Human Interaction and Emerging Technologies (IHIET-AI), Vol. 201, 2026, 213–225
https://doi.org/10.54941/ahfe1007175

Adaptive Autonomy in the Air Force: Testbed for Human-AI Collaboration

Tanya S. Paul[1], Colin Durrmeyer[2], James P. Bliss[3], Vincent Ferrari[2], and Daniel Lafond[4]

[1]Thales, cortAIx Labs, Palaiseau, France
[2]Centre de recherche de l'Ecole de l'air et l'espace, Salon de Provence, France
[3]Air Force Research Labs, Dayton, Ohio, United States
[4]Thales, cortAIx Labs, Quebec City, Canada

ABSTRACT

Autonomous and Agentic AI-enabled systems are fundamentally reshaping aerial operations—spanning intelligence, surveillance, reconnaissance, targeting, and weapons employment. Such rapid transformation raises a pivotal question: how can we preserve meaningful human control while ensuring that autonomy augments mission effectiveness? Past work has shown that important factors to consider for human-AI collaboration include reliability, dependability, predictability, and transparency as well as relational qualities like responsiveness and its effects on operator engagement in complex and uncertain task environments. In this paper, we introduce a new testbed called STAR-SKY (Sharing Task with Autonomous Resources) enabling investigations on how to dynamically adjust autonomy levels or task allocations between humans and machines can be done across multi-domain (SKY, LAND, SEA etc.). Its goal is to support gathering empirical evidence on how to calibrate and optimize that process based on complex factors such as human workload, fatigue, trust, situational awareness, doctrine, task complexity, authority, interdependence, and differences in mental models. We present initial benchmark data from a pilot study and show how the experiment design and metrics guide integration of meaningful human control within Human-Autonomy Teaming (HAT) systems in the Air Force. We conclude with a discussion on requirements and recommendations for human-in-the-loop experiments with the STAR-SKY simulation.

Keywords: Adaptive control allocation, Human-AI collaboration, Human-autonomy teaming, Dynamic AI tasking

INTRODUCTION

How can designers dynamically allocate tasks and authority between human and AI agents while keeping humans up to speed? Addressing this issue requires a strong interaction between cognitive sciences and informatics (Ferrari, 2019), knowledge about operational constraints as well as integration across engineering, human factors, legal studies, and AI ethics. This issue gave rise to many policies and strategic frameworks affirming the necessity of human oversight and control over those autonomous capabilities. For instance, DoDD 3000.09 states that "Autonomous and semi-autonomous weapon systems will be designed to allow commanders and operators to exercise

Received August 29, 2025; Revised November 28, 2025; Accepted December 8, 2025; Available online April 1, 2026

appropriate levels of human judgment over the use of force."(p.3) Similarly, the U.S. Naval Information Warfare Center–Pacific highlights top humanAI teaming research priorities as the following: (1) defining effectiveness metrics, (2) generating testbeds, (3) developing novel tasksharing paradigms, (4) AI understanding its human teammate, and (5) dedicated development teams—emerging from a National Academies of Sciences, Engineering, and Medicine consensus study supported by the 711 US Human Performance Wing Air Force Research Laboratory (Chiou and Lee, 2024 & Oswald, 2022). The work presented in this paper seeks to advance that agenda by focusing on two of the above priorities: testbed generation and task-sharing paradigm development. The new testbed, called STARSKY, was developed for evaluating a proof-of-concept AIdriven task allocator based on Paul and Lafond's work (2025; 2024) called STAR (Sharing Tasks with Autonomous Systems) within the context of aerial missions. Through the use of a 3D flight simulator, STARSKY enables three approaches to human-AI collaboration in aerial missions: i) Manual, where the human operator retains full control and makes all decisions; ii) Dynamic, where the AI issues takeover recommendations that the human may choose to accept; and iii) Automatic, where an autopilot mode can be activated, allowing the AI to take control autonomously according to predefined criteria. Based on Viswanatha (2025), Agentic AI can be defined as "an advanced form of artificial intelligence that demonstrates autonomous decision-making capabilities, independent goal-setting mechanisms, and adaptive problem-solving behaviors without continuous human intervention." Unlike conventional AI systems that operate within predefined parameters, Agentic AI exhibits a degree of autonomy that enables it to respond dynamically to changing environments and complex scenarios [2]." In this paper, we are interested in Agentic AI particularly present in aerospace, and embedded into aircraft (manned or unmanned) for the allocation of control among human vs AI agent. Particularly, we are interested in how we can implement AI for adaptive human meaningful control.

Human-Autonomy Teaming (HAT) Testbed for the Air Force

The successful integration of artificial intelligence as a teammate for operators relies on two fundamental dimensions. The first concerns the machine–AI interaction, the quality of which depends primarily on technical, industrial, and software performance (Parasuraman & Riley, 1997). The second, radically different, involves the human–AI interaction. This requires that operators be fully integrated from the design and deployment phases of these systems, to ensure their functional relevance, the fluidity of the interaction, the intelligibility of automated actions, and adaptability to the individual specificities of operators—each being unique by nature. These challenges are particularly acute in the military context, given the potential consequences of an error. This aspect is especially highlighted in the field of aeronautics, where missions demand extreme responsiveness, constant adaptability, and immediate decision-making capacity, all while operating in an extremely fast-paced, technical and complex environment. Pilots work in high-information-density settings, where physical and cognitive load can quickly reach critical levels, affecting perception, anticipation, and decision-making. In this context, artificial intelligence can and indeed seems poised to serve as a powerful decision-support assistant, capable of relieving the pilot of certain tasks so that they can focus on critical functions. However, for this cooperation to be effective, one essential parameter must be mastered: the autonomy relationship between AI and the human. The question of task allocation thus becomes central. It cannot be limited to an initial programming setup but must instead include the ability to adjust dynamically according to the complexity of the context, the state of the operator, and tactical objectives while ensuring human accountability in light of the fundamental implications of military actions.

STAR-SKY for the Future of Collaborative Air Combat Systems

STAR aims to address several key challenges associated with human–AI interaction at different levels: command and control (C2) systems through the efficient management of all available resources; AI-based recommendation generation enabling dynamic task distribution between humans and artificial intelligence; testing and training this model in a joint-forces simulator to evaluate "what-if" scenarios for mission preparation and team interoperability; assessing inter-agent coordination between human and autonomous operators according to their specific capabilities and interdependencies; and ensuring meaningful human involvement in the loop, in accordance with the requirements of the Future of Collaborative Combat Aircraft Systems (FCAS) and, more broadly, of collaborative combat doctrines. As complexity increases, STAR also has the capability to incorporate agentic AI to autonomously initiate actions in critical situations where time constraints may limit human decision-making. While the current design prioritizes human oversight, in high-pressure or time-sensitive scenarios, STAR may allow AI to bypass human action selection and assume direct control, optimizing mission success and safety in situations where rapid response is essential. To achieve these objectives, STAR relies on the development of an artificial intelligence system capable of performing advanced multimodal data fusion. This process includes the extraction, preprocessing, and transformation of raw data from multi-domain sensors into interpretable features. The evaluation of human/AI agent capabilities allows the deduction of dependencies between pilots and autonomous agents at all times. which is conducted using a co-active design approach Johnson et al. (2014) that emphasizes the interdependence between human and artificial agents rather than assigning tasks in isolation. Through an independence analysis, this step identifies which tasks e.g altitude control, obstacle detection, trajectory adjustments can be delegated to autonomous systems and which must remain under human control. Lastly, for the dynamic tasking machine learning techniques are employed to anticipate and determine the optimal agent to be tasked for a given task. Finally, STAR ensures that the human operator retains active and informed control. This preservation of the "human-in-the-loop" principle constitutes a fundamental requirement of the FCAS and of all collaborative combat doctrines. Designed with the dual goal of scientific and operational validation of human–AI cooperation principles, this study aims to deepen the understanding of control allocation modalities, particularly their effects on operational performance and the operator's cognitive load.

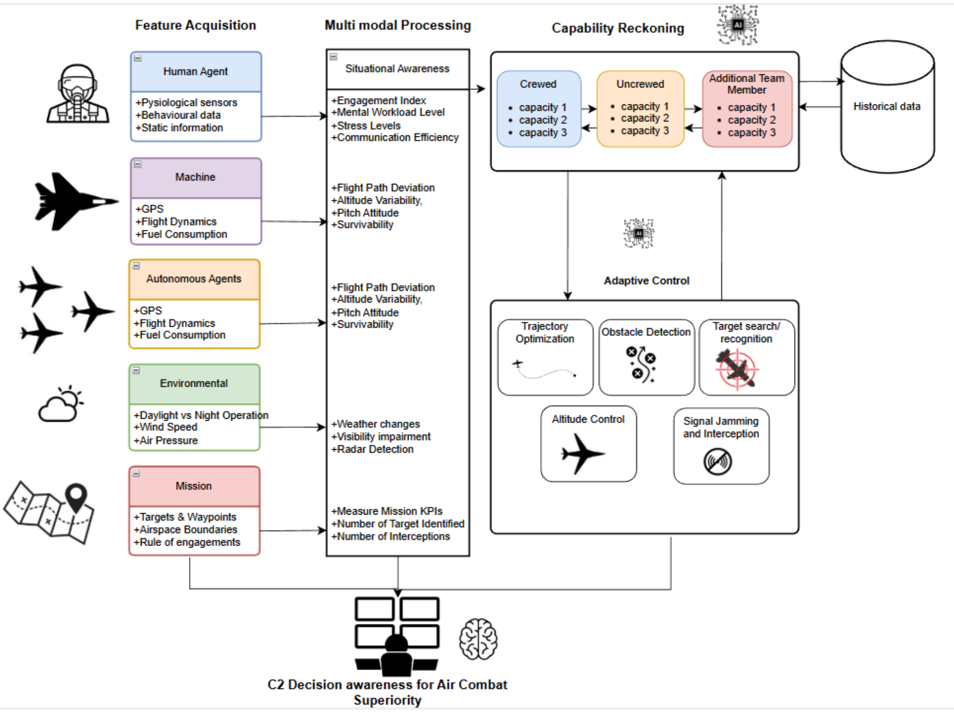

Figure 1: STAR (Sharing Task with Autonomous Resources) adapted into STAR-SKY.

Method for STAR-SKY Experiments

To enhance immersion and the credibility of the experimental scenario, each mission begins with the issuance of an Air Task Order (ATO), a document inspired by the flight plans used in air operations. In the context of this testbed, the ATO consolidates all the necessary information for mission execution. The ATO starts with a fictional geopolitical context, presenting a conflict between two states, followed by a series of tactical restrictions: prohibited overflight of certain urban areas (NFZ), the range of surface-to-air defenses, and a ban on entering the airspace of a neutral country, under penalty of sanctions. The core of the document details the assigned mission: the participant, acting as a pilot of a Rafale aircraft, is tasked with neutralizing several ground targets. The ATO specifies the radio identifiers, communication frequency, the targets to be engaged, the planned refueling sequence after the second strike, and the landing base. The ATO thus serves as a central tool for the mission, combining realistic constraints, operational objectives, and human responsibility within a collaborative human–AI framework. In this experimental scenario, participants are tasked with a series of complex, high-pressure tasks designed to simulate realistic operational challenges in modern air combat. These tasks include neutralizing ground targets within a limited timeframe, conducting an in-flight refuelling operation, and navigating through hostile environments while avoiding surface-to-air missiles (SAMs). Successful performance requires precise and swift decision making, coordination and multitasking, and situation awareness and threat management. Participants must also adhere to strict airspace regulations, demonstrating attention to detail and navigational expertise under non-visual conditions. A key component of the simulation involves human-AI collaboration, where participants interact with an AI-driven drone, making operational decisions based on the drone's data, thus testing their ability to integrate autonomous systems into tactical decision-making. Finally, the simulation challenges participants' cognitive load management and stress resilience, evaluating their capacity to perform effectively while handling several tasks simultaneously under pressure. These tasks collectively aim to replicate the cognitive and operational demands of real-world aerospace missions, offering insights into human performance in complex, high-stakes environments.

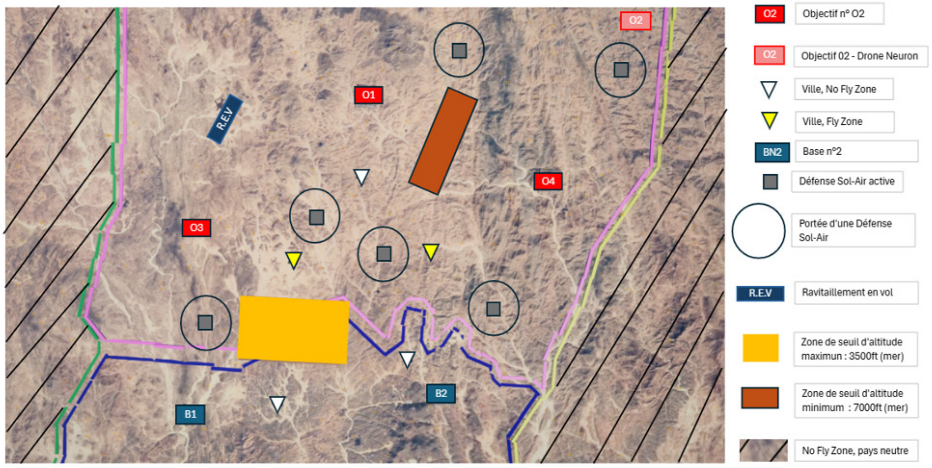

Figure 2: Example of Map of a fictive country, where we can see different objectives mentioned in the ATO (O1, O2, O3), Air Base (BN2), cities as triages, and zones with different flight restrictions e.g max altitude 3500ft or 7000ft, no fly zones.

Experimental Protocol

Initial Presentation: The experiment begins with a general introduction during which the supervisor presents the objectives of the study, the overall procedure, and the expectations for the participant. This phase aimed to establish a clear framework and to align the participant's understanding of the experiment with that of the supervisor.

Free Flight Training: The participant then accessed a ten-minute free training environment focused on familiarization with the aircraft controls. The virtual environment is set in the Grand Canyon, where five waypoints represented by red spheres are positioned around the participant's initial location. Although no constraints are imposed, participants were encouraged to reach these points in order to practice flight precision in preparation for future requirements. The participant may ask questions to the supervisor at any time.

Training for the Simulated Mission: A second ten-minutes training session follows, aimed at familiarizing the participant with the experimental materials (Air Task Order, maps, communication grid) and the main requirements of the simulated mission. In a simplified environment, the participant must perform a strike, an aerial refueling, communications, and a map annotation task using the Neuron drone. This scenario also includes compliance with no-fly zones and the avoidance of surface-to-air defenses. As in the previous phase, no data are recorded at this point.

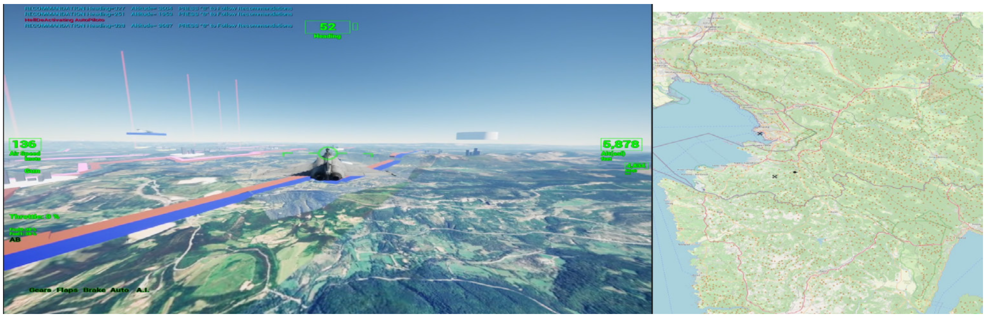

Figure 3: Example of Visual of the experimentation even though the Pilot had a first-person view. On the right of the 3D Simulation, we can see the Map which the C2 had to follow the mission objectives and course of actions.

Sensor Calibration or Baseline: Before the main sessions, physiological sensors described in Table 3 are installed on the participant. A calibration or baseline phase is then conducted: the participant observes a fixed image (a black cross on a white background) for five minutes in a quiet environment, establishing a reference line for cognitive analysis. At the end of this phase, the participant completes a subjective questionnaire.

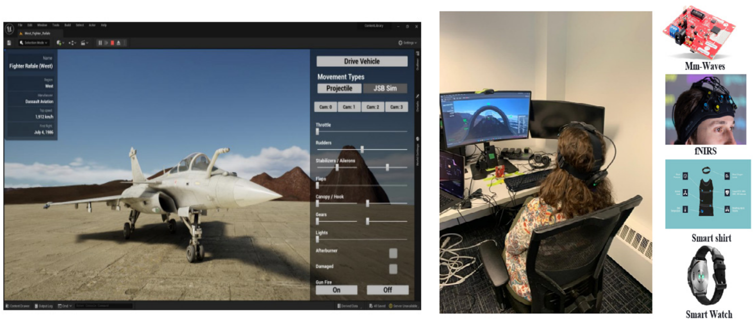

Figure 4: The left image shows the Rafale aircraft and one the right the pilot with the different neurophysiological signals and wearables listed in Table 3.

Human-AI Collaboration Modes: The core of the experiment consists of six flight sessions conducted under the three control allocation modalities, each repeated twice: Reference with Human Control 1 (R1), Manual 1 (M1), Dynamic 1 (D1), Reference 2 (R2), Manual 2 (M2), and Dynamic 2 (D2), defined in Table 1. This structure allows managing learning effects while ensuring a balanced distribution of conditions. Although the mission scenario remains identical to ensure comparability of performance, certain variables are modified from one session to another (for example, grid map coordinates, call signs, and radio frequencies). At the end of each flight session, the participant completes a subjective evaluation questionnaire as described in Table 2 assessing perceived cognitive load and overall experience such as the Task Load Index (NASA TLX), Situational Awareness Techniques (SART), Short Stress State Questionnaire (SSSQ) (Hart & Staveland, 1988; Endsley et al., 1998; Helton, Fields, & Thoreson, 2005).

Table 1. Dynamic control allocation modes.

Conditions	Descriptions
Reference	The reference condition. The human pilot has control over the full navigation capabilities of the aircraft without any assistance.
Manual Control	The human pilot has the ability to activate autopilot simplified which enables adaptive take-over altitude control and heading.
Dynamic Control	An Intelligent embedded AI sends a recommendation to the pilot, which can be accepted or rejected. This dynamic allocation enables smart/ intelligent take over (altitude/heading) and can take the human pilot to the next goal.

Table 2. Subjective questionnaires.

Questionnaires	Descriptions
NASA TLX	It is a multidimensional measure of perceived workload that evaluates six components such as mental demand, physical demand, temporal stress, and perceived performance.
SART	It assesses the participant's situational awareness, that is, their ability to perceive, understand, and anticipate key elements of the mission environment.
SSSQ	It measures the subjective state of stress across three dimensions: cognitive stress, engagement, and discomfort.

The combination of these three tools thus offers a detailed analysis of the participants' subjective experience, complementing the objective data collected during the mission. This approach enables a deeper understanding of the impact of different allocation modalities on perceived mental workload, situational awareness, and experienced stress levels. Secondly, several physiological sensors were used to collect data complementary to performance measurements and subjective questionnaires. These sensors were designed to continuously record specific biological markers likely to reflect the participant's cognitive, emotional, or physical state during the various simulated mission sessions.

Table 3. Neurocognitive sensing.

Sensors	Function and measures collected
Smartwatch	Fossil smartwatch that continuously monitored heart rate, heart rate variability (HRV), and respiration.
Bioharness	A chest-mounted physiological sensor from Zephyr that tracked heart rate, heart rate variability (HRV), and respiration rate.
Mmwave	Millimeter wave (mmWave) radar is a method for contactless monitoring of cardiorespiratory features, it is able to capture dynamic changes in position and physiological states without the need for wearable sensors.
Functional near-infrared spectroscopy	Octamon or Octamon+ devices from Artinis, placed on the forehead that monitored brain activity, specifically prefrontal cortex activation, through changes in blood oxygenation levels.

Data from these sensors are transmitted to a Pixel 4a smartphone via Bluetooth and monitored in real time on mission C2 dashboard. The integration of these sensors made it possible to build a rich corpus of physiological data synchronized with the other experimental metrics. These metrics aim to cross-reference physiological data with subjective and behavioral measures described in Table 4 to deepen the understanding of how different human–AI task allocation modalities affect operators' psychophysiological states.

Table 4. Behavioural metrics.

Metric	Descriptions
Communications	Evaluates the clarity and completeness of the participant's radio communications, including adherence to protocols and timely reporting throughout the mission.
Navigation	Ability to navigate safely within airspace constraints, avoid no-fly zones, and manage in-flight refuelling operations.
Cognitive Load & Stress Management	Assesses the participant's ability to manage multiple tasks and perform effectively under stress, evaluating their cognitive load and performance under pressure.
Mission Objectives	Neutralizing ground targets with limiting time, precision and decision-making ability under pressure. Evaluates how well the participant adheres to mission constraints and follows established operational procedures throughout the mission.
Threat Avoidance (SAMs)	The participant's ability to detect and avoid surface-to-air missile (SAM) threats, ensuring mission success and aircraft safety.
Human-AI Collaboration	Measures the participant's effectiveness in working with AI-driven systems (e.g., drones) to make decisions and integrate AI data into the mission execution process
Task Execution	Ability to complete mission objectives, such as neutralizing ground targets under time pressure, measuring precision, decision-making, and effectiveness. Adherence to mission constraints and established operational procedures throughout the mission.

Case Study Results and Discussion

The main objective was to assess whether an AI teammate, capable of temporarily taking control of the aircraft, could enhance pilot performance while reducing cognitive workload. To this end, three experimental conditions were compared: a reference condition with no assistance, a manual condition using a conventional autopilot, and a dynamic condition based on a simplified autonomous agent. The data extracted included objective measures such as strike success rate, navigation accuracy, and overall mission performance as well as subjective assessments of mental workload, alertness, and situational awareness. Table 5 presents a summary of the benchmark data collected.

Table 5. Benchmark data summary of the mean per condition.

Session	Performance (%)	Time (min) /10	Missed Strikes (%)	Workload NASATLX /20	Alertness (SSQ /5)	Situation Complexity (SART) /20
R	64.37	10.44	33.33	14.37	2.75	14.4
M	81.25	10.50	12.5	13.37	3	12.8
D	88.75	9.25	8.3	12.62	4.37	11.62

This study investigates the effects of dynamic task allocation between humans and artificial intelligence (AI) on performance, cognitive load, and mission success in a simulated aerial task. The results demonstrate that increasing autonomy in the system, from reference (R) to manual (M) and dynamic (D) control, leads to improvements in performance and efficiency. Cognitive load, as measured by the NASA-TLX, decreased with increasing autonomy, suggesting that dynamic allocation reduces mental demand and facilitates better focus on critical tasks. Additionally, participants reported increased alertness and a reduction in the perceived complexity of the task in dynamic conditions, further supporting the beneficial impact of AI assistance. These findings highlight the potential of dynamic AI systems to enhance both performance and cognitive efficiency, offering promising avenues for future research into more advanced autonomous systems in human-machine collaboration/ HAT.

Next steps and Recommendations

These capabilities correspond directly to Table 6's emphasis on dynamic allocation and cross-level coordination, other metrics could be transition warnings, and percentage of take-over processes to assess control distribution.

Table 6. Recommendations based on Levels of Autonomy in Cognitive Control (LACC).

LACC Level and Focus	Recommendation for Allocation of Control	Human (H) Role / AI Role (A)
Level 1: Physical Execution and Action	Delegate immediate, low-level, time-critical actions (e.g., motor control, sensor activation, path following) to AI, as these rely on precision and reaction speed.	H maintains supervisory oversight and intervenes when A diverges from H's intent and safety.
Level 2: Implementation Adjustment and Constraints	Share control between humans and AI for real-time adjustments such as parameter tuning, timing, and constraint management.	H defines operational boundaries and approves A's adaptive response. A can propose or recommend adjustments based on its understanding.
Level 3: Generic Functions and Plans Procedural Knowledge	Allow AI to execute standardized, well-understood procedures autonomously (e.g., checklists, standard operating sequences), while humans oversee exceptions.	H evaluates plan appropriateness and reframes procedures when context shifts.
Level 4: Values and Trade-offs Decision	Keep humans central in defining and prioritizing value-based trade-offs (e.g., safety vs. efficiency, risk vs. reward).	A's should display explainable and transparent reasoning for trade-offs. H should ensure give feedback and interact so A can learn from H.
Level 5 – Effects and Goals Functional Purpose	Assign responsibility for goal-setting and overall mission intent to humans, with AI assisting in monitoring progress and detecting goal conflicts.	AI can recommend goal adjustments under changing conditions as long as it's aligned with H'intent and initial boundary set.
Level 6: Frames Context and Situation Framing	Reserve framing and contextual interpretation—defining *what situation is occurring and why it matters*—for human cognition, supported by AI situational-awareness tools.	A should conduct data fusion comings from multi-modal sources to generate operational picture.
Cross Level Guidance Dynamic Allocation	Enable fluid transitions ("sliding autonomy") across levels as context and task demands evolve. Ensure transparency and maintain shared situational understanding between human and AI agents.	Both should adapt control dynamically balancing workload, transparency team and mission performance.

Lundberg and Johansson (2020) introduce the Joint Control Framework Score (JCF-S) for analyzing human-AI interaction and shared control in complex systems. The framework consists of three components: (1) Process Mapping (PM), which identifies core processes and control mechanisms; (2) the Levels of Autonomy in Cognitive Control (LACC) model, which describes cognitive interactions at varying levels of autonomy; and (3) the JCF-S, which visualizes dynamic control exchanges. It identifies "control joints" where human and machine processes intersect, enabling analysis of autonomy shifts, transparency, adaptive roles, and temporal control episodes. The framework supports structured assessments of human–autonomy systems by abstracting patterns across episodes. Its three-step approach begins with process mapping to identify relevant processes and agents, followed by LACC, which evaluates autonomy levels and interactions, and concludes with Human–Machine Interaction Temporal Analysis (HMI-T) to explore control evolution over time and its impact on coordination and workload. In military aviation, effective human-AI control allocation requires managing task transfers, situational awareness, and take-over coordination. Transition warnings (e.g., Level 2) ensure smooth shifts of control, while situational awareness (Level 6) supports shared understanding between agents. Measures such as trust, transparency, and explainability (Level 5) are critical for collaboration. For dynamic control allocation, AI should display the rationale behind decisions (e.g., unstable altitude control, Level 1) and offer insights into its performance predictions. Finally, HMI-T metrics should assess collaborative performance, communication of agent status, and coordination of task execution for smooth transitions.

CONCLUSION

The STAR system is aligned with the Joint Framework for Control (JFC) through its process mapping using interdependent analysis. These dependencies allow for pre-determined decisions and allocation to be guided not only by technical feasibility but also by operational doctrines and rules of engagement, which may explicitly prohibit certain functions from being performed by AI. In this way, process mapping provides a principled foundation for role allocation that reflects both system capabilities and mission-level constraints. STAR operationalizes the (LACC) framework by continuously evaluating dependencies between human and autonomous agents and optimizing autonomy levels to support cooperative performance. At the lower cognitive levels (LACC Levels 1–2), STAR automates physical and implementation tasks through advanced multimodal data fusion, allowing the AI to process sensor data and manage real-time control loops while the human retains supervisory authority. At the procedural level (Level 3), machine learning algorithms enable the AI to execute standardized operational functions and alert the operator when deviations occur, supporting flexible task reconfiguration. At higher cognitive levels, STAR assists in evaluating value-based trade-offs (Level 4) and monitoring mission goals and effects (Level 5), providing decision support that enhances transparency and trust. At the most abstract level (Level 6), STAR's contextual reasoning tools support the operator in interpreting complex situations while preserving the human's ultimate authority over strategic framing. The data collected in this study relate to the JCF-S by providing key performance and cognitive metrics that track the dynamics of human-AI collaboration. Objective measures like strike success rate and navigation accuracy, along with subjective assessments of workload and alertness, validate the control transitions and cognitive load shifts described in the JCF-S. These data demonstrate how STAR optimizes autonomy levels to reduce cognitive load and improve performance, aligning with the core principles of the JCF-S framework. By continuously redistributing control based on task demands and situational complexity, STAR-SKY seeks to maintain an active human-in-the-loop structure, optimizing workload, transparency, and mission performance through dynamic human-AI coordination.

Disclaimer: The opinions expressed are solely those of the author(s). They do not represent an endorsement by or the views of the United States Air Force, the Department of Defense, or the United States Government.

REFERENCES

Chiou, E. K., & Lee, J. R. (2024). Meaningful control in human-AI systems: Trusting AI agents. In *NATO HFM-RSY-377 on meaningful human control in information warfare*, October 2024.

Endsley, M. R., et al. (1998). A comparative analysis of SAGAT and SART for evaluations of situation awareness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42(1), 82–86. https://doi.org/10.1177/154193129804200119

Ferrari, V. (2019). Man-machine teaming: Towards a new paradigm of man-machine collaboration? In *Transformation of skills and uses induced by innovations in ISTE-Disruptive technology and defense innovation ecosystems* (Chapter 6, Part 2).

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139–183). Elsevier. https://doi.org/10.1016/B978-0-444-42425-5.50013-2

Helton, W. S., Fields, D., & Thoreson, J. A. (2005). Assessing daily stress with the Short Stress State Questionnaire (SSSQ): Relationships with cognitive slips-failures. *Proceedings of the Human Factors and Ergonomics Society*, 49, 886–890. https://doi.org/10.1177/154193120504900295

Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., Van Riemsdijk, M. B., & Sierhuis, M. (2014). Coactive design: Designing support for interdependence in joint activity. *Journal of Human-Robot Interaction*, 3(1), 43–62. https://doi.org/10.5898/JHRI.3.1.Johnson

Lundberg, J., & Johansson, B. J. E. (2020). A framework for describing interaction between human operators and autonomous, automated, and manual control systems. *Cognition, Technology & Work*. https://doi.org/10.1007/s10111-020-00637-w

Oswald, F. L., et al. (2022). The National Academies Board on Human-Systems Integration (BOHSI) Panel: Human-AI teaming: Research frontiers. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66(1), 130–134. https://doi.org/10.1177/1071181322661007

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, and abuse. *Human Factors*, 39(2), 230–253. https://doi.org/10.1518/001872097778543960

Paul, T. S., & Lafond, D. (2024). Meaningful human control in multi-agent systems through AI-based dynamic task allocation strategies. In *NATO HFM-RSY-377 on meaningful human control in information warfare*, October 2024.

Paul, T. S., Lafond, D., & Benzakine, P.-Y. (2025). AI-driven collaboration mode selection for manned-unmanned teaming in future combat air systems. In *2025 International Conference on Military Communication and Information Systems (ICMCIS)*, Oerias, Portugal, 1–10. https://doi.org/10.1109/ICMCIS64378.2025.11047730

Viswanathan, P. S. (2025). Agentic AI: A comprehensive framework for autonomous decision-making systems in artificial intelligence. *International Journal of Computer Engineering and Technology (IJCET)*, 16(1), 862–880.