Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
May 4, 2023
A Regional Conference of the Society for Industrial and Systems Engineering

Playbook Concepts and HMI Development

Cade Ballard, Cameron Lambeth, Pat Melampy, Erin McSoley, and Michael Kapralos

Department of Systems Engineering
United States Military Academy
West Point, NY 10996

Corresponding author's Email: cbo12gms@gmail.com

Author Note: The authors are a capstone group and cadets at the United States Military Academy at West Point. The authors are majors in the Systems Engineering department. The views expressed herein are those of the authors and do not reflect the position of the United States Military Academy, the Department of the Army, or the Department of Defense.

Abstract: This article illustrates the purpose of playbook concepts in Future Attack Reconnaissance Aircraft. As a team, we designed a playbook that was inserted into Lockheed Martin's human machine interface design for future cockpits. The purpose of the playbook is to offload cognitive requirements so that the pilot is able to effectively control the battlespace. We developed the playbook using ideas from human autonomous teaming scholarly articles to maximize efficient interaction between the human user and machine. After implementing the playbook into the interface, Lockheed Martin made a dynamic simulation to test two competing HMI concepts. We are currently doing a usability test to analyze which concept is more straight forward and easier to use.

Keywords: Human Machine Interface, Playbook

1. Background

1.1 Human Autonomous Teaming

The concept of Human Autonomous Teaming (HAT) was introduced in the 1990s. In 1991, NASA reported that "The concept of a joint human- intelligence system team [is] introduced. This team consists of an operator (onboard crew member or ground flight controller) and an intelligent system. As a team, the operator and intelligent system actively cooperate." (O'Neill et al., 2022). This idea shed light that automations and humans could interact as a team (O'Neill et al., 2022). It was not until a couple years ago that the term HAT was used to define this concept. Advancement in technologies and artificial intelligence allow the transition of HATs from being just a concept, to a used and tangible application (O'Neill et al., 2022). Because of this, research on HATs has grown exponentially (O'Neill et al., 2022). It is important to define HAT based of new research conducted in this field. O'Neill et al. define HATs as:

> "interdependence in activity and outcomes involving one or more humans and one or more autonomous agents, wherein each human and autonomous agent is recognized as a unique team member occupying a distinct role on the team, and in which the members strive to achieve a common goal as a collective. The "autonomy" aspect of human–autonomy teaming refers to the autonomous agent." (2022)

Simply put, HATs consist of humans and autonomous agents that work together to complete a mission. Many questions are raised with HATs in which we expose below using empirical research to gain a better understanding on HATs and their capabilities.

Function allocation is one of the most popular research questions within HATs. "Function allocation refers strategies for distributing system functions and tasks across people and technology" (Roth et al., 2019). Roth et al. give function allocation considerations when designing HATs based off a couple different methods. The first traditional theory is "Men are better at – Machines are better at" (MABA-MABA). Roth et al. agree that designers must consider the which agent would be better at certain tasks (2022). However, there are multiple downsides to this method. First, it is natural to place tasks on the automation agent first, leaving the human with tasks that are "leftover" (Roth et al., 2019). Designers focus more on the automation and its limits rather than the human's strength and abilities (Roth et al., 2019). A second critic is that "lists of what humans versus machines are better at can become quickly outdated as technologies continue to improve." (Roth et al., 2019). With the swiftly advancement in technology, lists must continuously be updated. The second traditional theory is Levels of Automation (LOA). A technology's LOA can be fully manual to fully automated (Roth et al., 2019). However, the drawback with LOAs is that it may generalize the team too much in which the HATs lose effectiveness (Roth et al., 2019). This method doesn't consider tradeoffs between the machine and human.

Roth et al. give new methods that should be considered when creating function allocations for HATs. HATs are most likely working in critical environments that may exceed capabilities in a constantly changing environment. Designers need to address this and create a design that can "accommodate fluid distribution of work" (Roth et al., 2019). This focuses on the application that the machine or human "could" do a task when needed or asked to complete (Roth et al., 2019). Flexibility is a key idea and arguably the most important attribute when designing function allocation for HATs (Roth et al., 2019).

While function allocation is a huge requirement in designing human automated teams, there are other questions that must be answered to optimize the effectiveness of the team. Roth et al. created an airlift asset allocation and detailed mission planning system that uses a HAT (Roth et al., 2018). The system has an automated scheduler that can schedule multiple missions while taking into account relevant information such as starting location, what to move, and latest delivery date (Roth et al., 2018). Three questions came from the experience with this new system. The first question asks what the human is likely to know that the automation system can't. In short, the human "would have access to dynamically changing, context=specific information that the automated scheduler would not be able to have access too" (Roth et al., 2018). This means that the human has better authority in directing and making changes when the schedule or mission changes. The human must be able to work with the automation to understand the change in mission so that they are on the same page quickly. The next question asked for the situations that aren't possible for the automation to work by itself (Roth et al., 2018). Simply put, it is impossible for designers to take into account every situation and code for it (Roth et al., 2018). The human would need the knowledge to create the solution without depending on the automation (Roth et al., 2018).

Similar to function allocation requirements, Roth et al. "observed the allocation of cognitive tasks is more fluid and finer-grained than traditional LOA taxonomies provided" (Roth et al., 2018). For the designer, this is called post-function allocation design decisions (Roth et al., 2018). These are decisions that arise from changes in the plan and must be made quickly while both the human and machine work together and communicate. To answer these questions, the user and automation must work together to create a solution. Roth et al. write that human automation interaction should expand "solution space" by coming up with alternative solutions that are "higher quality and more robust", in which the human nor automation could do by itself (2018). The team considered mechanisms to achieve this goal. The mechanisms are for the automation to give the user multiple solutions, which the user then conducts "what if" analysis based off the scenario and chooses the best option. A different mechanism is having the user come up with solutions then have the automation "flag" potential issues with the solution based of what the automation knows (Roth et al., 2018). A designer must consider these ideas when designing a product that allows the user and automation to communicate and complete a mission.

Human -autonomy teaming will continue to increase with the advancement in technology and research in the field (McNeese et al., 2019). In order to maximize team performance, the human and the machine must work well together to achieve the outcome. This means that they must have trust (McNeese et al., 2019). McNeese et al. conducted an experiment that focused on trust within HATs. The team used a remotely piloted aircraft system to conduct the experiment with three member teams that each had different roles in operating the system. They used a "wizard of oz" methodology to simulate the autonomous agent. This means that the team works with the autonomous agent, but that agent is really operated by an unseen human that acts as the autonomous agent (McNeese et al., 2019). The experiment had 44 participants which created 22 teams. Each team consisted of a photographer and navigator. The pilot was an experimenter, but participants were told that it was an autonomous agent (McNeese et al., 2019). To better represent the pilot as a "synthetic agent", the two-man team communicated and coordinated with the pilot with limited vocabulary (McNeese et al., 2019). McNeese et al., measured team performance and team trust to understand the team dynamic between human-human and human-autonomous teams (2019).

1.2 Cognitive Offloading and FARA Requirements

With the new Future Attack Reconnaissance Aircraft program, a key concern arises. There is a concern that pilots using FARA aircraft will not be able to cognitively process everything within the battlespace due to the advancement in technology and the new requirements that come with this new aircraft. Our goal with this project is to create an HMI interface that will allow the pilots to offload information and work with an automated agent to complete complex tasks in a critical environment. In the beginning of the semester, we read multiple scholarly articles about cognitive offloading and cognitive requirements for FARA. The main article we focused on was *Cognitive Requirements for FARA: Pilot Aiding in a Complex Decision Space* by Ernst et. al. The article went through cognitive task analysis and cognitive requirements for FARA pilots. Ernst. Et al wrote that:

FARA pilots will need to balance (a) searching an area of interest via direct visualization with (b) assessing an area distant from their location with a variety of sensing modalities. Pilots will need to maintain an awareness of the terrain, airspace, and threats in the immediate vicinity of their aircraft while performing similar functions for the UAS operating distantly.

Furthermore, as the reconnaissance mission unfolds, the FARA pilots will be responsible for making decisions regarding asset allocation and re-planning in response to the environment, adversary actions, and friendly operations. (Ernst et al, 2019)

FARA pilots are required to be proficient in more tasks compared to pilots today. Because of this, cognitive offloading is required so that the pilots don't feel overloaded on the battlefield. The article found that FARA pilots will require similar cognitive abilities to current Tactical Controllers (ITC) since they are essentially taking the place as mission managers in the aircraft. The mission manager will control up to 8 Air Launched Effects (ALEs) will the pilot maintains control of the aircraft. The mission manager uses these ALEs to conduct reconnaissance and even destroy enemy targets. They are deployed from the aircraft and are sent ahead of the FARA. With this new technology, pilots are able to have a better sense of the battlefield to make calculated decisions while processing more information than before.

2. MethodologyPlaybook Concept and HMI Development

The project began by conceptualizing what a playbook might look like within Lockheed Martin's current competing HMI designs. The first HMI design consisted of a map with an al le carte menu which allows the mission manager to select and move around the tools and functions needed at any given time. The al le carte HMI design also possesses a radial menu for when the mission manager clicks the screen. This radial menu is intended to be a quick method of accessing useful functions and tools. The second HMI design is a map with a northward-facing miniature map in the bottom right corner. This second design possesses a fixed menu on the right side of the display. This menu possesses the same functions and tools as the first HMI design. The project team was first tasked with conceptualizing what and where a playbook would be within these two competing HMI design concepts. The first step the team took was to determine how the mission manager might interact with the playbook within the respective HMI design. This led to the team deciding to make the playbook accessible on the radial menu in the first HMI design concept. The second design concept has the playbook feature in the fixed menu in the upper right-hand corner of the schedule planner. After determining the location of the playbook within the competing HMI design concepts the next task was to determine what the playbook would look like when selected by the mission manager. The project team determined this by grouping the different types of plays or mission sets the ALEs could conduct. The different play types identified were intelligence, surveillance, reconnaissance, and attack plays. Within each of the play types, there are differing plays that follow the mission set of each play type. When the mission planner calls a play, they are prompted to input variables such as objective/target, ALEs to select, the route to the target/objective, and when the ALE should arrive. Both HMI concepts require the same input variables to call a play. Once the input variables are set, the mission planner selects enter which now calls the play for the ALEs.

2.2 Testing Plan

At the beginning of the year our group was tasked with creating a test plan and methodology given very few guidelines on what data points Lockheed Martin wanted. Originally the plan was to test subjects on cognitive load alone and see how effective they could be under stress attempting to complete a mission set. There were initially two different mission sets that would be tested. As the semester progressed it was determined that it was not feasible to create a simulation with two full scenarios, so Lockheed put us in charge of creating a playbook for one mission subset. After this guidance, the testing plan was written.

Vague guidance stunted the growth of the testing plan due to the foggy scope of the project. The group decided to split the testing plan up into different sections so it would satisfy Lockheed's expectations. The different sections are testing, population, methodology, end state, and schedule. This allows the testing plan to lay out a detailed plan of action. It will help organize how the testing will be completed in a professional manner. The rationale behind this plan is to summarize exactly how the testing will be completed, why the testing is being done, and when the testing will be done.

The testing plan was initially developed to test the complexity of multiple ALEs in one environment. Testing included different stages of work that would test the cognitive load of the test subject. The significance of testing this way is that it will produce important data points for Lockheed Martin to continue their development of their HMI project. The West Point population is who will be tested. This is good because it is a diverse set of people that can give thoughtful feedback on the system. There are many aviators in the populous at West Point that can be tested along with civilians and military personnel that are not aviators.

After initial feedback from Lockheed, it was determined that the testing plan would need to be altered from its original form. One of the major improvements implemented is that of a new scenario in which two HMI concepts will be tested on the test subject. Lockheed wants feedback on two different HMI concepts. Originally this was not part of the testing priorities, however, when the stakeholder set these new conditions the testing plan needed to be updated. Additionally, they wanted a more narrowly tailored feedback survey at the beginning and end of testing to help strengthen their data points. For example, they would like the testing subjects to be asked if they liked one of the HMI concepts the most, and why? They were not as excited about the different ALEs being introduced in stages. The group is making a judgement call to keep this portion in the testing. It is to help enhance the data retrieval pertaining to the cognitive load portion of the testing.

In addition to the test plan, the group also created pre-and-post confidence surveys that are intended to be given to the testing population. These surveys hope to determine the population's confidence level before the test and to determine how confident they'd feel using the playbook again. The reason for collecting the pre-and-posttest confidence levels for the population is to determine if the playbook supports cognitive offloading for the mission planner. An increase in confidence after the test indicates that the playbook interface and the system adequately prepare the mission planner to direct ALEs to accomplish a given mission. Conversely, a decrease in confidence after the test indicates that the playbook interface and the system are not intuitive. This means that further refinement, improvements, and changes are needed to better support cognitive offloading for the mission planner. The pre-and-post confidence surveys were sent over to Lockheed Martin which they liked and showed to a Army research group conducting tests on a similar project. The group plans to refine these confidence surveys as the system gets created and a more complete picture on how the mission planner will interact with the system is created. Finally, to shed further light on the way forward, we plan on initializing testing during the month of March. Lockheed has given a hard date of the second week in February to have the software needed set up and operational.

3. Test and Data Analysis

The test conduct consist of three data sources, FaceX data, the HMI concepts data, and intermediate surveys through the testing process. The FaceX data is collected from six cameras that are attached to the test monitor. The data collected is eyetracking and gaze datapoint, stored in binary. The HMI concept data comes from the interaction between the participant and the two competing concepts, concept A and concept B. The test begins with the completion of two surveys that capture the participants military history and confidence in using the technology. The test then continues with the calibration of the FaceX camera. The cameras are calibrated to the specific participants gaze. This is accomplished through a calibration program which takes around 5 minutes. Once the cameras are calibrated and the data is recording, the participants moves to the HMI concepts and begins concept A, task 1. The task involve recon conducted by an ALE that reports and unknown enemy target. The participant must reassign the ALE a new task by using the playbook menu to call a ground reconnaissance play to expand the search for additional enemy targets. The second task of concept A is a continuation of the reconnaissance being conducted. The ALE reports a high value target which meets the engagement criteria. The participant must call an attack play from the playbook panel. Once this play is completed, concept A is complete. The participant must fill out two confidence surveys.

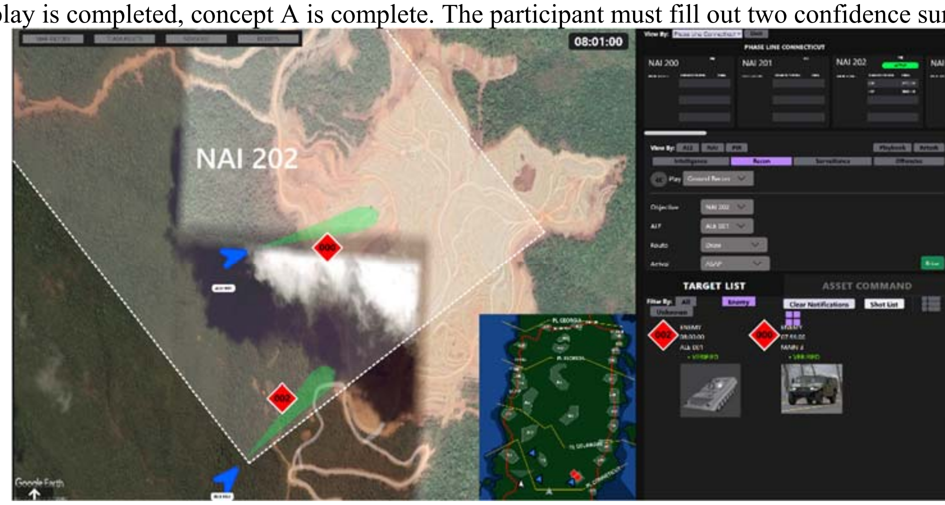

Figure 1. Concept A

For Concept B, the target panel is set up in the opposite side, and the mini map is larger on the right side. The playbook menu is displayed once the participant clicked anywhere on the screen. The menu is a radial menu that has the difference types of plays around the circle. Once the desired type of play is chosen the radial menu shifts to new options that consist of the object name, the timing of the call, the ALE tasked, and the path taken. Once all of the criteria of the mission is properly inputted the play can be saved and executed. The participants begins the mission when 3-4 allied targets are detected by the

ISBN: 97819384964-4-8

ALE. The search for more allied targets is expanded by calling a new play using the radial menu. The play calls an air surveillance pay with 4 ALEs searching in a segmented path. Once completed the participant moves onto task 2. Task 2 begins in a new OA where the ALEs detect a high payoff target that meets the engagement criteria. The participants must use the radial menu once again to call a new play to attack the detected targets with one ALE. This concludes concept B and the participant must fill out the remining 4 surveys.

Figure 2. Concept B

The data analysis of this test will be led by the Lockheed Martin analyzers after the conclusion of our testing. This project calls for 10 test subjects in order to get an array of different test subject that have different cognitive experiences. The analysis will ideally give our team and our partner Lockheed Martin team feedback in the useability of the different HMI concepts.

4. References

D. J. McAllister, "Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations," *ACAD MANAGE J*, vol. 38, no. 1, pp. 24–59, Feb. 1995.

Ernst et al. Cognitive Requirements for FARA: Pilot Aiding in a Complex Decision Space, December 31, 2019.

McNeese, N., Demir, M., Chiou, E., Cooke, N., & Yanikian, G. (2019). *Understanding the role of trust in human-autonomy teaming*. https://doi.org/10.24251/hicss.2019.032

O'Neill, T., McNeese, N., Barron, A., & Schelble, B. (2022). Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. Human Factors, 64(5), 904–938. https://doi.org/10.1177/0018720820960865

Roth, E., Depass, B., Harter, J., Scott, R., & Wampler, J. (2018). Beyond Levels of Automation: Developing More Detailed Guidance for Human Automation Interaction Design. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 62(1), 150–154. https://doi.org/10.1177/1541931218621035

Roth, E. M., Sushereba, C., Militello, L. G., Diiulio, J., & Ernst, K. (2019). Function Allocation Considerations in the Era of Human Autonomy Teaming. Journal of Cognitive Engineering and Decision Making, 13(4), 199–220. https://doi.org/10.1177/1555343419878038