Developing and testing two interfaces for Supplemental Data Service Provider (SDSP) tools to support UAS Traffic Management (UTM)

Jolene Feldman[1] and Lynne Martin[2]
NASA Ames Research Center, Moffett Field, CA 94035

Vimmy Gujral[3]
San José State University, NASA Ames Research Center, Moffett Field, CA 94035

Charles Walter[4]
ASRC Federal Data Solutions, NASA Ames Research Center, Moffett Field, CA 94035

Dorrit Billman[5]
NASA Ames Research Center, Moffett Field, CA 94035

Patricia Revolinsky[6]
NASA Langley Research Center, Hampton, VA 23681

Gregory Costedoat[7]
San José State University, NASA Ames Research Center, Moffett Field, CA 94035

Researchers conducted a usability study using two graphical user interfaces (GUIs) to explore how individuals interpret and interact with different preflight information displays, and to inform the development of Uncrewed Aircraft System (UAS) preflight planning predictive support tools to assess and mitigate flight hazards and risks. A series of preflight risk-assessment tasks were developed to evaluate participant performance using the Supplemental Data Service Provider-Consolidated Dashboard (SDSP-CD) and the Human Automation Team Interface System (HATIS) GUIs. Participants were trained to use both interfaces and their performance was evaluated. These evaluations focused on participants' preflight planning activities. Objective data on performance tasks across different scenarios involving multi-UASs, as well as self-reports of interactions and subjective experiences using

[1,2] Research Psychologist, Human Systems Integration Division.
[3] Researcher, Human Systems Integration Division.
[4] Senior Software Engineer, Human Systems Integration Division.
[5] Research Scientist, Human/Machine Systems.
[6] Aerospace Engineer, Aeronautics System Analysis Branch, AIAA Member.
[7] Human Factors Research Assistant, Human Systems Integration Division.

the GUIs were collected. Scores on the system usability scale (SUS) and on a simple task set were examined, as well as user feedback on open-ended questions, to inform development and identify potential improvements to the interfaces.

I. Introduction

The Uncrewed Aircraft Systems Traffic Management (UTM) concept aims to enable safe and efficient Uncrewed Aircraft System (UAS) operations in low-altitude uncontrolled airspace by providing services such as airspace configuration, flight planning, and flight monitoring. UTM will be capable of enabling low- and high-density operations of small (55 lbs. or lighter) uncrewed aircraft systems (sUAS) in urban, suburban, and rural environments. Uncrewed aircraft system operators (UASOs) will need easy access to important information, to build situation awareness and maintain safety in the airspace they are flying.

One aim of NASA's System-Wide Safety (SWS) project is to provide a suite of data-driven tools that fuse and analyze data collected from aviation systems, and other data sources, to predict hazards, and then mitigate the risks by allowing changes to plans, procedures, or designs [1-3]. Some of this information is being gathered in models that can show the user how these parameters will affect their upcoming flight. These models will become System-Wide Safety Services, Functions, and Capabilities (SFCs). These SFCs can assist operators to assess and test mitigations to these risks during preflight. Six SFCs that may expose flight risks are currently under development. These are battery reserve and end of discharge, global positioning system (GPS) strength, proximity to ground obstacles, ground risk assessment, radio frequency spectrum interference, and wind patterns. Supplemental Data Service Providers (SDSPs) provide the means through which these risk and hazard services can be run and displayed on a graphical user interface (GUI) as they are related to the vehicle's route of flight.

As part of the SFC development, the SWS project is investigating the interface requirements to display and organize this set of SFCs and report how operators might use an interface, focused on whether the GUIs provide information for situation awareness that UASO and other operators (e.g., Fleet Managers) will need before and during their flight operations to enable an acceptable level of In-time System-wide Safety Assurance (ISSA). These interfaces can provide a comprehensive overview of the potential hazards associated with a particular flight(s), which allow operators to determine whether a particular flight path is acceptable or not (requiring re-planning). The purpose of this research is to build and test interfaces for NASA SDSP prototypes with the same function that will contribute to ISSA and could offer additional information to UTM service suppliers and UASOs. The Supplemental Data Service Provider-Consolidated Dashboard (SDSP-CD) developed at NASA Ames Research Center and the Human Automation Team Interface System (HATIS) developed by the Human Automation Teaming Solution (H.A.T.S.) company are candidate GUIs that display the results of predictive SFC tools to the operator.

II. Background

The design and usability of a display affect the information operators attend to and their understanding of it in relation to operations [4]. Hazard data and risk predictions are informative but can be complex to read and understand. However, presented visually and in relation to flight parameters (such as flight path), the nature and significance of hazards become much more evident. Degani and colleagues [5] assert that, "For information to be useful beyond being a set of individual pieces of information, it must be integrated. Such integrated information can help enhance user understanding of relations among pairs, triplets, and even larger sets of tuples, thereby enhancing overall situation awareness". The SDSP-CD and the HATIS interfaces offer both a means to present hazard model output and offer an opportunity to test user understanding of the information, user decision making and the best ways to present such data to an operator. Decision-making for flight planning and in particular, assessment of risk is affected by situation awareness. This awareness is built from the elements that are perceived in the environment, how this information is comprehended, and the projection of the status of these elements in the near future [6]. Presenting salient and meaningful risk assessment information to operators is necessary to increase situation awareness and ultimately safety.

The goal of the present study was to determine whether these interfaces provide enough information for users to understand each GUIs capabilities, provide enough information to complete various risk and hazard assessment tasks successfully in simulated scenarios, and to also evaluate users' subjective experiences with each interface.

III. Usability Study

A. Method

This usability study was conducted on-site in the Airspace Operations Lab (AOL) at NASA Ames Research Center. Participants' performance and feedback were evaluated using two different GUIs to display SDSPs for hazard and risk assessment and mitigation in preflight planning UAS routes in different scenarios. The software for the SDSP-CD GUI was run on-site through NASA Ames servers. The software for the HATIS GUI was accessed remotely on Microsoft Teams with NASA Langley Research Center and run through servers located at NASA Johnson Space Center.

B. Participants

Sixteen participants completed the study. Eight participants were in-house colleagues (i.e., contractors or NASA civil servants) with subject matter expertise in the discipline of human factors (HF). The remaining eight participants were external recruits who possessed some level of UAS flying experience (e.g., recreational, military operations). In addition, these participants had a range of flight experience, including holding fixed-wing or instructor certifications while some held Part 107 certifications. All participants were naïve to the research purpose and procedure prior to the study. External recruits received financial compensation for participating, while in-house recruits volunteered to participate during their normal duty hours. This research was approved by NASA's Human Institutional Review Board, and informed consent was obtained from all participants.

C. Graphical User Interfaces

The SDSP-CD is a prototype display which shows the results of five services developed within SWS: (1) Battery reserve, or end of discharge, (2) Proximity hazards (Proximity to Threat, PtT), a calculation of proximity to ground obstacles – specifically buildings and trees, (3) Population hazard or human casualty risk assessment based on the local population density and the risk of UAS motor failure, (4) GPS coverage, an identification of GPS satellite availability, and (5) Radio Frequency Interference (RFI), a risk of signal disruption (Figure 1). For more details about the construction and development of these services, see [7-14]. Previous work evaluating users' feedback on the GUI elements of the SDSP-CD was completed and resulted in improvements to the interface [15] that was used in this study.

The HATIS interface also provides preflight planning risk assessment support. The HATIS currently displays the results of similar services as the SDSP-CD interface: (1) GPS/Navigation quality Minimum, (2) Battery State of Charge Minimum, (3) Population hazards, or the average Probability of casualty along a flight path (PcAve), as well as a maximum threshold exceedance at any one point along a flight path (PcMax), (4) Radio Frequency Interference Average, and Radio Frequency Maximum, and (5) Proximity to Threat minimum (Figure 2).

D. Scenarios

A total of four scenarios were developed. Two scenarios were developed to train participants to use each interface and two additional scenarios were developed to assess performance, after training, on various tasks using the interfaces. Scenarios for nominal operations, as opposed to emergency or critical response missions, were the focus. The location and mapped area for the training scenarios were the same, i.e., NASA Langley Research Center's City Environment for Range Testing of Autonomous Integrated Navigation (CERTAIN) area 1. The mapped task scenario area was different from the training scenarios but was the same for both interfaces, i.e., Langley Research Center's CERTAIN area 2. This prevented participants from seeing the same mapped areas in both the training and task scenarios. It also allowed for analogous conditions, so the tools could present comparable problems and conflicts with respect to preplanned flights between buildings, trees, and/or simulated people on the ground, etc.

The training scenario for the SDSP-CD involved three simulated time shifted, deconflicted sUAS flights (Figure 3). The missions included: Recon_1 as a perimeter patrol flight at an altitude of 40 meters, Recon_2 as a pipeline inspection flying at 100 meters, and Recon_3 as a search and rescue mission also flying at 100 meters. The training scenario for the HATIS included four simulated time shifted, deconflicted sUAS flights (Figure 4). The missions included: UAV401 as an aerial photography flight, UAV402 as an equipment transport flight, UAV403 as a pollution monitor and control flight, and UAV404 as a disease tracking and tree health monitoring flight.

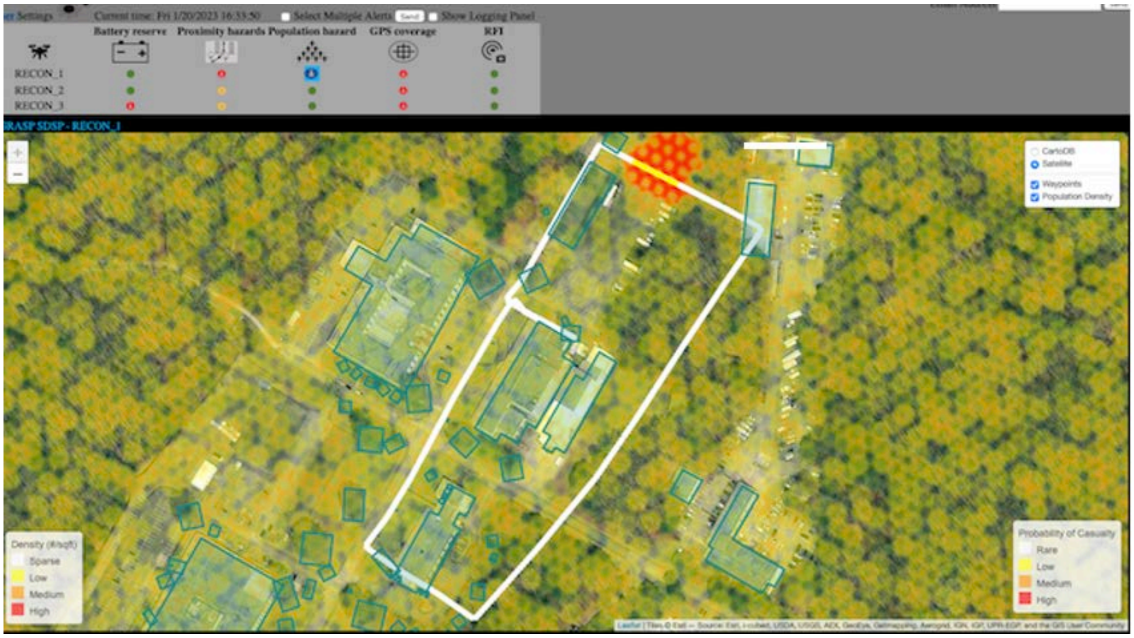

Fig. 1 Image of SDSP-CD display interface with Population hazard alert selected for the UAS with the callsign "RECON_1". Population density and probability of casualty heat map shown. (NASA image)

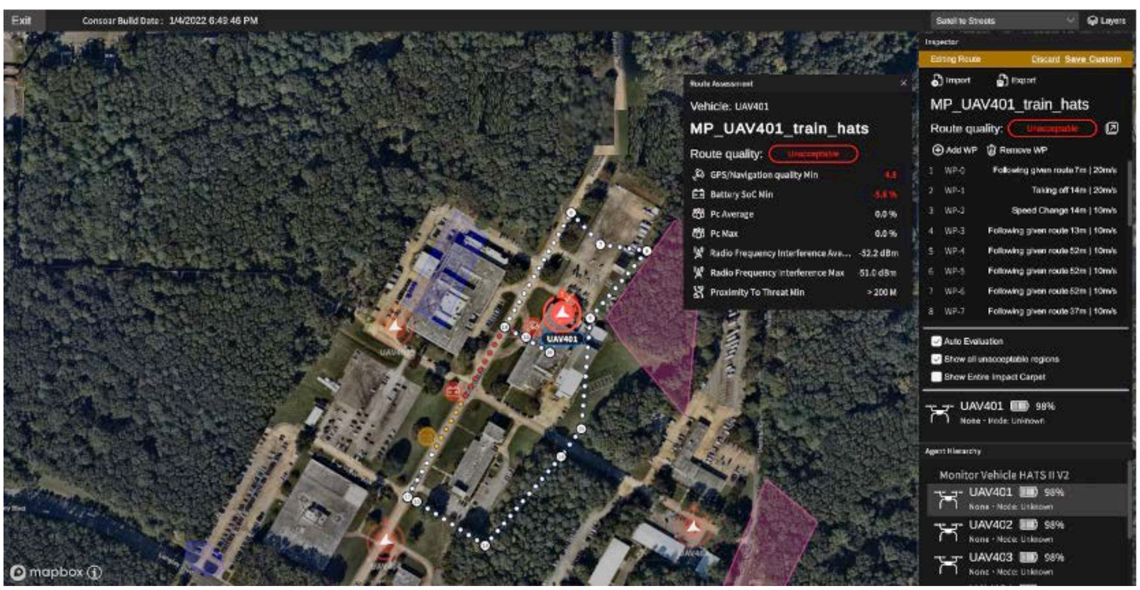

Fig. 2 Image of HATIS display interface with route assessment selected for a UAS with the callsign "UAV401". (NASA image)

Two scenarios were developed to assess performance. The Wildlife Management simulated missions included: UAV1 taking video of bird nesting flying at an altitude of 30 meters, UAV2 taking samples to a lab flying at an altitude of 35 meters, UAV3 tracking movement of foxes flying at an altitude of 35 meters, and UAV4 patrolling the whole area to look out for people and/or ground vehicles entering the restricted zone flying at an altitude of 100 meters (Figure 5). There were flight constraints associated with the missions (e.g., UAV1 had to fly below 50 meters to complete its mission).

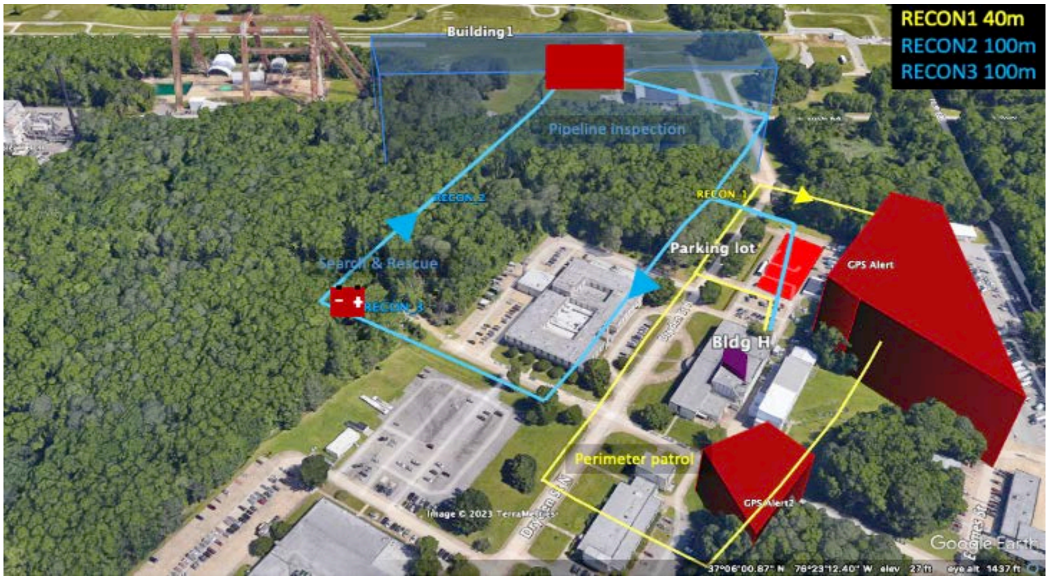

Fig. 3 Training scenario for the SDSP-CD. (Maps Data: Google, Image ©2023 TerraMetrics)

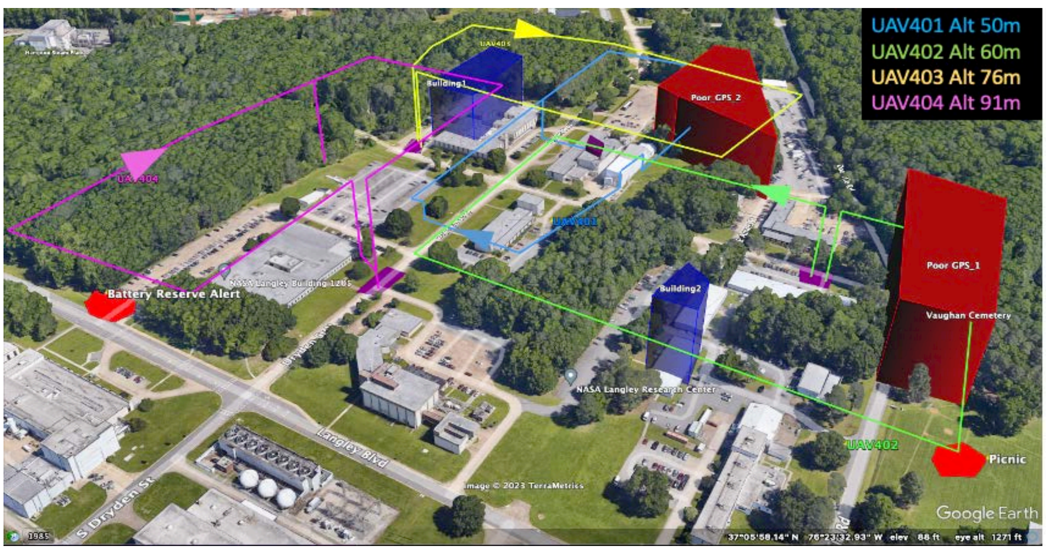

Fig. 4 Training scenario for the HATIS. (Maps Data: Google, Image ©2023 TerraMetrics)

The second scenario was a Package Delivery Mission (Figure 6). The simulated missions included: UAV1 and UAV2 taking samples to, as well as collecting samples from, specific buildings, with altitudes varying from 30-103 meters along their flight paths, UAV3 delivering lunches and snacks to specific buildings, with its altitude varying from 10-30 meters, and UAV4 picking up and dropping off mail from specific buildings, with altitudes varying from 10-30 meters along its path. Again, there were constraints associated with these missions (e.g., not fly above 122 meters and, due to the amount of UAS traffic, a one-way system was in place for UAV 1 and UAV2).

E. Procedures, tasks, and surveys

Participants were scheduled individually for an in-person 5-hour session. One interface system was trained at a time. There were two parts to the session for each interface, a scenario for training followed by a scenario for the task missions. Participants completed interactive training for both interfaces during the session, with either the SDSP-CD interface presented first or the HATIS interface presented first. These conditions and the task scenarios were counterbalanced into four conditions – with either the Wildlife Management Mission or the Package Delivery Mission being presented first, for either the SDSP-CD or the HATIS interface. The presentation of the scenarios and systems were counterbalanced into 4 conditions. Of the sixteen participants, two external participants (e.g., fixed wing pilots) and two internal participants (i.e., human factor participants) were assigned to each of the four conditions (Table 1).

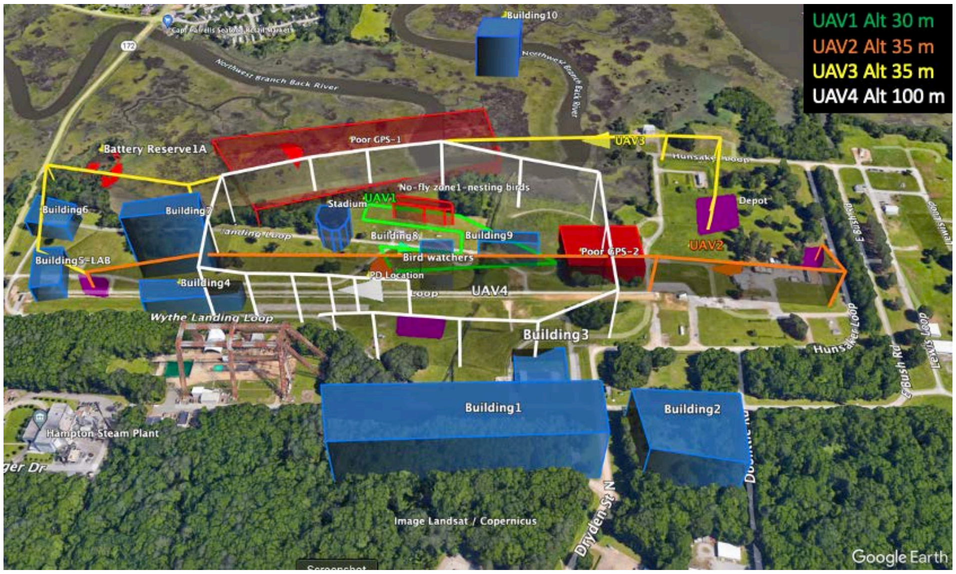

Fig. 5 Wildlife Management Scenario. (Maps Data: Google, Image ©2023 Landsat / Copernicus)

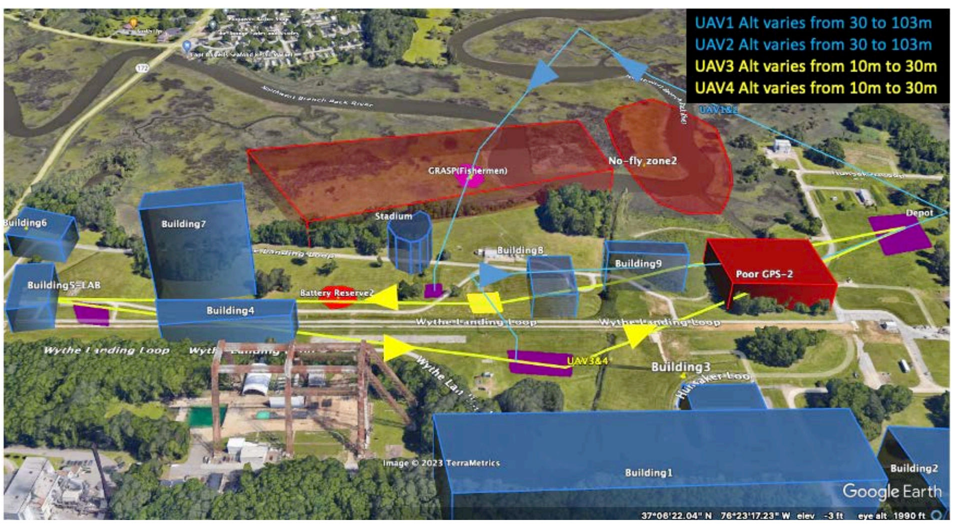

Fig. 6 Package Delivery Scenario. (Maps Data: Google, Image ©2023 TerraMetrics)

Training was presented through a Power Point slide deck and a brief 5-min pre-recorded overview video, where a researcher provided information on the features and capabilities of the interfaces. Training was interspersed with check for comprehension questions, hands on interaction with the interfaces, and opportunities for questions or review. The advantage of participants interacting with both interfaces was that it was an opportunity for them to directly compare their experience with each interface and the ways each GUI enabled them to complete their missions. Participants were trained to criterion; all activities during training provided practice for what participants would be expected to complete during the tasks.

Table 1 Table of the presentation order of the scenarios and systems with participant assignment

Presentation Orders: Scenarios ⟹ Systems ⇓	Wildlife first, Package second	Package first, Wildlife second
SDSP-CD first, HATIS second	4 participants*	4 participants*
HATIS first, SDSP-CD second	4 participants*	4 participants*

Note: * indicates 2 participants with the focus on HF expertise and 2 participants with the focus on UAS experience

Verbal responses and data entries by participants from a mix of tasks, prompted what-if questions, and questions about risks, paper maps drawings, responses to paper and pencil workload and usability surveys, and feedback on open-ended questions were collected. The NASA Task Load Index (TLX) was administered after two distinct, but contiguous blocks of tasks related to specific aspects of each GUI. The TLX is a brief survey that assesses subjective workload over multiple dimensions, resulting in an overall workload score [16]. The System Usability Scale (SUS) was administered at the completion of all tasks for each GUI. The SUS is a 10-item survey, with a five-point Likert scale (“Strongly Disagree to “Strongly Agree) that provides a high-level subjective general assessment of usability [17]. Some capabilities of the GUIs were demonstrated during the trainings (e.g., a feature to add or delete waypoints in HATIS, or the ability to change SDSP-CD threshold settings) but were not necessarily required to complete the tasks scenarios. Feedback on these features is included in the open-ended comments. Participants response times on all tasks were also recorded. All sessions were recorded through Teams with audio and screen recordings, but no video of the participants was captured.

More discussion about the TLX scores and additional analyses and results, as well as data from the additional tasks (i.e., re-routing flight paths on paper and the GUI) will be the subject of other papers. The focus for this paper is directed at participants’ ratings on the SUS, performance on a simple task set, and feedback on some of the open-ended questions as they relate to the tasks.

F. Coding methodology

Each of the participants’ responses were recorded and timestamped. Researchers developed a rubric for coding a participant’s responses as either “correct” (1 point) or “incorrect” (0 points), and a percentage correct was calculated. Two researchers identified, coded, and concurred on the criteria for a correct response for each question. When concurrence was not reached an additional researcher was consulted. Some items were dropped from scoring either due to unclear or poor wording of the question, or participants were not able to access information related to the question from the GUI due to unexpected software updates (e.g., GPS for HATIS was not available during some sessions). Subjective responses to more qualitative questions (e.g., Which alerts would you solve first and why? With the existing battery alert for UAV3 and UAV4, what options do you have and what would you decide to do?) – where there was more than one answer or way to solve a situation were reviewed. Additional comments and feedback related to some of the tasks are included below.

In the SUS, five of the items are positively worded and five are negatively worded. To score the SUS, the negatively worded item ratings are reversed and then the sum of the ten ratings is converted to a score out of 100. A higher score indicates the respondent found the interface more acceptable or user-friendly and a lower score indicates the respondent found the interface hard to use.

IV. Results

For SUS responses about participants’ experience using these interfaces, a score (not a percentage) above 69 indicates that the interface is rated as a “good” user-interface. The mean of participants’ scores on the SUS were above this level (SDSP-CD = 72.3, HATIS = 74.6) for both interfaces. Most participants rated the two displays similarly and positively, with participants A and P among the few showing big differences between their scores for the two displays, and participant E who rated both displays poorly (Figure 7).

Responses to open-ended questions confirm the SUS ratings that participants thought both interfaces were useful for preflight planning, particularly in illustrating where hazards and possible risks were along a flight path. When comparing the interfaces, participants primarily wanted flight path route editing capabilities and altitude information for the SDSP-CD and, for the HATIS, to be able to view all the flights and information in one at-a-glance view, as well the implementation of heat map features (that show the distribution of the hazard types over the local area). Participants' responses provide more detailed explanations about the items assessed by the SUS, including aspects of system usability, such as the need for support, training, and complexity.

Participants explained that the correspondence between the SDSP-CD dashboard (top grey table in Figure 1) and the location along the flight path trajectories on the map of the same information (lower area in Figure 1), such as the types of hazards (e.g., PtT, battery reserve) and values associated with specific segments of the flight path, were helpful. The ability to click on an icon of a specific service on the dashboard and locate where the problem was, the filtering feature, selecting multiple alerts (viewing or hiding certain services and/or vehicles), as well as the heat map (e.g., population hazard, shown in Figure 1) were also helpful. Some of the benefits of using this interface were being able to see and manage multiple vehicles together, which would support multiple flight efficiency (i.e., constructing more efficient and safer routes that are close together). Other useful features were having information about predicted obstacles based on user defined thresholds, giving the ability to detect, avoid, and remove risks – such as, reduce the likelihood of overextending battery or violating regulations (e.g., flying over people, Federal Aviation Regulations, FARs).

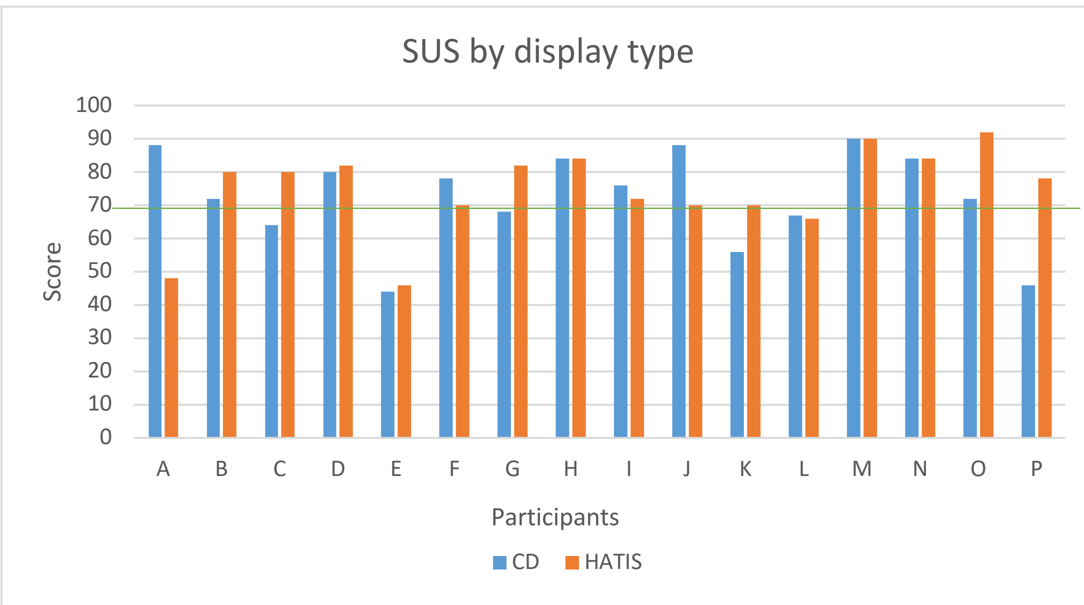

Fig. 7 Participants' SUS score by display type

After using the HATIS, participants reported that it was useful to have the ability to rubber-band the routes and have the interface re-evaluate the risks. Further, it was also reported that information was easy to read, with the overlay on a real map, and that it looked aesthetically polished. The route assessment or "pop out" box (Figure 2) was helpful, and participants liked having multiple units (vehicles) available in the same interface, being able to see the current aircraft's position and route and distinguish the levels of alerts. It was helpful that the flight paths changed color based on hazard threshold (e.g., battery from yellow to red along a salient waypoint sequence) to clearly show exceedances, the waypoints were numbered and could be moved, the active evaluation of the route updated as the route was changed, and that good information was provided (e.g., callsigns to ID aircraft, starting battery capacity/percentages). There were many benefits reported by participants when using the HATIS interface. It was beneficial being able to have a plan (and better planning) and manage multiple flights, seeing existing issues, and flying something deemed less risky (i.e., automatically assesses threat, health of each vehicle along its flight provided – e.g., battery life). Users liked the route analysis – "unacceptable", "acceptable", or "good". Being able to draw out a path around what needed to be avoided made it less demanding. One participant described how if flying as a remote pilot, the ability for operators to be able to see water below (e.g., creeks, lakes) or if a vehicle loses GPS signal and needs to return to its starting point is also very beneficial.

Conversely, there were some concerns expressed about both interfaces. Lacking information about altitude, take-off and landing points, and some confusion about what was shown on the map – "having to click in multiple places and then having to analyze it", were some of the challenges participants reported using the SDSP-CD. Some participants reported "too much clicking" in navigating the HATIS interface, which was a source of some frustration. This frustration was reflected by participant E in their low SUS score (Figure 7). It should be noted that, researchers had selected the option box for "auto-evaluation" of the flight paths in HATIS, which had already reduced the number of clicks required for route assessments. There was also some concern about hazard icons on the HATIS map being hidden by being layered one-on-top of the other (icons should be smaller or more transparent). Furthermore, participants wanted guidance on where to adjust a flight path and how to resolve the problem, as they were "guessing, incrementing" [the route] until the route was evaluated as "acceptable". The ability to adjust flights through HATIS was useful; however, the optimization of a route for one vehicle may cause a problem for another vehicle. This can become a huge problem if an operator is managing a large fleet of vehicles. More information was needed for ways to "fix the problem" and to prevent guesswork on adjusting unacceptable flight paths. One user stated that it would be helpful if the HATIS could provide a few suggested routes that an operator could choose between. The ability to adjust a flight path was viewed positively, but the challenge was knowing where to adjust the path (e.g., a user was randomly rubber-banding a flight path until eventually all hazards were cleared).

Participants provided numerous ideas, suggestions, and potential improvements about the information shown on the displays. They wanted more information depicted on the map (especially altitude, labeling no-fly zones, or restricted airspaces, building numbers and easy visibility of these, vertiports, direction of flight), details on population density hazards (e.g., different risk levels when people are in ground vehicles vs. out in the open), and estimates of the UAS battery range in both time and distance. They also wanted additional services to show weather and weather changes, terrain (or topographic data), and charging stations. Participants also asked for new features including the ability to overlay flight paths of multiple vehicles, the ability to rotate the map and see it in a 3-D view, and indication of emergency options. Other suggestions of possible features included: more details on why a service threshold was poor (e.g., reason for poor GPS coverage – as this could affect how they adjust the flight path), and knowing the recency of the data, and features like what are included on a visual flight rules (VFR) sectional chart.

There were questions and concerns about the reliability of the services for both displays. There were some questions about trusting the data and whether information will still be accurate when a vehicle is ready to fly. A user explained the need for operators to trust the data from the services, and how a lack of trust in the data could affect (i.e., increase) workload. Alternatively, a perception of fear of this being a "one-stop shop" was described where there is no need to look at other sources of information for confirmation or validation, and ultimately not wanting the system to be viewed as faultless. The threshold setting feature on the SDSP-CD was not used in this study and, unaware of this, participants wondered not only whether experts provided the correct data (e.g., thresholds), remarking that essentially the algorithm would determine the problems, but this assumes that the services are always available and that everything is reliable and online. Some participants commented that the anticipated data might not match reality, may reduce awareness (e.g., not able to see multiple operations at once, lack of awareness outside of their own operation), and perhaps give a false sense of security (e.g., GPS may be N/A). As an operator, participants asserted that they are trusting the information that they are being given (hoping it is accurate) and wondering if they fully understand what they are seeing on the display, and the extent to which they will make the proper adjustments. Participants also wanted to understand how well this interface(s) would integrate with other systems. One user described how if systems integrate well, that is helpful; but not if this is work that needs to be replicated.

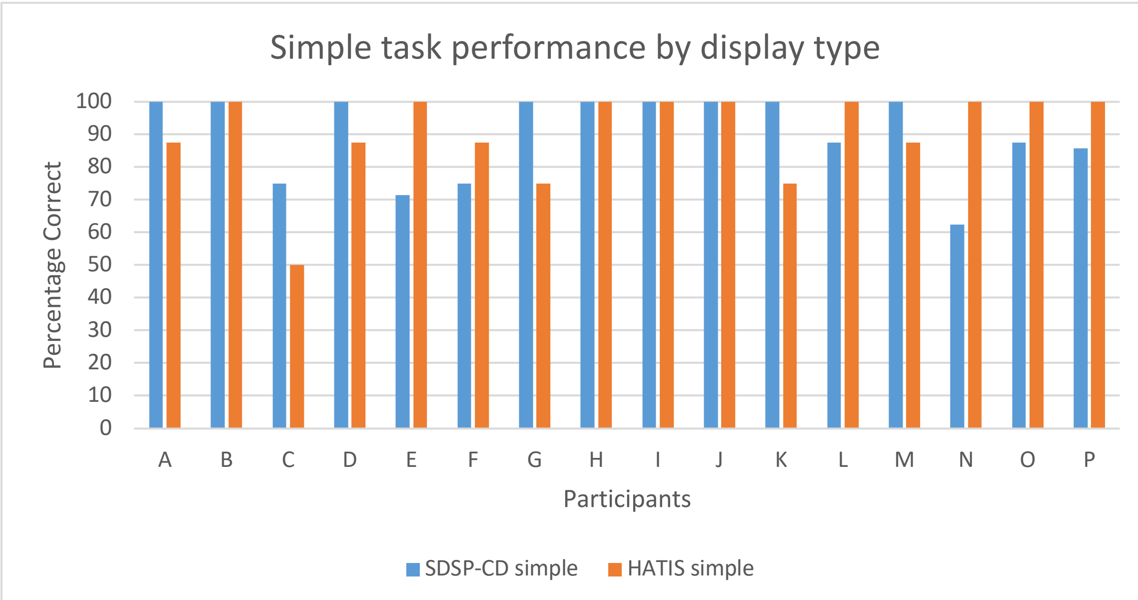

Fig. 8 Participants' scores on simple task set by display type

Participants were generally able to identify risks and hazards. The simple task set asked straightforward questions about what the participant saw on the interface. Some of these questions are listed in Figure 9. As a group, the mean percentage correct on the simple task set using the HATIS was 93%, and for the SDSP-CD it was 91%. While participants A and P showed big differences on the SUS scores between displays, both participants scored above 80% on the simple task set using both displays (Figure 8). Participant E, who rated the usability of both displays poorly, still reached a score of 70% on the simple task set using the SDSP-CD and 100% using the HATIS.

If we look at some specific participants, we see that participants C, G, and K got more tasks correct using the SDSP-CD than the HATIS, and that participants E and N got more tasks correct using the HATIS than the SDSP-CD. Some of these differences may be related to the concerns previously described about each interface, but also by looking more closely at the specific simple task set questions (Figure 9). These tasks may have been harder because of certain aspects of the displays.

When looking at responses from the Wildlife Management scenario more directly, question by question, there were more correct responses to question 1 when using the HATIS rather than the SDSP-CD (Figure 9). From the open-ended feedback, participants described how the route assessment button in HATIS, that evaluates the flight path and shows a "pop out" box with more detailed information, was helpful in distinguishing the levels of alerts. The flight paths on the map change color based on hazard threshold. For example, the battery path will turn from yellow to red along the sequence of waypoints to clearly show at what point the vehicle's battery has exceeded the accepted threshold. The route evaluation feature will indicate whether a specific flight path is "unacceptable" in red, "acceptable" in yellow, or "good" in the color white. Participants reported that the yellow to red color transition of the route provided beneficial information about the boundaries of operation and could reduce workload in deconflicting plans. The distinction of what yellow means, in this case "acceptable", is determined by the software – this was re-emphasized in the pop out box to the participants every time a flight path was evaluated.

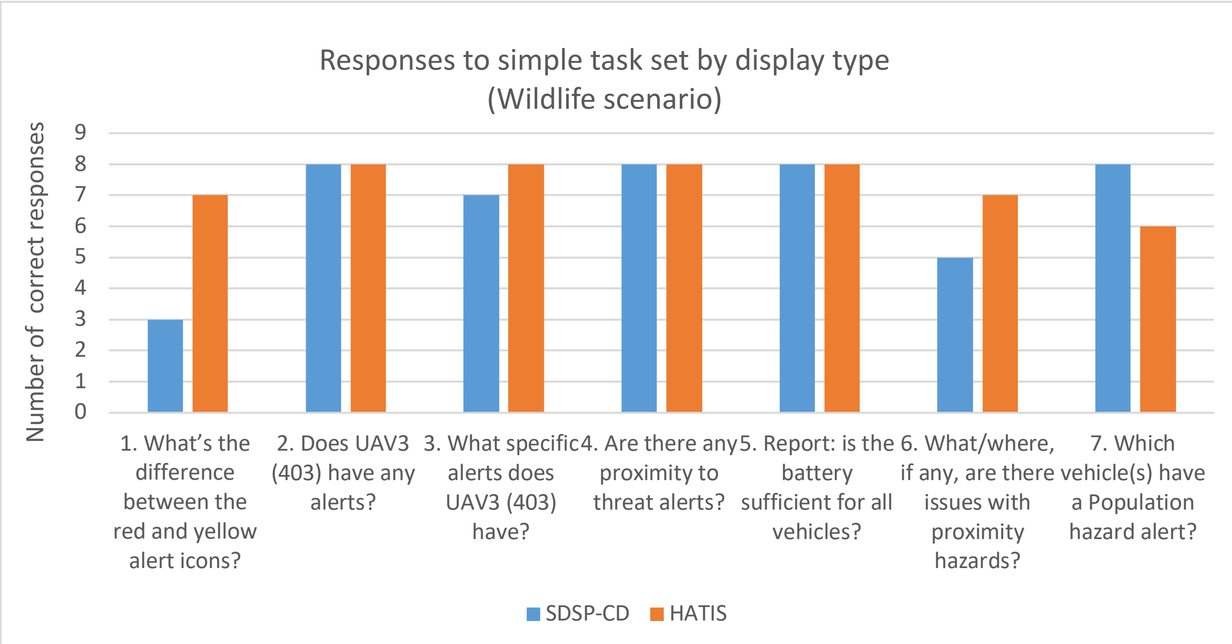

Fig. 9 Responses about the Wildlife Management Mission scenario by display type

In contrast, for the SDSP-CD, the panel displays the alerts and their associated colors, but without an auto-evaluation feature. Whether or not a route is acceptable, is left to the judgment of the operator. The interface helps with identifying the hazards but does not automatically determine whether the routes are acceptable or not, so this would depend on the operator's (i.e., fleet manager) threshold of risk. If an operator determined a flight path needed to be adjusted, they would have to resubmit another flight plan into the system. During SDSP-CD training, yellow alerts on the dashboard were defined as thresholds being exceeded, and so would not be considered an acceptable flight path but require re-planning. In addition, on the SDSP-CD map, yellow had other meanings (e.g., low density, low casualty, "ok" satellite signal, poor RFI). Participants commented on the inconsistencies of the legend information for each hazard type – for some the font was too small, the location of the legend box changed (lower right or lower left of the screen), and the levels of risk descriptions (difference between "good" vs. "ok") was unclear. These inconsistencies may have also contributed to the participants' unclear understanding about the meaning (i.e., severity) of a yellow alert and contributed to the lower relative score on the SDSP-CD for this question (Figure 9, question 1).

One possible explanation for why there was a difference in question 3 response means (Figure 9), to HATIS' advantage, is that when a vehicle is selected on the HATIS, information about each of the services, thresholds, and level of alert are indicated for that specific vehicle in the "pop out" box, while the remaining vehicles remain in a queue list on the panel adjacent to the map. For the SDSP-CD, users view all vehicles, services and associated alerts and severity by color on the dashboard, above the map. Participants have a top-level view of all risks and hazards to all vehicles. However, there was repeated concern from participants that when they clicked on an alert on the dashboard, the original color of the alert was masked (i.e., highlighted in blue). The original alert color (either red, yellow, or green), becomes highlighted blue on the dashboard when selected, as information for that specific vehicle and service displays on the map at the same time. This was an issue for tasks involving identifying alert severity. Participants wanted to be able to select alerts but still have the alert severity status clearly visible on the dashboard.

For questions 6 and 7, responses were nearly opposite of each other. Participants scored higher on identifying the proximity hazards and where they were using the HATIS, while scores were higher in identifying the population hazards using the SDSP-CD. There are advantageous features of each interface which may explain these differences. Question 6 asked for specifics about the location of a particular hazard. The map on the HATIS shows the flight path by waypoints and identifies where specifically along a route that a hazard exists – this could be detected more easily and determined by, or between, waypoints. Participants' comments described how HATIS allows knowing that something is wrong and where it starts to go wrong and provides flight-specific information.

Question 7 asked about more general information (not specifics, like where was the hazard). Many participants described that, acting as a fleet manager, the SDSP-CD dashboard allowed for an at-a-glance overview and allowed them to quickly ascertain problems (e.g., health of the fleet) and alerted mitigation of them. This comprehensive overview of vehicles, alerts, and status may have contributed to these results. This can also explain why during the Package Delivery mission, that all participants who were using the SDSP-CD to determine the alerts for UAV2 correctly responded to this question, but only half of the participants using the HATIS got this correct. Another feature of the SDSP-CD is a drill down icon to get more information about the population from a heat map, which provides more details about the level of risk based on the population density (which ranges from sparse to high), and the probability of casualty (rare to high). Participants commented on how they liked the heat map on the SDSP-CD and wanted to also have a heat map available on the HATIS.

When using HATIS, as compared to other interfaces, many participants reported they liked the ability to rubber-band or modify the trajectories which could provide a better sense of how to mitigate an alert, and more confidence that route issues could be solved more efficiently. However, this could also require a lot of trial and error to make successful changes. When attempting to quickly obtain information about the alerts, one user reported that there were more mouse clicks required with the HATIS than the SDSP-CD interface, when trying to get the same information. This became evident on one of the tasks where participants were asked to quickly determine the total number of alerts showing values over thresholds. Performance on this task varied for two reasons: masking and clicking. First, participants using the SDSP-CD could respond to this question quickly with one overall view; but as described previously, alerts could be masked by color which resulted in incorrect total counts – at least one of the red or yellow icons on the dashboard would often be selected and result in that alert being blue instead and not identified as an alert on the dashboard. Second, to get this information on the HATIS, a minimum of six mouse clicks are required (versus zero clicks for the SDSP-CD). Participants needed to click through each vehicle, identify the alert, remember the information, and continue through the remaining vehicles. This could be done accurately but took more time to complete.

V. Discussion

Based on user performance and feedback, consideration should be given to displays that provide data on all aircraft, hazards, and status in one easily comprehensible view. This information would consist of meaningful icons and consistent, well-defined modes that help support the goals of the mission and promote safety. Displays that enable quick access to accurate and reliable information, minimize potential user confusion, prevent inadvertently masked data, and facilitate reasonable workload levels should be considered. Displays should not only quickly and simply inform operators of the risks and hazards of a route but, particularly with dynamic displays, also provide operators with some level of guidance for flight re-planning. This may be in the form of heat maps for services or suggestions for alternate routes. Ideally, displays should integrate seamlessly into existing systems and account for the types of duties and tasks that will be expected of operators in real world operations.

These interfaces allow operators to know if a risk is present, where the risk is, but not what to do or where to go, which would be helpful for pre-flight planning. Like findings from Buck and colleagues [18], with the Remote Operations for Autonomous Missions (ROAM) UAS displays, developers need to consider displays that not only communicate that a problem exists, but also "how to alleviate a problem." This was similarly reported when participants used the SDSP-CD and wanted the "dynamics" of the HATIS – with its capability of rubber-banding the flight paths. However, users emphasized the need for a capability in solving "where to go" for flight planning (e.g., providing a radius map in HATIS like the heat map on the SDSP-CD). A user may be able to adjust a flight path by rubber-banding and solve, for example, a Probability of casualty (Pc) Average alert, but may still have a Pc Max alert to resolve. Additionally, an operator may adjust a flight path only to continue to drag the flight path to other veiled hazards again and again, which can result in inefficiency, increased workload, and potential frustration.

The importance of distinguishing between different levels and meanings of alerts and threshold settings among GUIs should also be considered. There are inherent differences in the interpretation of alerts between the interfaces. For the HATIS interface, when the display panel or pop out route evaluation assessment displays yellow, it means that the hazard (of any particular service) is approaching threshold, and that the flight path is considered "acceptable". However, on the SDSP-CD dashboard, if the icon button is yellow, the risk is slightly exceeding threshold and is considered an alert that the flight path needs to be re-planned. Additionally, the battery threshold setting is 15% for the HATIS compared to 25% for the SDSP-CD. These differences in threshold settings and interpretation of alert severity are examples of what should be clearly understood and communicated prior to operations.

It is also important to consider the consequences when operators select information on a display such that it masks critical display elements. For example, when any alert icon was selected on the SDSP-CD dashboard it would mask the color of the icon with a blue highlight color. When participants were asked to identify the number of alerts, and or types of alerts for specific vehicles, this masking effect prevented users from correctly assessing the information related to the hazards and risks.

Participants overwhelmingly stated that using the SDSP-CD prior to flight would indeed enhance safety. As mentioned in other summaries – being able to determine where and what types of hazards and whether a flight path needs adjustment (contingencies or other routes) to avoid hazards or minimize the impact of existing infrastructure (on-site, environmental) was reiterated. This means that an operator would be less likely to lose a signal or vehicle, exceed battery expectations (or know the location of risk), not fly over congested areas and impact something, and would know where the vehicle is during the mission. The task of the fleet manager to mitigate risks for a flight(s) is simplified in considering these hazards, as pieces of information are connected (e.g., not having to go between charts, Google Earth, multiple data sources), and do not overload the operator with information. One participant reported that the SDSP-CD depicts "the big 3": (1) internal risks (e.g., battery of the vehicle), (2) environmental risks to people and objects, and (3) support systems (e.g., GPS, compromised radio links). Risks could be monitored, particularly if the operator ran the plan every day.

Similarly, participants' descriptions emphasized that using the HATIS would ("definitely") enhance safety. This is because the interface allows operators to alter routes and permit evaluations (and try things out) of alternative trajectories prior to flights – an operator can consider a handful of threats simultaneously. This would be taxing if done manually. This interface allows more bandwidth for the operator to solve the issues (with time to check each vehicle in the system), help identify the hazard(s) with clear depictions of where hazards are encountered along the flight path, and reduce the risks to flight (i.e., avoid buildings, no fly zones, collisions). It also provides another layer of protection, as all the alerts are in one place with one tool, and there is continuous monitoring of each UAS. Safety concerns are (positively) addressed with this tool, but some level of hesitancy was noted with a user sharing that one can trust the data but understand that it is not infallible. This tool can create operations that are safe (a validation tool and "last safety check"), can limit or reduce risks (e.g., it will not violate no fly zones or hit buildings), saves time, and supports ensuring flight crew success. Using this interface could aid operators in identifying hazards when compared to using multiple sources and could also be used to understand how far off one is from an ideal flight path.

There was a range of limitations in this study, both expected and unexpected. One primary expected limitation was the system lag of the HATIS interface caused by operating it long distance, which resulted in a slightly negative impression of the HATIS experience for some users. The system lag was expected and conveyed to participants. They were informed that the system lag was not due to the program software itself, but due to inaccessible technical constraints. This system lag and minor difficulties may have affected other elements of the study. In some cases, the SDSP-CD also had brief minor technical difficulties which, although they were remedied quickly, still occurred.

The availability of HATIS' GPS service was unexpectedly disrupted during the study. The GPS service was unavailable half-way and through the remainder of the study; however, questions were changed to ask about alternate services and these were assessed instead (e.g., RFI or Population alerts) by users. This issue was spread mostly equally across the two groups of participants. Time stamp data was also collected during all tasks. However, these data may have been obscured due to additional time taken when participants asked the researcher to repeat the questions (i.e., due to unclear wording of questions, confusing terminology) and for some, difficulty navigating the cursor between monitors based on the computer set up.

There were discrepancies on some of the tasks – in particular, information that was portrayed on the paper maps differed from the GUI displays, and some questions had confusing wording or used unfamiliar terminology, despite running two pilot-participants to uncover these issues. Participants had some difficulties translating information between the mission maps and the GUI displays (e.g., trying to link an incorrectly identified group of people on the paper map with a population alert on the GUI). It is also possible that experimenters improved in training participants as the study progressed. These may have also had some effect on participants' performance on the tasks.

VI. Summary

In this study, the usability of the SDSP-CD and HATIS interfaces for sUAS preflight planning was explored. After training and completing tasks using each GUI, both interfaces were well regarded by users and indicated good usability. Overall, participants were also able to successfully assess risks and hazards using the interfaces when completing a variety of tasks during simulated nominal sUAS operations. Participants provided rich feedback based upon their experiences – including a wealth of information about both the capabilities and challenges unique to each interface and across systems. As fleet operators will be uniquely responsible for managing multiple potentially complex flights operating within uncontrolled airspace, informing the development of such interfaces is fundamental to the further advancement of preflight planning support tools to increase hazard detection, maintain situational awareness, improve workload management, and ensure safety in future operations.

Acknowledgments

This work was supported by the System-Wide Safety (SWS) project, under the Airspace Operations and Safety Program within the NASA Aeronautics Research Mission Directorate (ARMD). The authors would like to thank our pilot participants and human factor colleagues for their time, effort, and for sharing thoughtful feedback during this research study. We are also thankful for the continued support from the SWS Project team, Dr. Steve Young, Dr. Ersin Ancel, and Sloan Glover.

References

[1] Young, S., Ancel, E., Moore, A., Dill, E., Quach, C., Foster, J., Darafsheh, K., Smalling, Vazquez, S., Evans, E., Okolo, W., Corbetta, M., Ossenfort, J., Watkins, J., Kulkarni, C., and Spirkovska, L., "Architecture and information requirement to assess and predict flight safety risks during highly autonomous urban flight operations," NASA/TM-220440, 2020.

[2] Young, S. D., Ancel, E., Dill, E. T., Moore, A., Quach, C. C., Smalling, K., M., and Ellis, K. K., "Flight Testing In-Time Safety Assurance Technologies for UAS Operations," AIAA, Chicago, IL & Virtual, June 27-July 1, 2022.

[3] Ellis, K., Koelling, J., Davies, M., and Krois, P., "In-time System-wide Safety Assurance (ISSA) Concept of Operations and Design Considerations for Urban Air Mobility (UAM)," NASA/TM-2020-5003981, Hampton, VA, 2020.

[4] Aweiss, A. S., Owens, B. D., Rios, J. L., Homola, J. R., and Mohlenbrink, C. P., "Unmanned aircraft systems (UAS) traffic management (UTM) national campaign II," AIAA SciTech Forum, AIAA Infotech@Aerospace, Kissimmee, FL, January 8-12, 2018. doi.org/10.2514/6.2018-1727

[5] Degani, A., Barshi, I., and Shafto, M. G., "Information organization in the airline cockpit: Lessons from Flight 236," *Journal of Cognitive Engineering and Decision Making*, Vol. 7, No. 4, 2013, pp. 330-352.

[6] Endsley, M. R., "Toward a theory of situation awareness in dynamic systems," *The Journal of the Human Factors and Ergonomic Society*, Vol. 37, No. 1, 1995, pp. 32-64.

[7] Ancel, E. Capristan, F.M., Foster, J.V. and Condotta, R.C., "In-time non-participant casualty risk assessment to support onboard decision making for autonomous unmanned aircraft," AIAA Aviation 2019 Forum, 2019, p. 3053.

[8] Kulkarni, K. and Corbetta, M., "A Hybrid Battery Model for Prognostics in Small-Size Electric UAVs," Annual Conference of the Prognostics and Health Management Society, 2018.

[9] Corbetta, M., Banerjee, P., Okolo, W., Gorospe, G. and Luchinsky, D.G., "Real-time UAV trajectory prediction for safety monitoring in low-altitude airspace," AIAA Aviation 2019 Forum, AIAA 2019-3514, June 2019, p. 3514.

[10] Banerjee, P. and Corbetta, M., "In-time UAV flight-trajectory estimation and tracking using Bayesian filters," 2020 IEEE Aerospace Conference, IEEE, Big Sky, MT, 2020, pp. 1-9, doi:10.1109/AERO47225.2020.9172610.

[11] Teubert, C., Daigle, M. J., Sankararaman, S., Goebel, K. and Watkins, J., "A Generic Software Architecture for Prognostics (GSAP)," *International Journal of Prognostics and Health Management*. September 2017; Vol. 8, No. 2, 013.

[12] Moore, A. J., Schubert, M., Fang, T., Smith, J. and Rymer. N., "Bounding Methods for Heterogeneous Lidar-derived Navigational Geofences," NASA Technical Memo 22399, NASA Langley Research Center, Hampton, VA, 2019.

[13] Gilabert, R., Dill, E., and Uijt de Haag, M., "Evaluation of Improvements to the Location Corrections through Differential Networks (LOCD-IN)," Proceedings of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+), Miami, FL, September 2019.

[14] Banerjee, P., Corbetta, M. and Jarvis, K., "Probability of Obstacle Collision for UAVs in presence of wind," 2022 AIAA Aviation Forum, Chicago, IL, 27 June-1 July, doi.org/10.2514/6.2022-3460.

[15] Feldman, J., Martin, L., Bradley, J., Walter, C., and Gujral, V., "Developing a dashboard interface to display assessment of

hazards and risks to sUAS flights," AIAA Aviation Forum, Chicago, IL & Virtual, June 27-July 1, 2022. doi.org/10.2514/6.2022-3462
[16] Hart, S. G., NASA-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, October 2006, Vol. 50, No. 9, pp. 904-908.
[17] Brooke, J., "SUS: A "quick and dirty" usability scale," *Usability Evaluation in Industry*, edited by P. W. Jordan, B. Thomas, B. A. Weerdemeester, and A. L. McClelland, London, 1996, Taylor and Francis.
[18] Buck, B. K., Chancey, E. T., Politowicz, M. S., Unverricht, J. R., and Geuther, S. C., "A remote vehicle operations center's role in collecting human factors data," AIAA SciTech Forum, January 23-27, 2023, National Harbor, MD and Online.