roles were assigned to separate individuals, volume submissions and mission planning were not integrated and additional effort had to be made to coordinate the operations between the two. For example, a client operator may need to wait for available gaps in the existing GCSO regular procedures to deliver important UTM system notifications, which resulted in some take-off delays. It is based on this notion of integration between the client operations and GCSO that we applied a post hoc grouping of the ground control station configurations. A flight crew operated an integrated ground control station if a single individual occupied both the client operator and ground control station roles. In Table 1, these were flight crews with the letters "b" and "c" located within the same cell - flight crews FC7 to FC10.

All other flight crews carried non-integrated ground control stations. We conducted our analysis according to this classification.

5 Procedure

Each flight crew attended the flight test for three days. The first day consisted of a briefing and time for the crews to set up and test their equipment. The second two days were flight days when the crews flew warm-up flights and then a selection of the four scenarios that are described above. The warm-up flights served to verify connectivity between the aircraft and the ground station, and the client to the UTM system. Generally, crews flew two warm-up flights and two scenarios per flight day.

As part of the flight tests, a five-person human factors team collected data from the participants about their experiences flying in one location with multiple partners. They collected qualitative data from each group in the flight test as one researcher observed each crew. Data were collected in a number of ways, through observations of the participants during flights, brief questionnaires at the end of each scenario, and end of day group interviews. All these methods focused on five specific topics: flight crew workload during different phases of the flight, flight crew situation awareness, flight crew interaction, usage of the UTM clients and usability of the UTM system information. Participants were asked to rate their workload and situation awareness after each flight on a rating scale from 1 to 7 (very low to very high). Time permitting, they were also asked to discuss the flight they had just made. End of day interviews were focus group sessions, where flight crews discussed topics related to UTM. UTM reps also took part in a separate session and discussed the same topics. We focus our paper on the results generated from the subjective ratings collected for workload, situation awareness, flight crew communication, and how well flight crews thought the overall procedures, including the client operations, were integrated with those associated with the operation of the GCS. Analysis of the remaining human factors data are still underway and will be reported in a separate paper.

The 7 point Likert format rating scales were administered to GCSOs and PICs between each run. A run was defined by the start and stop of single test scenario; this was announced by the flight test director. Across 9 days, these ratings were collected for a total of 69 runs. These runs were not distributed evenly across flight crews due to schedule availability, unforeseen circumstances that made it unsafe to fly, or equipment failure. The distribution of runs is shown in Figure 2. Between integrated and non-integrated ground control configurations there were 26 and 43 runs respectively. Unfortunately, although interview and observation data was acquired for Flight Crew 4 (FC4), we were unable to acquire workload ratings from them for various reasons; consequently, the results reported here do not include input from FC4. Overall, the collection of ratings from the flight crews was inconsistent at best-so the total number of ratings does correspond to the number of runs.

Due to the safety risks associated with interacting with flight crews during live flight tests, human factors observers were instructed to administer data collection instru-

ments only when safe to do so and at the discretion of the flight crews. In some instances, human factors observers were not able to collect data. The analysis we show here is informal. The results we present in this paper serve only to assist with organizing and describing various flight crew configurations for informing future studies, and to highlight some potential display design challenges. We reserve formal investigations for future in-lab simulations where experimental control can be exercised.

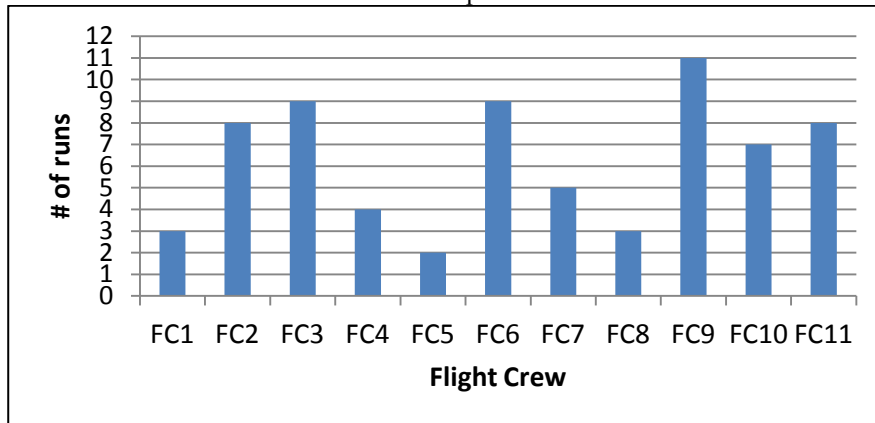


Fig. 2. Distribution of flights across flight crew.

6 Results

Sixteen workload ratings were collected for integrated GCSs versus 24 for non-integrated GCSs. Figure 3 shows that mean rated workload between the two crew-types was equally moderate. Average rated workload for integrated GCSs was 3.9 ($SD = 0.7$) and 3.6 for non-integrated ($SD = 1.5$).

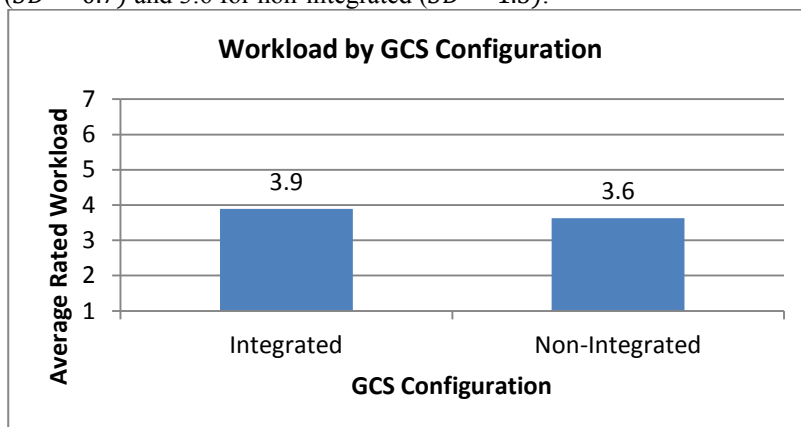


Fig. 3. Average workload rating by GCS configurations.

Overall quality of communication was reported to be very good for integrated and non-integrated GCSs (Figure 4), where the mean for integrated crews was 6.7 ($SD = 0.5$) and 6.6 for non-integrated ($SD = 0.8$). Lack of availability for ratings from

the flight crew and hardware failure resulted in some missing data here. There was one missing data point for the integrated group ($N = 15$) and 7 missing data points for the non-integrated group ($N = 17$).

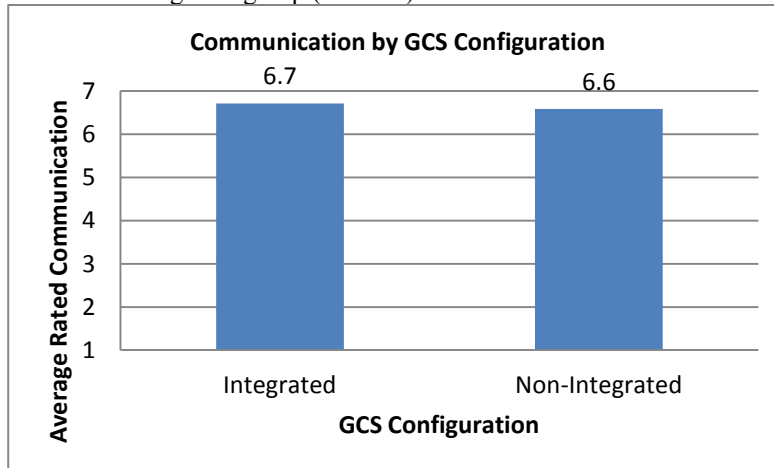


Fig. 4. Average communication rating by GCS configuration.

Consistent with overall quality of communication, on average the GCSOs indicated that there was little difficulty in relaying UTM information to the rest of the flight crew (Figure 5). This was the same for both integrated ($M = 2.3, SD = 2.4$) and non-integrated ($M = 2.4, SD = 2.2$). Again, lack of availability for ratings resulted in missing data with the integrated group ($N = 13$) and software issues inhibiting launch in the non-integrated group reduced the number of ratings significantly ($N = 9.0$).

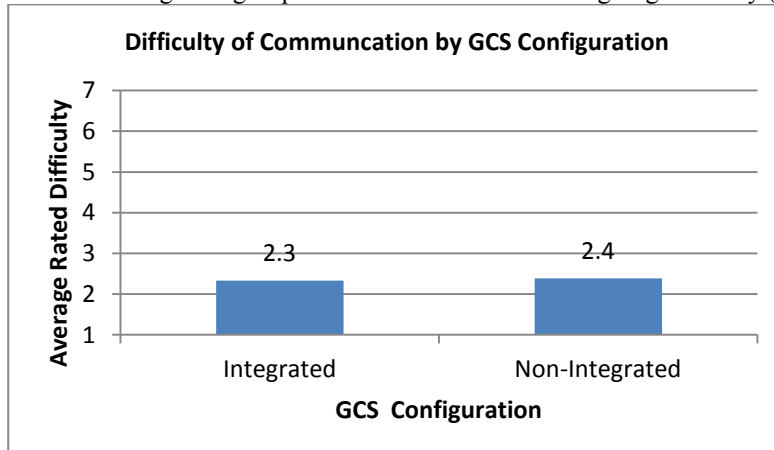


Fig. 5. Average difficulty of communication rating by GCS configuration.

We see the greatest differences between integrated and non-integrated crews with situation awareness (Figure 6). In the integrated group GCSOs felt that their situation awareness was good ($M = 5.1, SD = 2.1$). In contrast, GCSOs in the non-integrated group believed their situation awareness was mediocre ($M = 3.9, SD = 2.0$). Fourteen

ratings were acquired for integrated versus 18 for non-integrated crews. It is conceivable here, that with integrated displays, it was easier to maintain situation awareness because GCSOs did not need to divert attention from their map displays to view UTM system notifications.

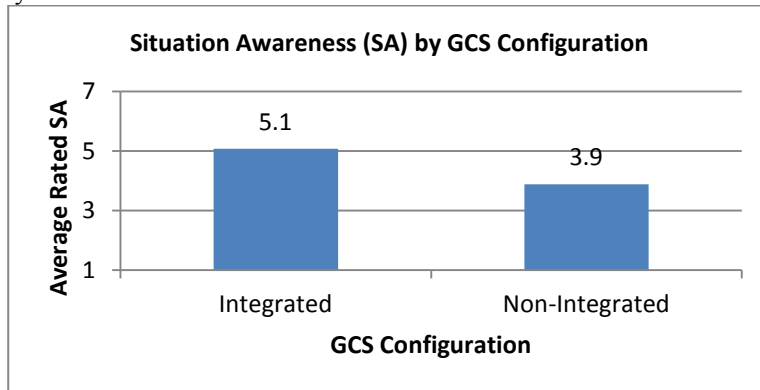


Fig. 6. Average situation awareness by GCS configuration.

Figure 7 shows ratings for how well GCSOs believed overall procedures were integrated with procedures for the ground station. The average rating for this dimension in the integrated group is 6.2 ($SD = 1.5$) and 5.4 ($SD = 1.3$) for the non-integrated group. The less positive view is not surprising. If a single operator was considering information from both the client and ground control station, it is reasonable that procedures for both the displays fall in line with a single set of procedures - even if the client and ground control were on separate, but proximal displays. Some additional reconciling of procedures would be expected if the client and ground control station operations were handled between two separate individuals. For this set of ratings we had 15 ratings for integrated and 18 for non-integrated crews.

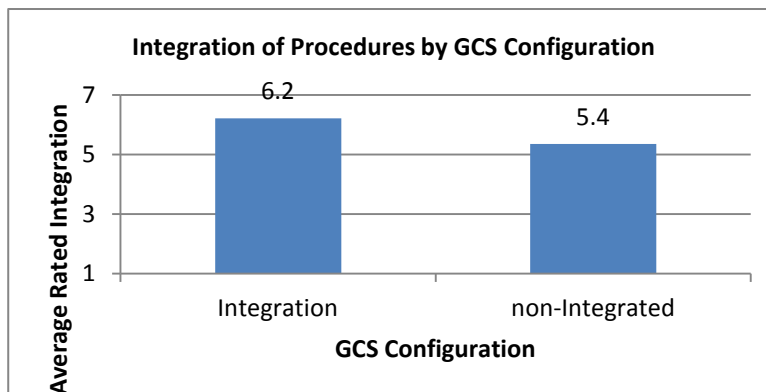


Fig. 7. Average rated integration of procedures by GCS configuration.

7 Conclusion

In this paper we offered a brief introduction to the UTM project and activities associated with the development of the UTM concept. In our results, we expressed that due to limitations in data collection opportunities and the unpredictable circumstances of a live flight test environment a considerable amount of data was missing from our explorative analysis of the UTM GCS configurations. However, TCL 2 offered an opportunity to make human factors contributions to the UTM development effort. We were able to field and evaluate early versions of the UTM ground control stations, as well as vet equally early versions of our data collection instruments and procedures. Collectively the data will inform human factors efforts in future UTM tests.

Overall, the subjective ratings revealed that with respect to workload and communication, there was little-to-no difference between the integrated and non-integrated GCS configurations. GCSOs felt workload was moderate in most circumstances. For communication, both GCS configurations were rated equally good. We speculate from field observations that when client and GCSO roles were not integrated, UTM representatives were able to compensate for any issues that germinated from that lack of integration.

Although the UTM representatives were an artifact of the flight test and did not operate regularly with the flight crew, the flight crews regarded them as essential members of the team. In almost all cases, because they had an established rapport with the UTM representatives, flight crews requested that the representatives initially assigned to them in shakedowns remain with them throughout the field tests. As we consider development of the displays and applications for UTM, it may be useful here to more closely evaluate the role of the UTM representative and how aspects of this role can be incorporated into automation used to assist interaction with the UTM system. It may also be useful, when the concept is extended in later TCLs, to evaluate the role of operators who will control large fleets - where the economy of employing a single operator will be pertinent and issues such as task switching will be a key concern to public safety and industrial organizations.

Ratings collected for situation awareness and integration of procedures seem to favor integrating the displays and combined roles for client operator and GCSO. This seems reasonable for commercial applications mentioned above, and for recreational activities where it would be economically impractical to require a multi-person crew to operate an aircraft. The principle challenge will be to identify what information the UTM system can provide for such displays without cluttering or intruding on central mission planning features.

As the development of the UTM system moves forward, the human factors effort will need to pay particular attention to assembling information requirements that will be sensitive to the privacy concerns of people and industry, but at the same time facilitate the sharing of an appropriate amount of information to support safe and effective use of the UTM airspace.

Acknowledgements

We extend our gratitude to all of the flight crews who participated in TCL 2 for their patience and support of the UTM human factors effort.

References

1. FAA: FAA Aerospace Forecasts, https://www.faa.gov/data_research/aviation/aerospace_forecasts/.
2. Unmanned Aircraft Systems (UAS) Service Demand 2015-2035 Literature Review & Projections of Future Usage. , Cambridge, MA (2013).
3. Kopardekar, P., Rios, J., Prevot, T., Johnson, M., Jung, J., Robinson, J.E.I.: Unmanned aircraft system traffic management (utm) concept of operations. Am. Inst. Aeronaut. Astronaut. (2016).
4. Prevot, T., Mercer, J., Martin, L., Homola, J., Cabral, C.D., Brasil, C.L.: Evaluation of High Density Air Traffic Operations with Automation for Separation Assurance , Weather Avoidance and Schedule Conformance. 11th AIAA Aviat. Technol. Integr. Oper. Conf. Including AIA. (2011).
5. Prevot, T., Homola, J., Mercer, J.: From Rural to Urban Environments: Human/Systems Simulation Research for Low Altitude UAS Traffic Management (UTM). 16th AIAA Aviat. Technol. Integr. Oper. Conf. Washingt. D.C. 13-17 June. 1–13 (2016).
6. Homola, J., Prevot, T., Mercer, J., Bienert, N., Gabriel, C.: UAS Traffic Management (UTM) Simulation Capabilities and Laboratory Environment. 1–7 (2016).
7. Johnson, M., Jung, J., Rios, J., Mercer, J., Homola, J., Prevot, T., Mulfinger, D., Kopardekar, P.: Flight Test Evaluation of a Traffic Management Concept for Unmanned Aircraft Systems in a Rural Environment. In: Twelfth USA/Europe Air Traffic Management Research and Development Seminar (ATM2017) (2017).