# Foundational Human-Autonomy Teaming Research and Development in Scalable Remotely Operated Advanced Air Mobility Operations: Research Model and Initial Work

Eric T. Chancey[1], Michael S. Politowicz[2], Bill K. Buck[3],
Kathryn M. Ballard[4], Vincent E. Houston[5], and Lisa R. Le Vie[6]
*NASA Langley Research Center, Hampton, VA 23681, USA*

James R. Unverricht[7]
*National Institute of Aerospace, Hampton, VA 23666, USA*

Meghan Chandarana[8]
*NASA Ames Research Center, Moffett Field, CA 94035, USA*

**To achieve the scalability envisioned for many Advanced Air Mobility (AAM) applications, uncrewed aerial system (UAS) concepts are being pursued with the goal of enabling fewer human operators to manage more increasingly autonomous vehicles. NASA's Transformational Tools and Technologies – Revolutionary Aviation Mobility ($T^3$-RAM) subproject has identified human-autonomy teaming (HAT) as a critical area of research required to support these operations. Under $T^3$-RAM, the HAT Foundational Research Activity has been tasked with providing basic research to identify HAT and human-automation interaction (HAI) principles that can be used to achieve scalable multi-vehicle UAS operations. This paper first outlines a research model to produce ecologically relevant basic research, then contextualizes completed and planned research and development activities within this model. Proposed research threads are presented, along with their practical and theoretical implications.**

## I. Introduction

The vision of Advanced Air Mobility (AAM) seeks to introduce a range of missions that will expand the capacity to transport people and goods to locations in rural and urban environments. Using aircraft of all sizes, this vision will be achieved in part by leveraging higher levels of automation and increasingly autonomous technologies, with applications ranging from commercial transport and air-taxi services (e.g., Urban Air Mobility [UAM]) to drone surveillance and inspection operations [1]. Yet, an expansion of air transportation services will also exacerbate (or be limited by) an approaching shortage of qualified pilots to support operational demands (see Ref. [2] for an early recognition of this in UAM; see Ref. [3] for a discussion on the factors related to the pilot supply-demand gap). To address this challenge, and promote scalability, remotely operated uncrewed aerial system (UAS) concepts are being considered [4, 5], with the goal of enabling fewer human operators to manage more increasingly autonomous vehicles

---

[1] HAT Foundational Research Activity Lead, Crew Systems and Aviation Operations Branch (CSAOB)
[2] HAT Foundational Research Activity Co-Lead, CSAOB, AIAA Member
[3] Aerospace Engineer, CSAOB, AIAA Senior Member
[4] Statistician, Engineering Integration Branch
[5] Senior Machine Intelligence and Data Systems Researcher, CSAOB, AIAA Senior Member
[6] NASA Langley Research Center Point of Contact for $T^3$ Autonomous Systems Enduring Discipline, AIAA Member
[7] Human Factors Psychologist, CSAOB
[8] Computer Engineer, Intelligent Systems Division, AIAA Member

(i.e., *m* humans-to-*N* vehicles, or *m:N* [6]). NASA's Transformational Tools and Technologies – Revolutionary Aviation Mobility ($T^3$-RAM) subproject has identified human-autonomy teaming (HAT) as a critical area of research that could be used to enable these envisioned types of operations [5]. We define the term HAT as a distinguishable set of two or more agents (human and machine) that interact dynamically, interdependently, and adaptively toward a common goal/objective/mission (definition adapted from Ref. [7, p. 7]). The concept of a human 'teaming' with a technology shifts the assumptions of traditional human-automation interaction (HAI) paradigms, as increasingly autonomous technology is required to assume many of the responsibilities and authorities traditionally held by humans within the overall system. Established HAI research, however, still provides a useful framework to understand and contextualize how humans and technology may be appropriately paired. Under $T^3$-RAM, the HAT Foundational Research Activity has been tasked with producing basic research (i.e., controlled laboratory studies, theory-development) to identify HAT and HAI principles applicable to achieving scalable multi-vehicle (m:N) UAS operations. In this paper, we first outline a model to produce ecologically relevant basic research. We then frame completed and planned research and development activities within the context of this model. This framing includes observational studies conducted during the development of a remote UAS operations center and ground control station (GCS) software and how that work has informed our theoretical perspective on HAT and HAI. The paper concludes by describing the development of a software application designed to conduct basic research with the goal of investigating increasingly autonomous multi-vehicle UAS operations.

## II. Grounding Basic Research in Practical Application

The overarching goal of the HAT Foundational Research Activity is to identify approaches that enable fewer remote operators to manage more increasingly autonomous vehicles, particularly those approaches that may be applied to the concepts being explored in AAM operations. This is a relatively broad goal, which is not constrained to a specific use case or concept of operation within the wider scope of AAM mission subsets. To accomplish this goal, the focus of this activity is to produce basic research that generalizes across specific applications. Basic research relies on generating theory and competing theories to determine the most accurate account of a topic under investigation. Frequently, basic research is conducted with controlled, randomized human-in-the-loop (HITL) experiments in which variables are independently manipulated or held constant to cleanly investigate the effects on human performance metrics, responses, and reactions. In this type of research, to maintain a high level of control, the experimental task is simplified by omitting factors present in operational settings. Often, this approach is taken to maximize internal validity, which concerns whether the relationship between two variables is causal in nature (i.e., an experimental manipulation causally affects an outcome measure [8]). Many factors in a laboratory environment may need to be controlled to attribute causality to an effect observed in support of a particular theory, but it can be difficult to identify *all* of the causally deterministic factors and how they relate to each other (this type of causal effect may be more accurately referred to as an *inus* condition, or insufficient but non-redundant part of an unnecessary but sufficient condition [8]). As a result, basic research can be susceptible to studying 'toy problems,' in which, ironically, the findings produced under such tightly controlled settings generalize to only that specific setting and are not universal findings as proposed [9, 10]. From this perspective, causal relationships are context-dependent, and the generalization of experimental effects can be at issue in any experimental setting. Indeed, although the strength of controlled experimentation is its ability to illuminate plausible causal inferences, the weakness of controlled experimentation is doubt about the extent to which that causal relationship generalizes beyond that specific experiment [8]. This points to the concept of external validity, or inferences about whether an effect holds over variation in persons, settings, treatments (e.g., experimental manipulations), and measures. Often, internal and external validity are discussed in ways that associate internal validity with controlled laboratory settings and external validity with field studies, yet the two are logically congruent (e.g., causal inferences from the field can be internally valid, and a researcher may wonder if findings from a field study generalize to a laboratory setting [a matter of external validity]; see Ref. [8]). Moreover, external validity is sometimes confused with ecological validity. Ecological validity is less a type of validity and more of an approach that calls for research that samples from settings and participants that reflect the 'ecology' of the application from which the effect is being studied (e.g., a study with pilot participants in a high-fidelity flight simulator has good ecological validity for relating to real-world flight operations [8]). Although the goal of the HAT Foundational Research Activity is to pursue basic, foundational research, it needs to generalize to ecologically relevant practical applications related to solving real-world problems (i.e., scalable m:N UAS operations). To support this, we adopt the research model outlined by Vicente as proposed in Ref. [10].

The model outlined by Vicente separates research into four broad types, which include highly controlled laboratory experiments (Type 1), less controlled but more complex laboratory experiments (Type 2), evaluations conducted in high-fidelity simulators or in the field (Type 3), and qualitative or descriptive field studies (Type 4). Table 1 provides a description of each type of research. Vicente offers a 'caricature,' however, of what he perceives is the 'traditional perspective' and provides a description of each. Type 1 research is often viewed as 'real' science and the only reliable method to discover pure, context-independent, and generalizable principles. Type 2 research is viewed as an inferior version of Type 1 research, in that not all variables are controlled, and effects may be confounded with those uncontrolled variables. Type 3 research findings are less coherent and interpretable than those in Type 2, as the task environment is too complicated to yield generalizable findings. Finally, Type 4 research is not considered scientific because no variables are held constant or manipulated, and therefore, hard quantitative data, amenable to inferential statistical analysis, cannot be usefully interpreted. Providing a critical analysis of this caricature of the traditional perspective, Vicente argues that if researchers prioritize Type 1 research at the expense of ignoring the other types and rarely or never travel into the field, then the findings of these studies will be based on *speculations* of the nature of the demands humans encounter in the real-world settings they are seeking to generalize to (i.e., issues of ecological and external validity).

**Table 1: Vicente's Research types and Descriptions**

| Research Type | Description |
|---|---|
| **Type 1: Highly controlled laboratory experiments** | Studies use simplified tasks in which variables are independently manipulated or held constant to investigate factors related to the phenomenon under study. Useful for competing theories to determine which provides a more correct explanation of the topic being studied. |
| **Type 2: Less controlled but more complex laboratory experiments** | Studies use tasks that are more complex than Type 1, losing experimental control to represent an operational setting and produce results that will better generalize to that operational setting more accurately. |
| **Type 3: Evaluations conducted in high-fidelity simulators or in the field** | Studies use tasks that are highly representative to the real-world environment and have less experimental control than Type 1 or 2 research. Useful to determine if the results obtained in Types 1 and 2 research overcome the myriad of additional factors that are not held constant. |
| **Type 4: Qualitative, descriptive field studies** | Field studies provide the ability to observe and document the pressing, significant issues that require research to understand. Also, field studies provide the basis to determine which factors should be manipulated in laboratory experiments (Type 1) and identifies which measures should be adopted to evaluate performance. |

These four types of research serve alternative purposes, which are complementary, and should be used to inform each other. Supporting this, Vicente proposes a model that links the different types of research in a manner that logically capitalizes on the strengths of each (Fig. 1). Instead of prioritizing only Type 1 research, the model suggests that basic research (Type 1), readily applicable to practical issues, should *begin* with observations in the real-world environment (Type 4) that the basic research is attempting to generalize to. Specifically, findings from Type 4 research can be used to meaningfully prioritize the topics and factors that should be experimentally investigated, as they pertain to operational problems and issues. This approach provides a greater chance to avoid producing basic research that is pursuant to solving problems that have no foundation in the real world. The model provides an ecologically sensitive approach to conduct defensible Type 1 research and encourages theory development that generalizes outside of the laboratory. As stated by Vicente, "If one does not conduct Type 1 research, then it is difficult to build theories that have solid empirical support; and without defensible theory, generalization from one application domain to another would be tenuous" [10, p. 327]. The following section provides a high-level overview of the Type 4 research that the HAT Foundational Research Activity has pursued and provides some initial overarching elements that will be investigated in Type 1 research activities.

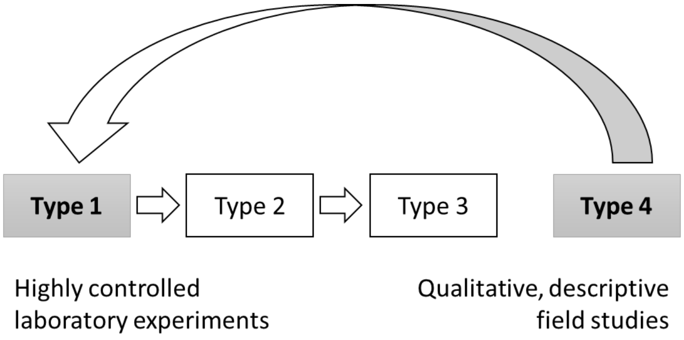

**Fig. 1 A Model for Ecological, Generalizable Foundational Research. Adapted from Ref. [10, p. 326].**

## III. Scalable UAS Operations Research and Development Activities

As previously stated, the HAT Foundational Research Activity has been tasked with producing Type 1 research to identify and apply HAT and HAI principles within the context of achieving scalable multi-vehicle (m:N) UAS operations. Yet to date, this activity has largely focused on producing Type 4 research to provide a real-world foundation for future Type 1 research. The following sections provide a more detailed outline of the Type 4 and Type 1 research and development efforts conducted by this activity as they align with the perspective outlined in Section II.

### A. Type 4: Observational Research and Testbed Development

To identify and address the challenges ahead for *realistically* scalable remote UAS operations, the HAT Foundational Research Activity has partnered with NASA's High Density Vertiplex (HDV) subproject under the AAM project. The HDV subproject is responsible for developing and testing technologies, concepts, and architectures that support the infrastructure needed for terminal environments around vertiports (i.e., ground or elevated areas used for the vertical takeoff and landing of aircraft in many AAM operations [11]). A central capability enabling this work is developing and standing up the ROAM (Remote Operations for Autonomous Missions) UAS Operations Center, a facility used to remotely command and control simulated and live vehicles [12]. The HAT Foundational Research Activity has worked closely with HDV to provide human factors research and design support with the specific purpose of exploring and advancing remotely operated UAS concepts in this environment.

Located at NASA Langley Research Center (LaRC), the buildup of ROAM has been an evolutionary process which began with a simple workstation setup that allowed an 'operator' to monitor small uncrewed aerial system (sUAS) telemetry data (see Ref. [13] for initial human factors assessment in ROAM). This early instantiation provided a testbed to identify the needs and capabilities required to transition this facility into a remote sUAS command-and-control environment. Much of this initial transition relied on interviewing ground control station operators (GCSOs) and flight crews supporting various sUAS research activities conducted at the LaRC CERTAIN (City Environment for Range Testing Autonomous Integrated Navigation) flight range [14]. During this buildup, the HAT Foundational Research Activity designed and delivered the MPATH (Measuring Performance for Autonomy Teaming with Humans) ground control station (GCS), which is the primary software and interface component used at the GCSO position in ROAM [15]. Modified from the open-source QGroundControl software used to control Micro Air Vehicle Link (MAVLink)-enabled sUAS [16], the MPATH interface facilitates onboard automation transparency and improves upon baseline usability through a modular 'widget design' that increases the spatial proximity of related information. Importantly, to support ongoing human factors data collection, MPATH logs user interaction data (e.g., button clicks and associated time stamps) and dynamic positions of both areas of interest (i.e., regions on the MPATH display) and objects of interest (e.g., the dynamic movement of a specific vehicle icon) for post-hoc association with eye movement data.

In addition to interviews with subject matter-experts, we have conducted both simulated [17] and extended visual line of sight (EVLOS)* [18] human-in-the-loop observational studies. Lessons learned from these Type 4 research activities have contributed to more confident transitions in the complexity of ROAM operations, such as three GCSOs simultaneously controlling three live sUAS (i.e., multiple concurrent 1:1 human-to-vehicle ratios) and, more recently, a demonstration in which a safety pilot (located at the CERTAIN range) handed off a live vehicle to a GCSO simultaneously controlling a simulated vehicle from ROAM (i.e., 1:2 [1 live, 1 simulated] human-to-vehicle ratio). Ultimately, the goal of ROAM is to enable a crew of human operators to remotely manage and control multiple highly automated sUAS under beyond visual line of sight (BVLOS) conditions. Adhering to the research model outlined in Ref [10], the primary objectives of these Type 4 research activities were twofold: 1) Conduct human factors and usability assessments that uncover as many workstation, display element, and operational issues as possible (primarily used to support immediate recommendations and improvements), and 2) Generate hypotheses from observations made during these studies, which will be tested in future Type 1 laboratory experiments. In addition to supporting the buildup of ROAM [12], developing the MPATH GCS [15], and conducting HITL studies [17, 18] we have identified several Type 1 research avenues to pursue (see Table 2 for a selected list). The next section describes emerging HAT theory, as well as existing HAI perspectives, in the context of the Type 4 research activities outlined in section III-A. This section concludes with a brief description of a new Type 1 research testbed capability designed to study HAT and HAI issues relevant to scalable multi-vehicle UAS operations.

---

[*] Under EVLOS conditions, the GCSO relies on trained observers in the field to keep the vehicle within their visual line of sight (see Ref. [38] for description of an example EVLOS arrangement).

**Table 2: Selected Type 4 Issues Identified for Type 1 Research Threads**

| Issues Identified | Description |
|---|---|
| **Automation Transparency** | Automation transparency was a consistent theme that emerged during Type 4 research activities. Human-to-automation transparency is the information that automation needs to share with the human about its perspective on the task it is completing, how it is completing it, and awareness of its intentions [19, p. 41]. Transparency is a critical design factor to support appropriate mental models, trust, and subsequent use of automation [20]. Transparency issues were evident during instances in which the detect and avoid aid would take control of the vehicle and provide a resolution that was not communicated to the GCSO. |
| **Switching Between Vehicles (Multi-Vehicle Operations)** | During a demonstration in which a single GCSO was simultaneously responsible for a live vehicle and a simulated vehicle from a single GCS, it became apparent that it is crucial to ensure (technologically, procedurally) that the operator is aware of which vehicle is under control before any commands are issued. Multi-vehicle control paradigms in which an operator is switching among multiple vehicles needs further investigation to discover the nature and types of errors likely to be committed. |
| **Nuisance and False Alarms** | An excessive number of alerts was observed and commented on by the GCSOs during live and simulated operations in ROAM within the GCS software. Nuisance alarms occur when the response bias of a signaling system is set too liberally (i.e., minimal evidence is required to signal an event; e.g., Ref. [21]). Frequently receiving and dismissing alarms can lead to instances in which important information is overlooked, sometimes referred to as 'alarm fatigue.' Similarly, a liberal response bias may be selected to maximize the chance of signaling an abnormal event, at the expense of generating many false alarms [22]. Referred to as the 'cry-wolf' effect, false alarms tend to be particularly damaging to trust in sensor-based signaling systems, which can also lead to reduced system compliance and slowed response times [23]. |
| **Proximity Compatibility Principle** | Because ROAM is a research testbed, new information sources are often being integrated and displayed. This integration has sometimes come at the expense of violating the proximity compatibility principle, where tasks that are highly relevant should have display elements that are proximally close to prevent dividing attention [24]. For example, in the HITL studies described in Ref. [17, 18], the GCSO had to integrate information from separate map interfaces across two monitors. This setup can lead to the visual attention failure of change blindness, or failing to notice that something is different from what it was [25]. The HAT Foundational Research Activity is currently developing a machine learning-based solution to combat change blindness in ROAM [26]. |
| **Displaying Vertical Position** | MPATH has received relatively high usability ratings [15], yet displaying vertical position has been identified as an issue (particularly as it relates to providing contextual distance from other vehicles and terrain). Several user interface designs have been proposed, such as providing a cross-sectional view that shows the ownship's vertical relationship between the ground and other vehicles. Recently, during the 1:2 demonstration in ROAM, a supplemental prototype 3D display was available to the GCSO, which was anecdotally used "about as much as MPATH" (GCSO comment). Though 3D displays may be a viable option in future ROAM operations, the concept of *naïve realism* has shown that although users prefer these types of displays, generally simpler displays lead to better performance [27]. |
| **Shared Situation Awareness between Field and Operations Center Groups** | The larger flight operations team consisted of at least two geographically separated groups: the ROAM group and the CERTAIN flight range group (and sometimes a third group located at NASA Ames Research Center in California). Team situation awareness (SA) [28] needs to be explored in these types of geographically distributed environments. Specifically, issues related to establishing the roles and responsibilities among the groups and group members, standardized communication channels and language, and shared mental models among group members were identified during Type 4 research activities. |

### B. Type 1: Foundational Research and Testbed Development

To promote scalability, many of the remote UAS operations envisioned within AAM will be supported by increasingly autonomous systems, which will progressively shift responsibility and authority toward the technology and away from the human [29]. The operations conducted in ROAM are also following this path, and onboard automation such as detect and avoid, crash management, and autopilot systems have been leveraged to begin to enable fewer GCSOs to manage more vehicles in this environment. Beyond the specific issues identified in Table 2, more generalized findings have emerged with implications for HAT theory development in AAM operations. The next section contextualizes Type 4 findings in emerging HAT theory and existing HAI perspectives.

### 1. *A Theoretical Perspective on HAT in Scalable Remote Operations*

Much as our definition of HAT is borrowed from the interpersonal teaming literature, we have also leveraged interpersonal teaming theoretical models. The model in Fig. 2 is adapted from Ref. [30], which outlines the "Big Five" components of *teamwork*, or the "set of interrelated thoughts, actions, and feelings of each team member that are needed to function as a team and that combine to facilitate coordinated, adaptive performance and task objectives resulting in value-added outcomes" [30, p. 562]. These components promote team effectiveness, where the coordinating mechanisms of trust and shared mental models are required to facilitate these teamwork components (see Ref. [20] for detailed overview of the relationship between trust and mental models developed by the HAT Foundational Research Activity). Although we have adapted this model for HAT, the relationships among the variables in Fig. 2 are still under investigation, as it is plausible that interpersonal (human) teams likely differ in important ways that may not easily translate to HAT. Indeed, from the perspective outlined in Ref. [30], teams do not simply engage in *taskwork*, or "interactions with tasks, tools, machines, and systems" [31, p. 90]. This clarification points to similar differences between HAI paradigms (e.g., [32]) and emerging HAT concepts, with the focus being on the difference between automation and autonomous systems.

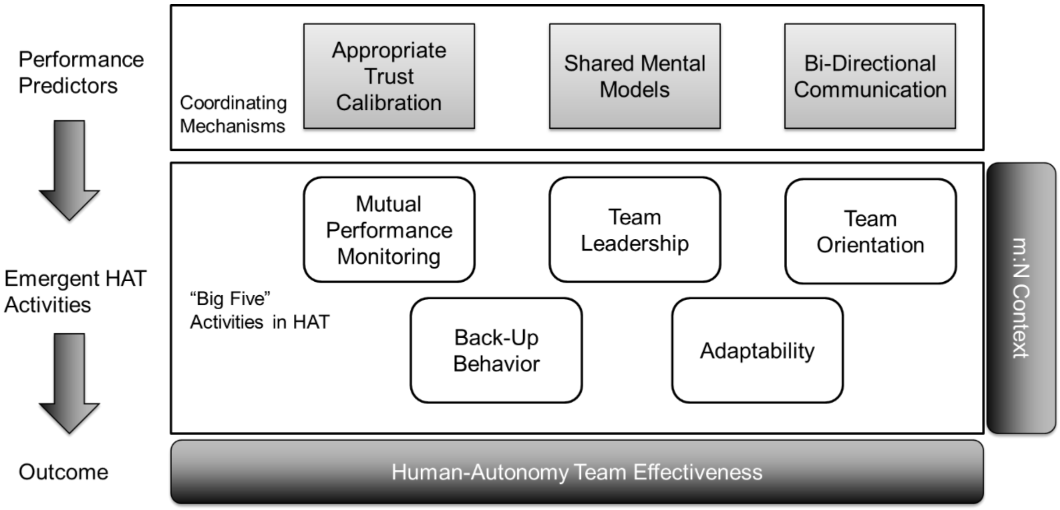

**Fig. 2 Preliminary Representation of a Proposed HAT Framework. Adapted framework from Ref. [30].**

The term 'increasingly autonomous' is frequently invoked to describe the technology that will be leveraged in AAM operations [29], which spans "the spectrum of system capabilities that begin with the abilities of current automatic systems, such as autopiloted and remotely piloted (non-autonomous) unmanned aircraft, and progress toward the highly sophisticated systems that would be needed to enable the extreme cases" [33, p. 2]. Defining technology in this way implies that these operations will be supported by both automation and autonomy, and yet the theoretical implications for how to approach these two forms of technology vary, and a distinction needs to be made. Kaber (see Ref. [34, p. 419]) provides a useful conceptualization of autonomy based on three tenets: 1) If a system is not hardened to an environment, it is not viable or autonomous; 2) If a system requires monitoring (and the possibility of human intervention), it is not independent or autonomous; and 3) If a system is not responsible for mission goals or control of resources, it is not 'self-governing' or autonomous. Clearly, based on these tenets, the role of the GCSO in ROAM is to provide an oversight function for the vehicles operating at the flight range, indicating automation and not autonomy is supporting these operations. This clarification is not intended to indicate that we are not able to gain insight into HAT from the Type 4 research conducted in ROAM. Instead, we are simply acknowledging that because AAM operations will rely upon 'increasingly autonomous' systems to achieve scalability, much of the important work conducted in HAI paradigms is still relevant and can provide useful theoretical and empirical touchpoints to guide the transition to a HAT paradigm shift. Consequently, the planned Type 1 research that will be pursued by the HAT Foundational Research Activity will address both HAI- and HAT-related concepts to help provide ecologically relevant factors to achieve scalable remote vehicle operations. To accomplish this, we have developed a new Type 1 research testbed, the Human-Autonomy Teaming Task Battery (HATTB), which will support experimental investigations in HAT and HAI paradigms.

### *2. HATTB: A Type 1 Research Testbed*

Targeted for eventual public release, the HATTB is a software application designed to create and conduct controlled HITL experiments with the purpose of offering the researcher the ability to systematically manipulate and measure variables pertinent to multi-vehicle, HAT/HAI-enabled operations (Fig. 3). The design philosophy of the HATTB follows that of the Multi-Attribute Task Battery [35, 36], which is a programmable battery of tasks that simulate pilot responsibilities during flight. Specifically, the goals of the HATTB are to provide: 1) An easily accessible tool for many researchers to study scalable multi-vehicle UAS operations, 2) An ecologically relevant testbed environment that consists of tasks abstract enough to be used by non-expert participants, and 3) A tool that encourages and enables researchers to replicate the findings from studies using a common testbed environment. This approach was taken to accelerate Type 1 research investigating HAT and HAI paradigms, which may be used to achieve scalable multi-vehicle UAS operations (i.e., the quantity of HATTB studies addressing these topics will not be limited to NASA researchers and resources supporting individual efforts).

**Fig. 3 Screenshots of the HATTB.**

Acknowledging the need to study multi-vehicle management in HAI and HAT paradigms, the HATTB was designed around the stages and levels of automation framework outlined by Parasuraman, Sheridan, and Wickens (see Ref. [37]). Parasuraman et al. define automation as “a device or system that accomplishes (partially or fully) a function that was previously, or conceivably could be, carried out (partially or fully) by a human operator” [37, p. 287]. Although this definition specifically references automation, the use of “partially or fully” encompasses the notion of levels of automation (LOAs) that can range from fully manual to fully autonomous. Moreover, this framework emphasizes the role of the human in determining overall system performance (i.e., “functions previously carried out by a human”), and maps LOA onto a simplified version of human information processing stages that cover input functions (i.e., information acquisition, perception and working memory) and output functions (i.e., decision making, action implementation). The purpose of adopting this framework to guide the design of the HATTB was to provide a method to structure HAI and HAT experimental paradigms.

Variables of interest and the overall design of the HATTB were derived from observations and studies conducted in Type 4 research activities within the ROAM environment and the HATTB tasks are generally abstracted versions of GCSO task elements. Several tasks are available in the HATTB that can be experimentally manipulated to investigate many potential research questions. Tasks include a spatio-temporal map task in which the goal is to manage/control and deconflict multiple vehicles (with configurable automated/autonomous aids), a signal detection-based search task (with a configurable signaling system), a compensatory tracking task, and a system monitoring task. Scenarios can be augmented with supplemental informational displays, which can also be experimentally manipulated. Following the Parasuraman et al. framework [37], automated aids cover both input and output functions, and the level of automation can be manipulated along with the reliability and error characteristics of those aids. The types of automated aids chosen for the map task reflected those leveraged in ROAM, which include a detect-and-avoid system capable of rerouting around other vehicles and geofenced areas, sensor-based signaling systems capable of indicating potential conflicts that need to be avoided, and informational display elements (e.g., vehicle battery indicator, waypoint paths, data tags). The purpose of the HATTB is not to act as a complete solution to study all potential research threads so it may be an incomplete solution for some researchers. Therefore, the end goal of this testbed is to release it as an open-source software application to allow researchers the flexibility to explore experimental paradigms not natively available in the existing HATTB application.

## IV. Conclusions

This paper has outlined an approach to conduct ecologically relevant, generalizable foundational research as it relates to achieving scalable and increasingly autonomous UAS operations. To accomplish this, initial research and development work has focused on observational studies surrounding the buildup of a remote UAS operations center and GCS. This work has led to generating research questions and hypotheses that can be tested in more controlled laboratory settings, as well as the development of a new testbed designed for conducting experiments. However, the work outlined in this paper provides only the foundation for future research, and many of the topics and issues raised have yet to be tested.

## References

[1] National Academies of Sciences, Engineering, and Medicine, "Advanced Aerial Mobility: A National Blueprint," The National Academies Press, Washington, DC, 2020.

[2] J. Holden and N. Goel, "Uber Elevate: Fast-forwarding to a future of on-demand urban air transportation," 27 10 2016. [Online]. Available: https://www.uber.com/elevate.pdf/. [Accessed 12 November 2019].

[3] C. L. Caraway, "A looming pilot shortage: It is time to revisit regulations," *International Journal of Aviation, Aeronautics, and Aerospace,* vol. 7, no. 2, 2020.

[4] E. T. Chancey, M. S. Politowicz and L. R. Le Vie, "Enabling Advanced Air Mobility operations through appropriate trust in human-autonomy teaming: Foundational research approaches and applications," in *2021 AIAA SciTech*, Virtual Meeting, 2021.

[5] J. B. Holbrook, L. J. Prinzel III, E. T. Chancey, R. J. Shively, M. S. Feary, Q. V. Dao, M. G. Ballin and C. Teubert, "Enabling Urban Air Mobility: Human-autonomy teaming research challenges and recommendations," in *2020 AIAA AVIATION*, Virtual Meeting, 2020.

[6] V. V. Aubuchon, K. E. Hashemi, R. J. Shively and J. M. Wishart, "Multi-vehicle (m:N) operations in the NAS - NASA's research plans," in *AIAA AVIATION Forum*, Chicago, IL, 2022.

[7] E. Salas, T. L. Dickinson, S. A. Converse and S. I. Tannenbauem, "Toward an understanding of team performance and training," in *Teams: Their training and performance*, Norwood, NJ, Ablex, 1992, pp. 3-29.

[8] W. R. Shadish, T. D. Cook and D. T. Campbell, Experimental and Quasi-Experimental Designs for Generalized Causal Inference, Belmont, CA: Wadsworth Cengage Learning, 2002.

[9] U. Neisser, "Chapter 3: Ordinary Seeing," in *Cognition and Reality: Principles and Implications of Cognitive Psychology*, San Francisco, W. H. Freeman and Company, 1976, pp. 33-50.

[10] K. J. Vicente, "Heeding the legacy of Meister, Brunswik, & Gibson: Toward a broader view of human factors research," *Human Factors,* vol. 39, no. 2, pp. 323-328, 1997.

[11] National Aeronautics and Space Administration, "High Density Vertiplex Subproject," 23 August 2021. [Online]. Available: https://www.nasa.gov/aeroresearch/programs/iasp/aam/hdv/description/. [Accessed 9 November 2022].

[12] B. K. Buck, E. T. Chancey, M. S. Politowicz, J. Unverricht and S. C. Geuther, "A Remote Vehicle Operations Center's Role in Collecting Human Factors Data," in *AIAA SciTech*, National Harbor, MD, 2023.

[13] M. S. Politowicz, E. T. Chancey and L. J. Glaab, "Effects of autonomous sUAS separation methods on subjective workload, situation awareness, and trust," in *AIAA SciTech*, Virtual, 2021.

[14] J. Brown, "CERTAIN: City Environment Range Testing for Autonomous Integrated Navigation," NASA Langley Research Center, Hampton, VA, 2016. [Online]. Available: https://ntrs.nasa.gov/api/citations/20160009254/downloads/20160009254.pdf.

[15] M. S. Politowicz, E. T. Chancey, B. K. Buck, J. Unverricht and B. J. Petty, "MPATH (Measuring Performance for Autonomy Teaming with Humans) Ground Control Station: Design Approach and Initial Usability Results," in *AIAA SciTech*, National Harbor, MD, 2023.

[16] "QGC - QGroundControl - Drone Control," [Online]. Available: http://qgroundcontrol.com/. [Accessed October 2022].

[17] J. Unverricht, E. T. Chancey, M. S. Politowicz, B. K. Buck and S. Geuther, "Eye glance behaviors of ground control station operators in a simulated Urban Air Mobility environment," in *41st AIAA/IEEE Digital Avionics Systems Conference (DASC)*, Portsmouth, VA, 2022.

[18] J. Unverricht, E. T. Chancey, M. S. Politowicz, B. K. Buck, S. Geuther and K. Ballard, "Where is the human in the loop? Human factors analysis of extended visual line of sight unmanned aerial system operations within a remote operations environment," in *AIAA SciTech*, National Harbor, MD, 2023.

[19] J. B. Lyons, "Being transparent about transparency: A model for human-robot interaction," in *Trust and Autonomous Systems: Papers from the 2013 AAAI Spring Symposium*, 2013.

[20] E. T. Chancey and M. S. Politowicz, "Designing and Training for Appropriate Trust in Increasingly Autonomous Advanced Aerial Mobility Operations: A Mental Model Approach - Version 1," NASA Technical Memorandum, Hampton, VA, 2021.

[21] J. Welch, "An evidence-based approach to reduce nuisance alarms and alarm fatigue," *Biomedical Instrumentation & Technology,* vol. 45, no. s1, pp. 46-52, 2011.

[22] R. D. Sorkin and D. D. Woods, "Systems with human monitors: A signal detection analysis.," *Human-Computer Interaction,* vol. 1, pp. 49-75, 1985.

[23] E. T. Chancey, *The effects of alarm system errors on dependence: Moderated mediation of trust with and without risk [Unpublished doctoral dissertation],* Old Dominion University, 2016.

[24] C. D. Wickens and C. M. Carswell, "The proximity compatibility principle: Its psychological foundation and relevance to display design," *Human Factors,* vol. 37, pp. 473-494, 1995.

[25] R. A. Rensink, "Change detection," *Annual Reviews in Psychology,* vol. 53, pp. 245-277, 2002.

[26] V. E. Houston, E. T. Chancey, M. S. Politowicz and N. E. Jennings, "A preliminary development of the Intelligent Change Detection System (ICDS): Using machine learning to combat change blindness in remote operation environments," in *AIAA SciTech,* National Harbor, MD, 2023.

[27] H. S. Smallman and M. St. John, "Naive Realism: Misplaced faith in realistic displays," *Ergonomics in Design,* pp. 6-13, 2005.

[28] E. Salas, C. Prince, D. P. Baker and L. Shrestha, "Situation awareness in team performance: Implications for measurement and training.," *Human Factors,* vol. 37, no. 1, pp. 123-136, 1995.

[29] A. Pritchett, M. Portman and T. Nolan, "Research & Technology Development for Human-Autonomy Teaming - Final Report: Literature Review and Findings from Stakeholder Interviews," NASA Langley Research Center, Hampton, VA, 2018.

[30] E. Salas, D. E. Sims and C. S. Burke, "Is there a "Big Five" in teamwork?," *Small Group Research,* vol. 36, no. 5, pp. 555-599, 2005.

[31] C. A. Bowers, C. Braun and B. B. J. Morgan, "Team workload: Its meaning and measurement," in *Team Performance Assessment and Measurement: Theory, Research, and Applications*, M. T. Brannick, E. Salas and C. Prince, Eds., Hillsdale, NJ, Lawrence Erlbaum, 1997, pp. 85-108.

[32] R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," *Human Factors,* vol. 39, no. 2, pp. 230-253, 1997.

[33] National Research Council (NRC), Autonomy research for civil aviation: Toward a new era of flight, Washington, DC: National Academies Press, 2014.

[34] D. B. Kaber, "A conceptual framework of autonomous and automated agents," *Theoretical Issues in Ergonomics Science,* vol. 19, no. 4, pp. 406-430, 2018.

[35] J. R. Comstock and R. J. Arnegard, "The Multi-attribute Task Battery for Human Operator Workload and Strategic Behavior Research (Tech. Memorandum No. 104174)," Hampton, VA., 1992.

[36] Y. Santiago-Espada, R. R. Myer, K. A. Latorella and J. R. Comstock, "The Multi-Attribute Task Battery II (MATB-II) Software for Human Performance and Workload Research: A User's Guide (NASA/TM-2011-217164)," Hampton, VA, 2011.

[37] R. Parasuraman, T. B. Sheridan and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans,* vol. 30, no. 3, pp. 286-297, 2000.

[38] N. Rymer, A. Moore, S. Young, L. Glaab, K. Smalling and M. Consiglio, "Demonstration of two extended visual line of sight methods for urban UAV operations," in *AIAA AVIATION FORUM,* Virtual Event, 2020.