Usability Evaluation Technical Report

A Usability Evaluation of the Communication Tool prototype prepared in Adobe XD for High Density Vertiplex (HDV) team

Report Prepared by: Shraddha Swaroop

Table of Contents

Executive Summary

The National Aeronautics and Space Administration (NASA) High Density Vertiplex (HDV) subproject conducts simulation and live flight tests on a periodic basis in which it is necessary for the broad team that ranges between two NASA bases in Virginia and California to work together and know what the other is doing to do their job.

A high level of team situation awareness is required to allow the many roles and people playing those roles to be aware of what steps have been completed so the next one can be started in a timely fashion that can keep the simulation or live test running in a seamless fashion. Qualitative user interviews revealed that 60 percent of the participants saw a gap in communications. That was a key problem that needed to be solved so this team created an Adobe XD prototype of a proposed tool to be used in the next live flight scheduled for March 2023.

This report includes workflow analysis and usability testing results for this prototype. Usability testing is a method for understanding usability of a digital product or website by asking participants to do certain critical tasks to examine the usability of an interface. It is an easy, inexpensive and repeatable process that is scientifically shown to reveal problems that the target audience is having with an interface as well as areas that are working well for the user. This activity aligns with the fifth objective in the project plan which says "create a plan to research, test and develop a tool by November 2022.

Methods

Our team of one human factors graduate student and NASA HDV data manager worked together to compile the information contained in this report. The human factors graduate student created a usability test plan (attached as **Appendix A**) with the following six steps:

1. **Determine User Goals:** Determined via User Needs Analysis of *n* = 6 HDV employees (the complete User Needs Analysis is attached as part of **Appendix B**)
2. **Determine Customer Requirements:** Determined from information as a result of a series of qualitative user interviews as shown in **Appendix A.**
3. **Determine Goals of Evaluation:** Determine how well the existing iteration of the prototype provides readily-available, easily-accessible and important information to current and researchers (task performance). Identify usability issues that hinder access to relevant information or cause frustration in those using this tool (user satisfaction).
4. **Determine Evaluation Methods:** Usability evaluation
5. **Detail Methods Being Used**: Determined by the scope of the project, availability of participants, and time constraints:
   - Moderated formative usability testing (appropriate for iterative

design, limited participants, and quick turnaround time)

6. **Identify Test Metrics:**
   - Time on Task
   - Task success rate
   - Self-reported task difficulty
   - System usability scale for Websites (SUS) – Form attached as **Appendix E**

A moderated formative usability test was conducted with a convenience sample of six current HDV team members including team leads, researchers and developers.

Overview of Performance, Efficiency, and User Perception Metrics

Following is an overview of the data. A detailed discussion of each is contained in the body of this report.

User Perception: System Usability Scale

SUS score by participant. The gold bar represents the minimum score required to obtain marginal usability. The green bar represents the minimum score required to obtain acceptable usability. The average SUS score was more than 83. All of the participants were in the acceptable range which was above the green bar of 70 and nobody was in the marginal usability zone between 50 and 70.

Overview of Efficiency: Time on Task, Task Success Rate, and Task Difficulty

Most of the participants did well in terms of time on task, success rate and task difficulty on all the tasks with the exception of Task 3.

Task	**Average time* (in second) spent on task**	**Average task success rate*** **	**95% Confidence Interval****	**La Place Point estimation**	**Average reported task difficulty**
1	22.20	1	.64 - 1.00	.88	4.83
2	58.28	1	.64 - 1.00	.88	4.33
3	84.68	1.3	.23 - .91	.63	4.25
4	19.61	1	.64 - 1.00	.88	4.00

*Geometric mean **Adjusted Wald

***Coded as 1= success; 2= success with difficulty; 3= failure

Task Key:

Task #1: Log in as a GCSO researcher and obtain security clearance
#2: Find the upcoming task for Range Safety Officer
#3: Find upcoming task for yourself, the GCSO researcher
#4: Check off the next test for yourself, the GCSO researcher, to let the team know that step is complete.

Summary of Design Recommendations

- The Flight Overview list should include duties and roles for everyone beyond ROAM including pilots on the ground and those at NASA Ames to give everyone proper team situation awareness
- This tool should include the ability to filter for different roles
- This tool should include an avatar or some way to indicate the role(s) that logged in
- This tool should have an editing capability to edit the Flight Overview List for each run on the Home Page as needed as well as a capability to leave notes for the team on the My Tasks page
- This tool should feature task numbering on the My Tasks page for duties that correlates to the design of the list on the Home Page

Introduction

The National Aeronautics and Space Administration (NASA) High Density Vertiplex (HDV) subproject conducts simulation and live flight tests on a periodic basis in which it is necessary for the broad team that ranges between two NASA bases in Virginia and California to work together and know what the other is doing to do their job.

A high level of team situation awareness is required to allow the many roles and people playing those roles to be aware of what steps have been completed so the next one can be started in a timely fashion that can keep the simulation or live test running in a seamless fashion. Qualitative user interviews revealed that 60 percent of the participants saw a gap in communications. That was a key problem that needed to be solved so this team created an Adobe XD prototype of a proposed tool to be used in the next live flight scheduled for March 2023.

This report includes usability testing results for this prototype. Usability testing is a method for understanding usability of a digital product or website by asking participants to do certain critical tasks to examine the usability of an interface. It is an easy, inexpensive and repeatable process that is scientifically shown to reveal problems that the target audience is having with an interface as well as areas that are working well for the user. This activity aligns with the fifth objective in my project plan which says "create a plan to research, test and develop a tool by November 2022."

As such, the purpose of this report is to apply human factors principles and usability guidelines to the iterative design of the Communication Tool prototype and to inform management and designers about the prototype's current usability with the aim of making improvements where they are most necessary in an effort to improve the ease of which researchers access important information. This report is organized as follows:

- The Background section describes the purpose of this report.
- The Method section details the steps our team of one human factors graduate student and HDV data manager took to reach our conclusions and recommendations.
- The Results section outlines the results of the formative usability evaluation along with a detailed analysis of each of the performance and self-report metrics used in the evaluation.
- The discussion, conclusion and recommendations section summarizes our team's findings and makes recommendations based on the information contained in the foregoing sections.

Background

Team situation awareness includes two or more members of a team dealing with multiple information resources to accomplish a shared goal (Salas, 2004). Teams are required to collectively detect and interpret cues, remember, reason, plan, solve problems, acquire knowledge and accomplish the goal as an integrated unit (Cooke, 2004). What is critical is that every individual team member have the information each needs to accomplish his/her responsibilities (Endsley, 1995) which is what this communication tool aims to help researchers achieve.

However, different roles in a simulation and flight test are performing myriad other duties in addition to using this communication tool so the concern that this may increase workload and distract from other duties is a concern that will be addressed in the usability study. The team is interested in identifying and addresses current design issues in the prototype that may impact usability for key users of the Tool, such as researchers, developers and tech leads. There are some fixed limitations of what can be developed for this tool. For instance it must stay within the security protocols that NASA has set for all its systems.

As such, the information contained in this report is meant to be a comprehensive usability evaluation of the Tool and may be used by anyone charged with designing and maintaining the Tool to create a usable website within the limitations of its current structure.

Method

Our team of one human factors graduate student and the HDV data manager worked to compile the information contained in this report; however, we each prepared individual and separate written reports.

Usability Test Plan

Our team's first step was to create a usability test plan, which allowed us to identify *six* important pieces of information that we used to guide our testing, analyze our results, and formulate our recommendations. The complete and detailed Usability Test Plan is attached as **Appendix B**. An overview of the usability plan is as follows:

1. **User Goals:** Determined via User Needs Analysis of $n = 6$ researchers/developers (the complete User Needs Analysis is attached in **Appendix B**)

   The major characteristics of likely users are:

- Most are HDV researchers ($n = 4$)
- Most are male ($n = 4$)
- Most are age 20 - 30 ($n = 4$)
- Most have 0 to 5 years of experience in their role ($n = 4$)

- All enjoy middle to strong team situation awareness (n = 6)

2. **Customer Requirements:** Determined from information compiled by both members of the team after the human factors graduate student conducted qualitative user interviews. A complete and detailed list of customer requirements is attached as **Appendix A.**
3. **Goals of Usability Evaluation:** Determined from steps 1 and 2 above:
    - Determine how well the existing prototype provides readily-available, easily-accessible and important information to current HDV team members (task performance).
    - Identify usability issues that hinder access to relevant information or cause frustration for users of the communication tool (user satisfaction).
4. **Identify Appropriate Evaluation Methods:** Determined by the scope of the project, availability of participants, and time constraints:
    - Moderated formative usability testing (appropriate for interactive design, limited participants, and quick turnaround time)

5. **Details of the Evaluation Methods:**
    - Usability tests conducted with six HDV team members from both NASA-Ames and NASA-Langley bases

6. **Identify Test Metrics:**
    - Time on Task
    - Task success rate
    - Self-reported task difficulty
    - System usability scale for Websites (SUS) in Appendix E.

Scenarios and Test Tasks

Based on the results of our qualitative interviews and our identification of customer and user goals, our team derived four tasks corresponding to four hypothetical use scenarios as follows:

Task #1: Log in as a GCSO researcher.

Scenario #1: You are a GCSO researcher named Bert Johnson working on your first live flight. You are seated in your station at ROAM about to start the live flight test for HDV. Please log in so you can get more information. When you obtain clearance, tell me what the message says

Task #2: Review tasks and find the next upcoming task for the Range Safety Officer.

Scenario #2: As a researcher, you know there are a lot of roles working on this test. It is imperative for you to understand who is doing what so you know when your tasks are due. For this task, I would like you to navigate to find information on the range safety officer's next upcoming task. Tell me what that person is doing and what number that task is.

Task #3: Locating your tasks

Scenario #3: Now you want to make sure you complete everything that needs to be done by your role. How do you find out what those tasks are? Tell me what your next task is.

Task #4: Complete task

Scenario #4: You have just completed a task, now you want to make sure to check it off so your team members and the FTD know you are finished. How would you do that?

Procedure

Participants were a convenience sample of HDV team members. The human factors graduate student scheduled and conducted each moderated usability test for all six participants. The usability tests were conducted via Zoom and Teams, were recorded and transcripts were generated. Participants followed the below structure:

- Complete consent form to participate in the test and to be recorded.
- Completed pre-trial demographic questionnaire
- Spent a few minutes looking around the Communication Tool prototype to gain familiarity with it before the test tasks began
- Once participants were ready, the task scenarios were read to the participant and provided for them in the chat function of Zoom and Teams.
- Each task was timed from start to finish by the moderator with a smartphone stopwatch.
- After each task, participants were asked to rate the difficulty of each task:
    - "Rate your experience on this on a scale of 1-5, with 1 being very hard and 5 being very easy.
- At the end of the test session, participants completed (via a Google Forms link) the System Usability Scale (SUS) and verbally answered four open-ended questions to provide insight into their perceptions of the tool's usability (see **Appendix E** for the SUS scale and **Appendix H, I, J** for the open-ended questions).

Usability Metrics

Data relating to three categories of usability metrics were collected: user preferences, task performance, and task efficiency.

User Preferences Metrics

• System Usability Scale (see **Appendix E** for the scale questionnaire)

• Post-study open-ended questions (see **Appendix J** for the

questions)
• Post-task open-ended question (see **Appendix H** for the questions)

Task performance and Task Efficiency Metrics

• Time on task
• Task success (binary success/fail)

Results, Discussion, Recommendations

Participant demographics are presented first, followed by summary data, and a detailed analysis of the three least problematic tasks and the three most problematic tasks.

Participant Demographics

- Gender: Female (2), Male (4)
- Ages: 20-30 (4); 31-40 (2)
- Roles: Tech lead (1); Researchers (4); Developer (1)
- Experience: 0 to 5 years (4); 6 to 10 years (1); 11 to 15 years (1); 16+ years
- Base: NASA-Ames (4); NASA-Langley (2)
- HDV Flight or Simulation experience: 0 to 5 years (4); 6 to 10 years (0); 11+ years (1); Other (0)
- Team situation awareness: (1=weak and 5 is strong): 1 (0); 2(0); 3(3); 4(2); 5(1);

Quantitative Results

An overview of the data for each metric utilized to determine usability of the Website is as follows:

System Usability Scale

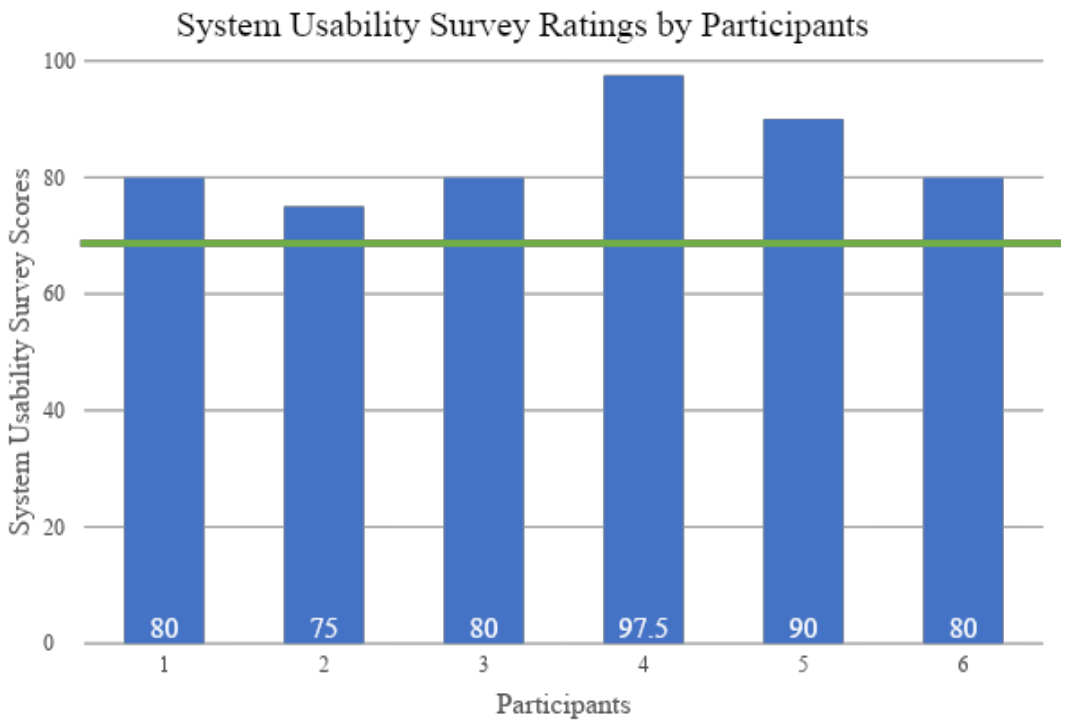
The SUS is a self-report measure that asks participants 10 questions relating to their perceptions of the communication tool prototype. The odd-numbered questions are positively-worded and the even-numbered questions are negatively-worded.

The SUS is scored as follows:

- Odd-numbered questions: subtract 1 from the participant's rating
- Even-numbered questions: subtract the participant's rating from 5
- Add up the new scores then divide by 2.5
- Scores range from 0 – 100
- Scores < 50 = unacceptable usability
- Scores 50 – 70 = marginal usability
- Scores > 70 = acceptable usability

The SUS score for each of our participants appears below. The gold bar represents the minimum score required to obtain marginal usability. The green bar represents the minimum score required to obtain acceptable usability. Here, the average SUS score was almost 84, which is well above acceptable

usability.



Completion time, accuracy rate and self-reported task difficulty

Data are summarized below by task. For a look at individual participant data, See Appendix H.

Task	Average time* (in second) spent on task	Binary success rate	95% Confidence Interval**	La Place Point estimation	Average reported task difficulty**
1	22.20	1	.64 - 1.00	.88	4.83
2	58.28	1	.64 - 1.00	.88	4.33
3	84.68	1.3	.23 - .91	.63	4.25
4	19.61	1	.64 - 1.00	.88	4.00

*Geometric mean **Adjusted Wald

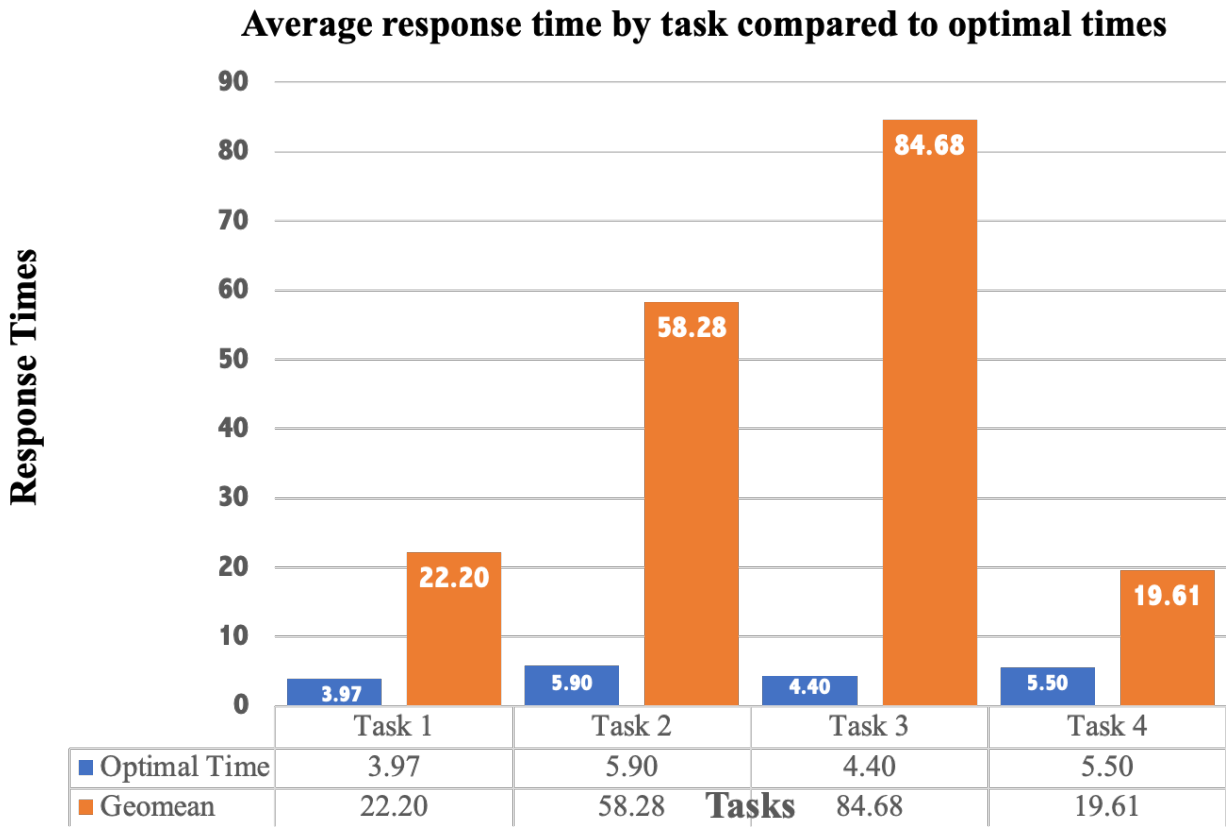
Time on Task Summary Data

To find the central point of the time on task data, we used the geometric mean, as it is the most appropriate when analyzing centrality of highly skewed and variable data such as time-related data. The optimal task time, range, and

standard deviation are also provided in the chart below.

Participant	**Task 1**	**Task 2**	**Task 3**	**Task 4**
1	30	105	130	35
2	29	45	27	14
3	15	29	420	7
4	18	27	41	7
5	14	107	107	37
6	82	99	57	64
Geomean	22.20	58.28	84.68	19.61
Optimal Time	3.97	5.90	4.48	5.50
Range	14-82	27-107	27-420	7-64
Standard Deviation	25.77	38.93	147.30	22.37

Task Key:
Task #1: Log in as a GCSO researcher and obtain security clearance
#2: Find the upcoming task for Range Safety Officer
#3: Find upcoming task for yourself, the GCSO researcher
#4: Check off the next test for yourself, the GCSO researcher, to let the team know that step is complete.



The two tasks with the highest geomean completion times were as follows:

- Task 2 (Find the upcoming task for Range Safety Officer): 58.28 seconds
- Task 3 (Find upcoming task for yourself the GCSO researcher): 84.68 seconds

The above summary results allowed our team to easily identify the single most problematic task and the two least problematic tasks, each of which are discussed in detail below. Because our test tasks mainly related to finding information on the prototype tool, we identified Time on Task, and Success Rate as the most important measures on which to focus our detailed evaluation.

Listed below are the three least problematic tasks. As these tasks had relatively high success rates, relatively low task times, and low difficulty ratings, the remainder of this report will focus on the one task that we deemed to be the most problematic.

Three Least Problematic Tasks

Task	Success Rate	Geomean time on task (seconds)	Optimal time* (seconds)	Mean difficulty rate (out of 5)
#1: Log in as a GCSO	6/6	22.20	3.97	4.83

researcher and obtain security clearance				
#2: Find the upcoming task for Range Safety Officer	6/6	58.28	5.90	4.33
#4: Check off the next test for yourself, the GCSO researcher, to let the team know that step is complete.	6/6	19.61	5.50	4.00

*Note: Optimal time was determined by team members who were familiar with the fastest way to accomplish the task. Participants were unfamiliar with where to find the information and were also thinking out loud as they were completing the tasks. As such, the discrepancies between the optimal time and participants' actual times were expected.

Single Most Problematic Task

As stated above, our team focused on Time on Task and Success Rate as the most important metrics on which to focus our detailed evaluation of the single most problematic task which was Task #3.

Task	**Success Rate**	**Geomean time on task (seconds)**	**Optimal time* (seconds)**	**Mean difficulty rate (out of 5)**
#3: Find upcoming task for yourself, the GCSO researcher	4/6	84.68	4.48	4.25

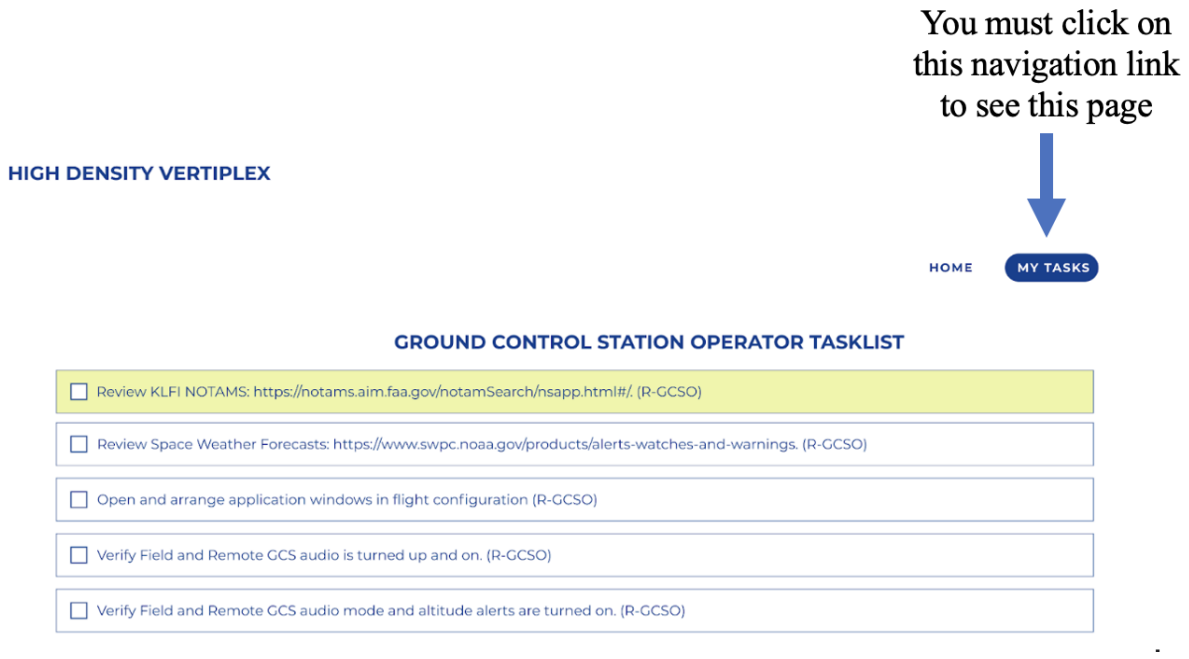
Task # 3: Find upcoming task for yourself, the GCSO researcher

Success rate: 4/6; (One participant could not find this information immediately and eventually made it to the My Tasks tab to complete the activity but exceeded the 2 minute time limit; the other participant did find the information but also exceeded the time limit; the other four found this information on the Home Page instead of the My Tasks page.

Geomean task time: 84.68 seconds

Optimal time: 4.48 seconds

Difficulty score: 4.25/5.0



Half of the subjects saw the upcoming tasks for the GCSO researcher on the Home page and either didn't see the My Tasks button or were waiting for it to light up to go there. One person did eventually get there and the other did not. The participant did not get there but did go there for Task 4 when was trying to check off the task as completed. The information about the next task can be found on the Home Page but it is static. The only way to interact with it is to go to the My Tasks page by clicking the My Tasks link in the navigation.

Four participants were able to complete this, but the majority expressed frustration at having to go to a whole new page to see different tasks based on different roles:

> "So that's also something that I guess wasn't maybe obvious to me. "
> "I would think that the interface would make it more clear which ones are mine versus theirs, right? Cause that

Design recommendation

- Filters should be available on the home page to show which roles are assigned to which duties to make it easier to follow along.

Conclusion

Although users had some difficulty finding important information on the prototype, we also asked them to share their thoughts about the prototype in general and to highlight some things they liked the most and the least about the Website. We used the test data along with these comments to make our five recommendations.

Comments on workload

- "It would be just slightly more workload: Having to monitor this to make sure that, like if we're up, doing what we need to do."
- "I think that kind of coordination will help with the workload cause part of the workload anyway is just trying to get statuses on what step the other person's on."
- "Highlighting the current step is extremely useful and it offloads a lot of the workload from the researcher to keep track of where they are. I do believe that that's a benefit because the researchers don't have the key track in their working memory."

Comments on team situation awareness

- "It would help with situation awareness because you would have one source of reference for where the where the checklists are"
- "So you know, but for my perspective for situation awareness 100%, I think that this checklist would make it a lot easier for me to know where I was."
- "I think there is value in having an overarching list like this, where you got every task that needs to be accomplished for a given run. Because I think for awareness it's important to know. Where am I in this step? Am I up next in a task that I need to do?"

Suggestions for improvement from participants

"It would be good to have numbers associated with tasks on My Tasks. I don't see numbers associated with them, so I don't know if these are in order. It looks like that top. One is highlighted here as well, so that sort of indicates that that's where we are in the flow. But again it would be good to have the numbers."

"What would be useful here (on Home Page) is to have tabs up here to split the checklist based on our roles. So if I clicked on RSO tab, I would see all their tasks. They can narrow it down to who they want to see."

"I imagine we would need to create a Master Checklist for all

Summary of recommendations

In compiling this report and making recommendations to the department, we considered the intended users' goals and the customer's goals and took into consideration some suggestions for improvement made by the participants. Discussion of our analyses that formed the basis of our recommendations are included in the Results section above. A generalized summary of those recommendations are as follows:

- The Flight Overview list should include duties and roles for everyone beyond ROAM including pilots on the ground and those at NASA Ames to give everyone proper team situation awareness
- This tool should include the ability to filter for different roles
- This tool should include an avatar or some way to indicate the role(s) that logged in
- This tool should have an editing capability to edit the Flight Overview List for each run on the Home Page as needed as well as a capability to leave notes for the team on the My Tasks page
- This tool should feature numbering on the My Tasks page for duties that correlates to the design of the list on the Home Page

References

N.J. Cooke, Measuring team knowledge, N.A. Stanton, A. Hedge, K. Brookhuis, E. Salas, H. Hendrick (Eds.), Handbook of Human Factors Methods, CRC Press, Boca Raton, USA (2004)

M.R. Endsley, Towards a theory of situation awareness in dynamic systems. Hum. Factors, 37 (1995), pp. 32-64

E. Salas, Team methods, N.A. Stanton, A. Hedge, K. Brookhuis, E. Salas, H. Hendrick (Eds.), Handbook of Human Factors Methods, CRC Press, Boca Raton, USA (2004)

Taylor, R.M., 1990. Situational awareness rating technique (SART): the development of a tool for aircrew systems design. In: Situational Awareness in Aerospace Operations (AGARD-CP-478). NATO-AGARD, Neuilly Sur Seine, France, pp. 3/1 –3/17.

Appendix A: Customer requirements

Use case: Researchers and tech leads involved in a HDV flight test or simulation said a lack in transparent communication about a preceding step being accomplished could delay them in the step he/she is responsible for. This lag caused the system to log out and restarting the system caused lateness in accomplishing their step in the process.

- Create a communication tool that allows researchers to the status of what step is being conducted in the live flight test or simulation.
- This tool shall cue researchers to engage specific tasks as appropriate for their role.
- This tool shall also tell the researcher how preceding steps affects their own tasks.
- This tool shall not distract or interfere with research activities.

Appendix B: Usability Test Plan

Usability Test Plan for HDV Communication Tool

Version 1, November 21, 2022

Table of Contents

1 Executive Summary

This document describes the usability test plan for HDV communication tool Interface for researchers. This document serves as a concise summary of the test, specifying the procedure, tasks, the roles and responsibilities of everyone involved, and describing the timelines.

Section 2 (on the next page) describes the participants and responsibilities of everyone involved in the test, including the testing team, the observers and the test participants.

Section 3 describes the evaluation procedure, including the location and dates, the test facilities and how the test sessions will be organized.

Section 4 describes the tasks that will be used in the test. This section describes the core functions that will be tested.

Section 5 describes the usability data that we will collect from the test and contains the agreed usability metrics that we will test the product against.

The appendices contain our working documents for the test, such as questionnaires, consent form, user flow, and screengrabs of interface for use during test.

2 User needs analysis

1. Identify the different classes of users (flight test director, researchers, tech leads, developers.).
    i. Researchers for (ASI) and Fleet Operations (FO) teams
    ii. Tech leads for ASI, VVSO and FO teams
    iii. Flight test director
    iv. Developers for ASI, VVSO and FO teams
2. Characteristics of users (that will affect how they use/interact with product)
    i. Gender: Male or female
    ii. Age: Between 18 and 35 years old
    iii. Expertise: Worked on High Density Vertiplex (HDV) subproject
    iv. Frequency of usage: None
3. Operational procedures
    i. Determined by the users
4. Performance criteria (speed, accuracy, quality, consequences of not meeting criteria)
    i. Completion time
    ii. Accuracy
5. Task demands (physical, perceptual, cognitive, health and safety)
    i. Physical, perceptual, cognitive, health and safety
6. Environment where product will be used
    i. Over Teams call with training provided beforehand
7. Availability of technical assistance
    i. None except what can be found on prototype

3. Participants and responsibilities

3.1 PARTICIPANTS

Based on a Nielsen-Norman Group advice for qualitative data in a usability study, it would be good to test between 5 to 7 participants in the usability test. Recruitment will be done informally. Participants will be researchers, developers and tech leads of the High Density Vertiplex (HDV) team who are available on test days.

3.2 RESEARCHERS

Researcher will be conducting a formative moderated usability test which will the following goals:

- Measure time it takes to complete actions
- Evaluation of system usability and situation awareness

Research questions:

- 1: Do researchers and tech leads differ in performance while using interface?
- 2: What impact does this tool have on subjectively assessed workload?
- 3: Do situation awareness scores increase or decrease after the use of this interface?

3 Evaluation Procedure

3.1 LOCATION AND DATES

The usability test will be carried out remotely via Teams and recorded. We plan to test participants according to the schedule shown in Table 1.

Table 1: Schedule for the consent (Appendix A), demographics survey (Appendix B) usability test and System Usability Survey (Appendix C).

	Dec. 2	Dec. 5	Dec. 7	Dec. 8	Dec. 14
9:30 am or anytime subject is available	Participant 1	Participant 2	Participant 4		Participant 6
1 pm or anytime subject is available		Participant 3		Participant 5	

3.2 TEST SESSIONS

Each participant session will be organized in the same way to facilitate consistency. Users will be asked to commence each task from the login page of HDV communication tool prototype in Teams call.

Ask participant to fill out Consent form: **https://forms.gle/pJuj1E1B4gpzRPuj7**
Ask participant to fill out Demographic survey: **https://forms.gle/w6A4ZkkPtP4AfHy27**

Moderator will begin the session with the following: "Thank you for participating in our study today. For the next 30 minutes, please consider you are using our new HDV communication tool.

During the course of this session I'll ask you to complete some activities and then ask you some questions related to those activities. We want to see how you interact with it and solicit your feedback as a user. Any difficulties you encounter can be valuable feedback to us, so don't worry about making mistakes or struggling to complete an activity! There are no right or wrong responses, there's just potential for XTM client interface to improve.

If you get stuck at any time, you are welcome to ask any questions.

Do you have any questions for me before we begin? If not, let's get started."

Tasks

In the first part of the usability test, the researcher will ask participants to carry out various tasks with the product. The researcher will prompt participants to keep up a good flow of comments as they work. The tasks include:

1. Log in as a GCSO researcher
2. Determine the first task for the range safety officer is on the Flight Plan Overview
3. Determine the first task you must complete as a GCSO researcher
4. Complete the task so the test can run smoothly

Questionnaires

At the beginning and end of the test session, participants will complete the Situation Awareness Response Technique questionnaire shown in Appendix C. System Usability Survey (Appendix D) will be completed by participants when simulations are complete for the day.

4 Test tasks

The test tasks are described below.

<table>
<tr><th colspan="2">Task 1: Log in as a GCSO researcher</th></tr>
<tr><td colspan="2">Moderator: I am going to share a wireframe link to you via Microsoft teams chat. Please click on the link and share your screen. Please don’t click on anything until we start the tasks.

Share link : https://xd.adobe.com/view/0b1d5e28-04bc-45e3-b308-14b2a407246c-5e4e/?fullscreen

Ready? You are a GCSO researcher named Bert Johnson working on your first live flight. You are seated in your station at ROAM about to start the live flight test for HDV. Please log in so you can get more information. When you obtain clearance, tell me what the message says</td></tr>
<tr><td>Moderator:

Knowledge Probe: How did you expect this interface to work?

Perception Probe: Did you expect the log in display to look like this?</td><td>Observation/Comments:
Note any confusion, hesitations, or difficulties during the process.

Straightforward. Some kind of immediate feedback. Delay second gues

Verbal Comments:</td></tr>
<tr><td>Subjective Q: Rate your experience of logging into the interface on a scale of 1-5, with 1 being very hard and 5 being very easy.
[if low] Explain</td><td>1=Very Hard 5=Very Easy
Verbal Comments:

5</td></tr>
</table>

Q: Was there anything in the UI that confused you? **[If yes]** Explain	❑ Yes ❑ No **Verbal Comments:**
Response Time: Accuracy: (Success/Fail/Succeed w Difficulty)	

Task 2: Review tasks and find next upcoming task for the Range Safety Officer	
Moderator: As a researcher, you know there are lot of roles working on this test. It is imperative for you to understand who is doing what so you know when your tasks are due. For this task, I would like you to navigate to find information on the range safety officer's next upcoming task. Tell me what that person is doing and what number that task is.	
Moderator: Navigate to the section where you would expect to find information on Range Safety Officer's upcoming task.	**Observation/Comments:** Note any confusion, hesitations, or difficulties during the process**.** **Verbal Comments:** **Unexpected to take me to anothe page**

Knowledge Probe: Where do you expect to see the information on the Range Safety Officer?	**Verbal Comments:**
PROBE: Is this adding to your workload? Is this helping your team situation awareness?	**Verbal Comments:**
PROBE: What do you think this information can be used for?	**Verbal Comments:**
Q: Rate your experience of navigating to find information on the Home Page of 1-5, with 1 being very hard and 5 being very easy. **[if low]** Explain	1=Very Hard 5=Very Easy **Verbal Comments:**
Q: Was there anything in the UI that confused you? **[If yes]** Explain	❑ Yes ❑ No **Verbal Comments:**
Response Time: Accuracy: (Success/Fail/Succeed w Difficulty)	

Task 3: Locating your tasks
Moderator: Now you want to make sure you complete everything that needs to be done by your role. How do you find out what those tasks are? Tell me what your next task is.

Moderator: Now you want to make sure you complete everything that you need to do during this live flight test. How do you find out where those tasks are?	**Observation/Comments:** Note any confusion, hesitations, or difficulties during the process. ❑ *Refers to the My Tasks page to check for tasks* **Verbal Comments:**
Q: What are your thoughts on finding out what you need to do?	**Verbal Comments:**
Q: Rate your experience of finding out what a GCSO researcher has to do on a scale of 1-5, with 1 being very hard and 5 being very easy. **[if low]** Explain	1=Very Hard 5=Very Easy **Verbal Comments:**
Q: Was there anything in the UI that confused you? **[If yes]** Explain	❑ Yes ❑ No **Verbal Comments:**
Response Time: Accuracy: (Success/Fail/Succeed w Difficulty)	

Task 4: Complete task
Moderator: You have just completed a task, now you want to make sure to check it off so your team members and the FTD know you are finished. How would you do that?

Moderator: You have just completed the initial task on your list, how would you make sure that your team members and FTD know it is done?	**Observation/Comments:** • Note any confusion, hesitations, or difficulties during the process.
Q: Rate your experience letting your team know your task is finished on a scale of 1-5, with 1 being very hard and 5 being very easy. **[if low]** Explain	1=Very Hard 5=Very Easy **Verbal Comments:**
Q: Was there anything in the UI that confused you? **[If yes]** Explain	❑ Yes ❑ No **Verbal Comments:**
Response Time: Accuracy: (Success/Fail/Succeed w Difficulty)	

5 Data collection

5.1 DATA LOGGING

Note down any difficulties faced by the participant and associated comments. Each observation will comprise an observational code and a short description of the behaviour. This is the observational code that will be used:

S	Start task
E	End task
G	General comment
P	Positive opinion
N	Negative opinion
X	Usability problem
*	Video highlight — an "Ah-ha!" moment
B	Bug
F	Facial reaction (e.g. surprise)
A	Assist from researcher
Q	Gives up or wrongly thinks finished
H	Help or documentation accessed
M	Misc (general observation by researcher)[GCM(JSURF1]

5.2 USABILITY METRICS

Accuracy

Measurement of if the participant was able to finish the task or if he/she failed the task. Success will get a score of 1 and failure will get a score of 0.

Response time

The number of seconds it takes each participant to do each task in seconds. This will be used to calculate a geomean for each task.

Questionnaires

We will administer two questionnaires (see Appendix C and D).

Appendix C: Consent Form

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
HIGH DENSITY VERTIPLEX (HDV) SUBPROJECT
DATA VISUALIZATION TEAM

Consent to Participate in Usability Testing

I am conducting a usability test as part of a class at California State University Long Beach called PSY689, Practicum, Fall 2022. If I agree to participate, I will log into Teams and be asked to complete a questionnaire on my background. Then I will perform various tasks on an Adobe XD prototype. Finally, I will be asked to complete two follow-up questionnaires. The entire process should take less than one hour. I am aware that I will receive no compensation for my participation.

I am aware that I will be observed and videotaped while using the software program. This tape will be used to analyze the usability of the prototype only. It is not in any way a test of my ability to operate a computer or my knowledge of aviation or air traffic control

My participation in this test is voluntary. My relationship or status with NASA will not be affected by my decision to participate. I may refuse to participate or discontinue my participation at any time without penalty. There are no foreseen risks that will result from my participation in this study. I will probably not receive any direct benefits from this study, but the knowledge gained by the researcher may improve the design of future software products and flight test experience.

The videotapes will be viewed by the members of the Data Visualization Team. If I agree, portions of these tapes may also be shown at conferences on computer interface design.

If I have any other questions or wish to report a research-related problem, I may contact Shraddha Swaroop at 757-286-6125, shraddha.swaroop@nasa.gov.

Based on the foregoing, typing my name and the date below and clicking accept indicates my agreement to participate.

Based on the foregoing, typing my name and the date below and clicking accept indicates my agreement to participate.
o Agree
o Reject

Full name:

_____

Today’s date:

_____

Appendix D: Demographic survey

Participant questionnaire

Thank you for participating in our study. This questionnaire seeks demographic information about you and your role on the High Density Vertiplex (HDV) subproject. The information you provide is strictly kept confidential. Completion should take no more than 5 minutes. Thanks for your time - your expertise is valuable to our project.

What is your gender?
o Male
o Female
o Prefer not to say
o Other

How old are you?
o Less than 19
o 20 to 30
o 31 to 40
o 41 to 50
o 51 and over

What is your role at HDV?
o Tech lead
o Researcher
o Developer
o Other:

How many years of experience do you have in this role?
o 0 to 5
o 6 to 10
o 11 to 15
o 16+ years
o Other:

Which base do you work out of?
o NASA-Ames
o NASA-Langley

o Other:

How many times have you been involved a simulation or flight test for High Density Vertiplex (HDV) team?
o 0 to 5
o 6 to 10
o 11 +
o Other:

How would you rate team situation awareness on the AOA Flight Test?

Weak 1 2 3 4 5 Strong

Appendix E: System Usability Scale (SUS)

System Usability Scale

Instructions: After you put in your participant ID, and date, please give us your honest opinions on each of the statements. For each of the following statements, mark one box that best describes your reactions to the website today. This should take no more than 4 minutes. Your help and experience on the site will greatly help our project. Thanks!

Participant ID*

_____

Date*

_____

1. I think that I would like to use this tool frequently.*

Low 1 2 3 4 5 High

2. I found this tool unnecessarily complex.*

Low 1 2 3 4 5 High

3. I thought this tool was easy to use.*

Low 1 2 3 4 5 High

4. I think that I would need assistance to be able to use this tool.*

Low 1 2 3 4 5 High

5. I found the various functions in this tool were well integrated.*

Low 1 2 3 4 5 High

6. I thought there was too much inconsistency in this tool.*

Low 1 2 3 4 5 High

7. I would imagine that most people would learn to use this tool very quickly. *

Low 1 2 3 4 5 High

8. I found this tool very cumbersome/awkward to use.*

Low 1 2 3 4 5 High

9. I felt very confident using this tool. *

Low 1 2 3 4 5 High

10. I needed to learn a lot of things before I could get going with this tool.*

Low 1 2 3 4 5 High

Appendix F: User Flows

Users

1. **Go to Login page**
   1A. Put in username and password
   1B. Select role

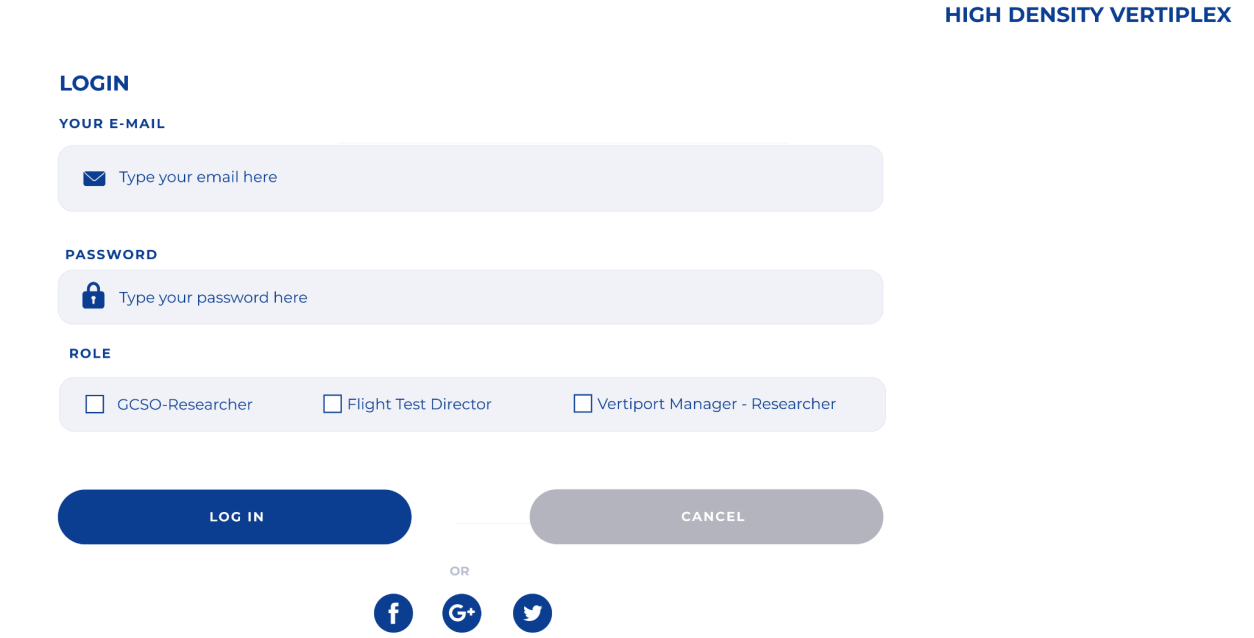
2. **Go to Home page**
   2A. Scroll down list of checklist items
   2B. Find the first one for Safety Range Officer (#21).
   2C. His/her task is to: Verify communication devices are charged

3. Preview the GCSO tasks on the My Tasks page.
   3A. Find out what needs to be done.

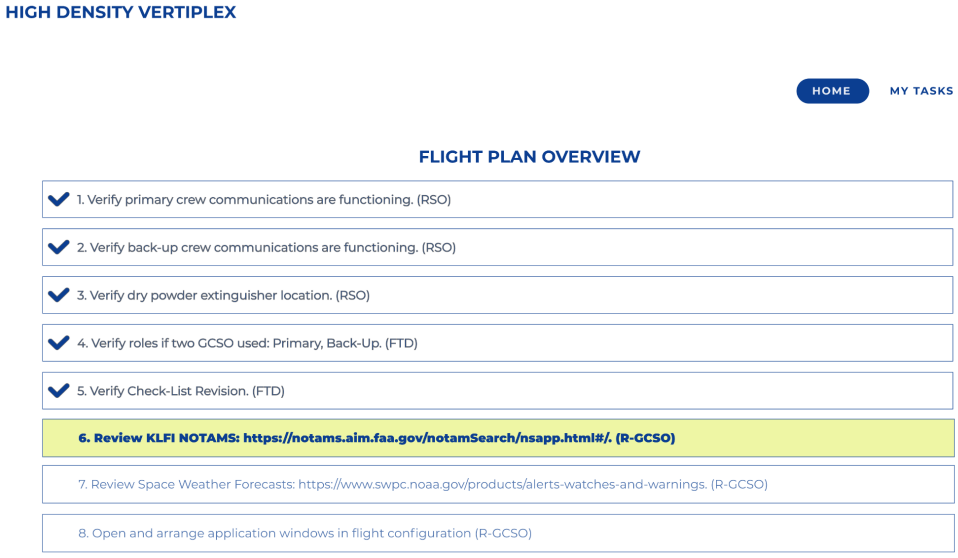
4. Preview the GCSO tasks on the My Tasks page.
   4A. Hit checkmark to show it is done.

Appendix G: Prototype screengrabs

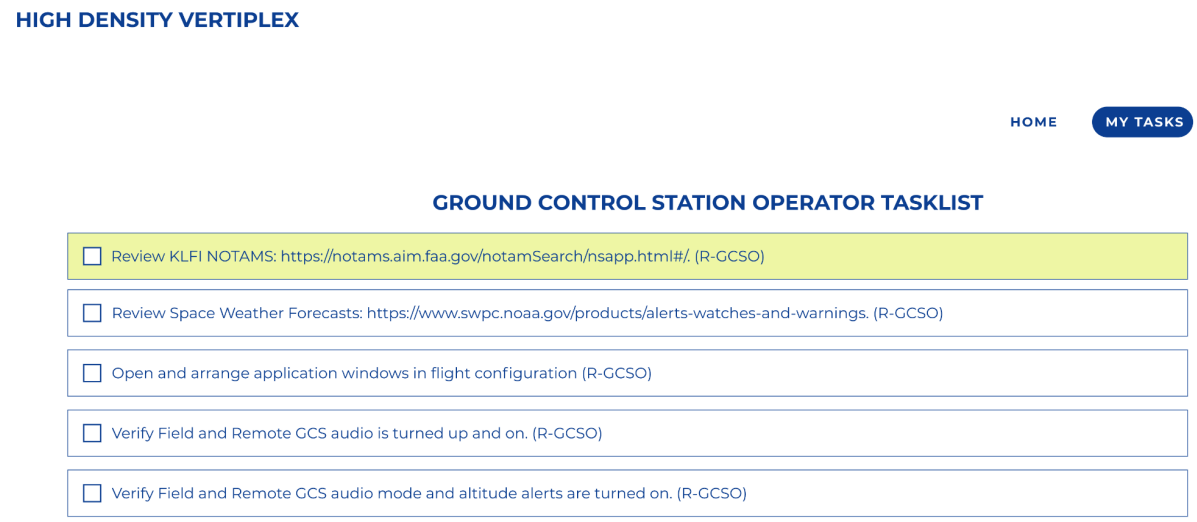
Login page



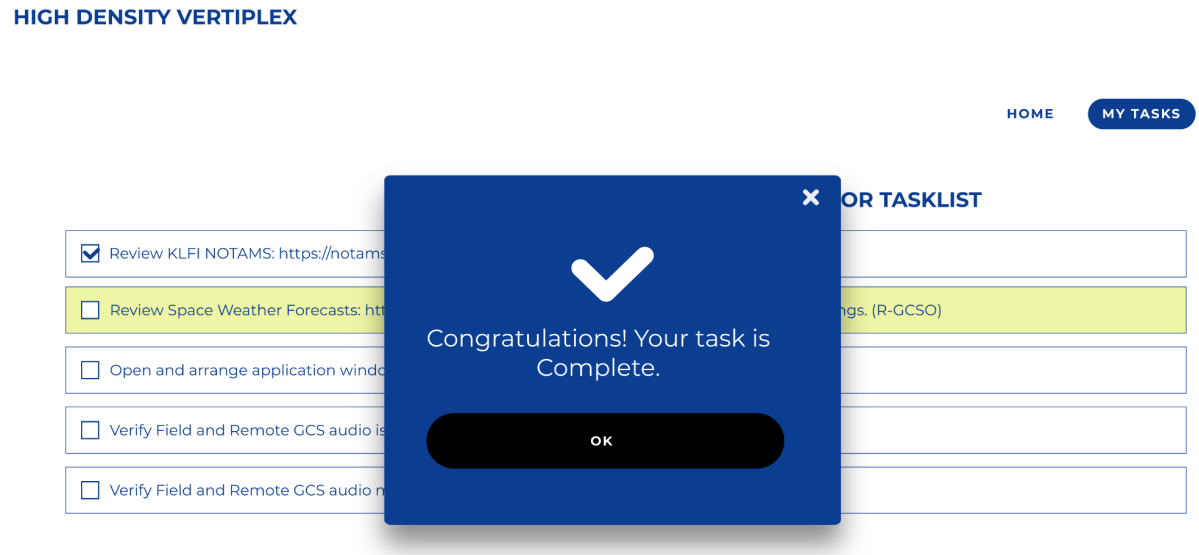
On Home Page look for Range Safety Officer task



On My Tasks Page look for GCSO list of tasks



Check off first task on GCSO's list



Appendix H: Post-task open-ended questions

Task 1
Knowledge Probe: How did you expect this interface to work?
Perception Probe: Did you expect the login display to look like this?
Q: Was there anything in the UI that confused you? **[If yes]** Explain

Task 2
Knowledge Probe: Where do you expect to see the information on the Range Safety Officer?
Perception Probe: Is this adding to your workload? Is this helping your team situation awareness?
Knowledge probe: What do you think this information can be used for?
Q: Was there anything in the UI that confused you? **[If yes]** Explain

Task 3
Knowledge Probe:What are your thoughts on finding out what you need to do?
Perception Probe: Did you expect the login display to look like this?
Q: Was there anything in the UI that confused you? **[If yes]** Explain

Task 4
Q: Was there anything in the UI that confused you? **[If yes]** Explain

Appendix I: Open-ended questions after all tasks

- How are you feeling about this tool?
- What increases/decreases to situation awareness did you experience?
- What increases/decreases to workload did you experience?

Appendix J: Post experiment open-ended questions

- What was your most/least favorite part of the tool?
- Did you think the tool was simple and straightforward or did it have a lot interrelated parts you had to keep track of?
- As far as your attention load went, were you focused on only one aspect of the tool or were you concentrating on my things at once?
- How much information were you able to get from the tool? Did you understand very little knowledge or were you able to get a lot of knowledge and use it?
- How alert were you while using the tool? Was it low or were you alert and ready to go (high)?

Appendix K: Data on Time on Task, Task Success and Task Difficulty (by participant)

P1	**Task 1**	**Task 2**	**Task 3**	**Task 4**
Time on task (seconds)	30	105	130	35
Task success	1	1	1	1
Task difficulty rating	5	4	5	4

P2	**Task 1**	**Task 2**	**Task 3**	**Task 4**
Time on task (seconds)	29	45	27	14
Task success	1	1	1	1
Task difficulty rating	5	5	4.5	4

P3	**Task 1**	**Task 2**	**Task 3**	**Task 4**
Time on task (seconds)	15	29	420	7
Task success	1	1	0	1
Task difficulty rating	4	4	3	4

P4	**Task 1**	**Task 2**	**Task 3**	**Task 4**
Time on task (seconds)	8	27	41	7
Task success	1	1	0	1
Task difficulty rating	5	5	5	4

P5	**Task 1**	**Task 2**	**Task 3**	**Task 4**
Time on task (seconds)	14	107	107	35
Task success	1	1	1	1
Task difficulty rating	5	4	3	5

P6	**Task 1**	**Task 2**	**Task 3**	**Task 4**
Time on task (seconds)	82	99	57	64
Task success	1	1	1	1
Task difficulty rating	5	4	5	3

Task Key:
Task #1: Log in as a GCSO researcher and obtain security clearance
#2: Find the upcoming task for Range Safety Officer
#3: Find upcoming task for yourself, the GCSO researcher
#4: Check off the next test for yourself, the GCSO researcher, to let the team know that step is complete.