Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 1205–1215
https://doi.org/10.54941/ahfe1008350

A Multidimensional Framework for Assessing Human–AI Collaboration in Military Operations: The CogARMY Teamwork Index (CTI)

Tanya S. Paul[1], Daniel Lafond[2], Dale Richards[3], Gianluca Borghini[4], and Johan De Heer[5]

[1]Thales CortAIx, Palaiseau, France
[2]Thales CortAIx, Quebec City, Canada
[3]Thales CortAIx, Glasglow, United Kingdom
[4]La Sapienza University, Italy
[5]Thales CortAIx, Amesterdam, The Netherlands

ABSTRACT

The increasing integration of autonomous systems and artificial intelligence (AI) into military operations is transforming teamwork, requiring new approaches to assess performance in hybrid human–AI teams operating in dynamic, high-stakes environments. Existing teamwork assessment frameworks offer robust theoretical foundations for evaluating leadership, communication, coordination, and adaptability, but were developed primarily for human-only teams and are not fully suited to Human-AutonomyTeaming (HAT), defined by O'Neill et al. (2022) as interdependent collaboration between one or more humans and one or more autonomous agents, each fulfilling distinct roles while pursuing shared mission objectives. This paper proposes a CogArmy Team Index (CTI), a multidimensional framework for assessing teamwork effectiveness across human-only and hybrid human–AI military teams. Developed through the Cognitive Training and Teamwork Assessment of Army Personnel (CogARMY) initiative and NATO Science and Technology Organization Human Factors and Medicine research task group HFM-374, the CTI extends established teamwork theory with autonomy-specific dimensions, including trust calibration, explainability, human–AI coordination, adaptive task allocation, and human–AI synchronisation. The framework adopts a multilevel Input–Process–Output architecture linking contextual inputs to teamwork processes, coordinating mechanisms, autonomy management, regulation, and mission outcomes. The CTI leverages a multimodal, data-driven methodology integrating behavioural, physiological, and AI-derived metrics, overcoming limitations of self-report measures while enabling assessment of both human and autonomous team members. Composite scores combine theory-driven and empirical weighting to support adaptability across operational contexts. The proposed framework offers a scalable, scientifically grounded approach for evaluating and optimising teamwork in next-generation military environments, supporting operational readiness and future human–AI collaboration.

Keywords: Human-autonomy teaming, Teamwork assessment, Military human factors, Trust calibration, Multimodal assessment

Received July 20, 2026; Revised September 27, 2026; Accepted October 3, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

INTRODUCTION

Teamwork is critical to military operations: effective coordination, communication, and cohesion directly influence mission success and operational efficiency (Essens et al., 2005; Salas et al., 2005). Traditional assessment methods provide useful data but often fail to capture nuanced interactions, adaptive decision-making, and the high-pressure demands of modern military teams (Turnage et al., 1990), and struggle to separate individual skill from teamwork effectiveness under realistic mission conditions. The growing presence of synthetic agents in decision-making further requires assessment tools capable of addressing hybrid and multi-agent environments. At the same time, advances in neuroscience, wearable sensing, data analytics, and artificial intelligence create new opportunities for deeper, more objective insight into team dynamics. This paper proposes the CogARMY Teamwork Index (CTI), a multidimensional framework for assessing teamwork effectiveness across both human-only and hybrid human–AI military teams, developed through the Cognitive Training and Teamwork Assessment of Army Personnel (CogARMY) initiative and NATO STO HFM-374. Section 2 defines teamwork for hybrid and non-hybrid teams; Section 3 reviews current assessment methods and their limitations; Section 4 surveys enabling technologies; and Section 5 presents the CTI architecture and its six-layer structure. By combining established teamwork theory with recent advances in AI-enabled and neurophysiological assessment, the CTI aims to provide a reliable, scalable tool for guiding team design, monitoring readiness, and improving mission outcomes across diverse operational contexts.

DEFINING TEAMWORK FOR HYBRID AND NON-HYBRID TEAMS

Before reviewing assessment methods, it is useful to distinguish team, teamwork, and team performance. A team is a distinguishable set of two or more individuals who interact interdependently and adaptively to achieve shared objectives (Salas et al., 1992). Teamwork refers to the collaborative processes through which members coordinate knowledge, skills, and resources toward common goals, whereas team performance reflects the extent to which those goals are achieved. The two constructs are related but distinct: teamwork describes how a team functions, while performance describes what it accomplishes. Effective teamwork does not guarantee high performance, since external factors such as time pressure, uncertainty, fatigue, and task complexity can constrain outcomes a distinction that is especially relevant in high-risk domains such as military operations, where teams may demonstrate strong teamwork despite operating under conditions that limit overall performance. Because teamwork is a dynamic social process rather than an outcome, its assessment requires examining communication, coordination, cooperation, conflict management, and inclusive participation, which offers more actionable feedback for team development than performance metrics alone. Teamwork research has expanded considerably since the mid-twentieth century (Benne & Sheats, 1948; Bales, 1950). Belbin (1993) proposed that

effective teams comprise complementary behavioural roles rather than individuals with similar characteristics; this framework remains relevant to hybrid human–AI teams, where roles such as Coordinator or Implementer can be interpreted as functional contributions distributed according to the capabilities of each team member, human or artificial, while responsibilities involving ethical judgement or safety-critical decisions remain under human supervision.

Communication is fundamental to effective teamwork regardless of team composition: clear, timely information exchange supports coordination and shared situational awareness, whereas communication failures consistently reduce effectiveness (Cooke et al., 2001). Human–agent teams differ according to how authority and decision-making are distributed among members (Scholtz, 2003), with humans and agents assuming roles including supervisor, operator, teammate, or observer. Early human–robot interaction research prioritised maximising robot autonomy while minimising human workload (Kaupp & Makarenko, 2008); although efficient, this approach can overlook shared situational awareness, transparency, and the human oversight needed to supervise autonomous systems effectively. Task allocation in hybrid teams must therefore consider cognitive responsibilities such as monitoring and decision-making, not only physical actions. Trust shaped by both interaction quality and the alignment of agent behaviour with human expectations (White & Richards, 2006; Hancock et al., 2011) remains a critical factor in human–agent collaboration, much as interpersonal trust underpins effective human-only teams (Dirks & Ferrin, 2001). Several frameworks have identified the behavioural competencies underlying effective teamwork. The "Big Five" model (Salas et al., 2005) identifies Team Leadership, Mutual Performance Monitoring, Back-up Behaviour, Adaptability, and Team Orientation as core competencies. Complementing this, Lavelle et al. (2020) developed the Temporal Observational Analysis of Teamwork, an observational method analysing verbal and non-verbal behaviours in real time across five domains: Team Leadership, Shared Mental Models, Requesting and Providing Assistance, Team Performance Monitoring, and Team Attitudes. Collectively, these frameworks converge on a consistent set of core dimensions mutual performance monitoring, back-up behaviour, adaptability, team orientation, leadership, shared mental models, closed-loop communication, mutual trust, information sharing, and cohesiveness that provide the conceptual foundation for the assessment methods reviewed next.

Current Methods to Assess Teamwork

The limitations of traditional methods have driven the development of technology-enabled approaches capable of more objective, continuous, multidimensional evaluation. Advances in simulation, AI, physiological sensing, neuroimaging, and multimodal data integration are transforming teamwork assessment by enabling real-time analysis of behavioural, cognitive, and physiological interaction advances that are particularly relevant to hybrid human–AI teams, where collaboration extends beyond human–human interaction to include intelligent autonomous agents.

Simulation-based assessment provides a controlled, operationally representative environment for evaluating teamwork without the risks of live operations, ranging from Synthetic Task Environments to Virtual Reality and Live-Virtual-Constructive training (Cooke & Shope, 2017; Fernandez et al., 2017). Compared with field exercises, simulation enables repeatable scenarios, objective performance logging, and systematic evaluation of communication and coordination, though effectiveness depends on selecting an appropriate level of fidelity for training objectives rather than physical realism alone (Jones, 2021). A significant recent development is the integration of synthetic teammates autonomous agents participating in collaborative tasks enabling controlled experimentation on human–AI collaboration, with mixed human–agent teams shown to achieve performance comparable to all-human teams under certain conditions (Demir & Cooke, 2014; McNeese et al., 2018). Nonetheless, simulations cannot fully reproduce the physical, psychological, and environmental stressors of real operations.

AI and machine learning (ML) are increasingly used to automate behavioural coding and improve the scalability of observational assessment, supporting analysis of communication patterns, action recognition, and multi-person tracking in complex operational environments (Sreenu & Durai, 2019). Natural language processing and large language models further enable automated analysis of verbal interactions even with limited training data (Idalski Carcone et al., 2019; Spain et al., 2025). Supervised learning has been used to replicate expert instructor assessments of trainee performance (Benesch et al., 2027), wearable physiological sensors have distinguished novice from expert pilots during complex flight manoeuvres (Ferrari et al., 2018), and NLP techniques have enabled automated measurement of closed-loop communication directly from transcribed speech (Robinson et al., 2024). These technologies are best viewed as augmenting rather than replacing expert human observers, and together with digital twins and intelligent simulation environments extend assessment beyond retrospective analysis toward predictive, adaptive training.

The deployment of AI and autonomous systems in military operations also requires assessment methods that extend beyond human-only teamwork models. HAT frameworks emphasise adaptive collaboration in which authority, task allocation, and decision-making are dynamically distributed between humans and autonomous agents (Singh et al., 2021; Zhao et al., 2020). Traditional autonomy taxonomies classify technological capability rather than collaborative effectiveness (Torkjazi & Raz, 2022); more recent work instead advocates adaptive autonomy, whereby autonomy levels continuously adjust to mission requirements, operator workload, trust, and system confidence (Paul & Lafond, 2024; Paul et al., 2024, 2026). Related concepts adjustable autonomy, mixed-initiative systems, and Coactive Design similarly emphasise flexible allocation of authority and reciprocal interaction between human and autonomous teammates (Johnson et al., 2011, 2014), highlighting that future assessment must evaluate trust calibration, shared situational awareness, and the evolution of decision authority throughout mission execution, not task performance alone.

Wearable physiological sensors provide continuous, objective measures of operators' cognitive and affective states such as heart rate, respiration, and electrodermal activity widely applied to estimate stress, workload, and fatigue, all of which influence teamwork and operational performance (Sciaraffa et al., 2021). Neurophysiological techniques offer complementary insight: electroencephalography (EEG) has long been used to monitor mental workload, attention, and fatigue (Borghini et al., 2012), and inter-brain synchrony measured via EEG or fNIRS has been associated with cooperative behaviour, shared attention, and team effectiveness (Wang et al., 2025). Most such studies remain confined to laboratory hyperscanning conditions, but advances in portable sensing and multimodal signal processing are steadily improving feasibility for realistic military environments. No single technology captures the full complexity of teamwork; there is accordingly increasing emphasis on integrated frameworks that combine behavioural observation, communication analysis, physiological sensing, and AI-driven analytics (Cooke et al., 2013; Fiore et al., 2014; Klein et al., 2003), which together with psychometric data can support adaptive task allocation, readiness assessment, and dynamic human–AI collaboration (Paul et al., 2024).

Towards a Teamwork Index for Hybrid/Non-Hybrid Military Teams

The proposed CTI addresses this gap by extending established teamwork theory to accommodate collaboration between humans and intelligent agents. Rather than replacing existing frameworks, the CTI preserves their theoretical foundations while introducing autonomy-specific dimensions explainability, adaptive autonomy, trust calibration, human meaningful control, and dynamic task allocation that reflect the distinctive characteristics of HAT. Core teamwork processes remain centred on communication, coordination, leadership, adaptability, mutual performance monitoring, and backup behaviour, while coordinating mechanisms encompass shared mental models, situational awareness, communication quality, and trust, the last conceptualised as a dynamic, bidirectional construct incorporating both human confidence in AI and measurable indicators of AI reliability. The CTI is designed as a scalable framework with two complementary implementation pathways: a high-technology configuration that exploits multimodal sensing, wearable physiology, AI-assisted behavioural analysis, and simulation data for continuous, objective assessment; and a low-technology configuration that operationalises the same conceptual dimensions using structured behavioural observations, Behaviourally Anchored Rating Scales, self-report instruments, after-action reviews, and mission performance metrics. Maintaining identical conceptual dimensions across both pathways ensures consistency of interpretation while allowing deployment from laboratory simulations to field exercises and operational units. Methodologically, the CTI follows an Input–Process–Output framework linking operational context, team composition, task characteristics, and autonomy levels to teamwork processes and coordinating mechanisms, ultimately influencing mission outcomes including effectiveness, safety, resilience, and efficiency. Indicators will undergo iterative validation through simulation, exercises,

and operational studies to establish reliability, construct validity, and predictive relationships with team effectiveness, with composite scores derived from a combination of theory-driven and data-driven weighting that adapts to mission characteristics and operational demands. As autonomous capabilities evolve, the CTI is intended to remain an adaptive framework that incorporates emerging evidence, technological advances, and evolving concepts of HAT.

The CogARMY Teamwork Index (CTI)

The CTI is organised into six hierarchical layers representing the principal determinants of teamwork effectiveness in human and hybrid human–AI teams (Table 1). The framework begins with operational inputs describing team composition, mission characteristics, autonomy level, and environmental constraints, which influence the core teamwork processes derived from the Big Five model (leadership, coordination, adaptability, backup behaviour, team orientation, and mutual performance monitoring). These processes are supported by coordinating mechanisms shared understanding, communication quality, situational awareness, and mutual trust. To accommodate hybrid teams, an additional layer captures autonomy management, including explainability, trust calibration, adaptive task allocation, human oversight, and dynamic redistribution of decision authority. A regulation layer evaluates cognitive and physiological stability through indicators such as workload, fatigue, stress, and error management, while a final outputs layer evaluates mission success, decision quality, safety, resilience, efficiency, and coordination quality. Together, these layers provide a comprehensive yet flexible structure supporting objective, multimodal assessment across diverse military operating environments, compatible with both traditional and AI-enabled assessment methodologies.

Table 1: Proposed CTI layers.

Layer	Component	Dimensions	Description
L1	Inputs	Team composition; task complexity; autonomy level; skills (human & AI); operational constraints; mission uncertainty; command structure; resource availability	Defines the operational conditions that shape teamwork performance (mission demands, team structure, and capability distribution across human and AI agents).
L2	Core Teamwork Processes	Leadership & coordination; mutual performance monitoring; backup behaviour; adaptability; team orientation; team composition	Core observable teamwork behaviours extended for Human-Autonomy Teaming, capturing coordination between humans and AI agents under dynamic conditions.
L3	Coordinating Mechanisms	Shared understanding; mutual trust; communication quality; situation awareness alignment	Key mechanisms enabling coordination and reducing errors in human–AI and human-only collaboration.

(Continued)

Table 1: Continued.

Layer	Component	Dimensions	Description
L4	Autonomy & Interaction Control	Explainability of AI decisions; trust calibration; task allocation decisions; handover timing; level of human control; AI initiative level	Describes how control is distributed and adjusted between humans and AI during task execution.
L5	Regulation & Performance Stability	Stress level; cognitive load; fatigue; vigilance; error detection; attention maintenance; performance degradation	Captures human and system regulation under pressure, uncertainty, and workload variation.
L6	Outputs	Mission success; decision accuracy; response time; safety; efficiency; resilience; team coordination quality	Final operational outcomes reflecting combined human–AI or human-human performance in task execution.

Practically, assessment proceeds through the six layers as interconnected components of team effectiveness rather than a single technique. Layer 1 (Inputs) characterises the operational context team composition, mission complexity, autonomy level, command structure, and environmental uncertainty providing the basis for interpreting subsequent teamwork behaviours. Layer 2 (Core Teamwork Processes) evaluates leadership, coordination, adaptability, mutual performance monitoring, backup behaviour, and team orientation, extending the Big Five framework to HAT. Layer 3 (Coordinating Mechanisms) captures communication quality, shared situational awareness, shared mental models, and mutual trust. Layer 4 (Autonomy and Interaction Control) the element distinguishing the CTI from traditional teamwork models evaluates explainability, trust calibration, adaptive task allocation, human oversight, initiative sharing, and authority transfer, capturing the quality of collaboration rather than simply the capability of autonomous systems. Layer 5 (Regulation and Performance Stability) assesses the cognitive and physiological resilience of individuals and teams, drawing where available on wearable physiological sensors, neurophysiological monitoring, behavioural observation, or validated subjective instruments, allowing implementation across both technologically enabled and resource-constrained environments. Layer 6 (Outputs) integrates information from the preceding layers into measures of mission success, decision quality, safety, efficiency, resilience, and coordination quality. Because the CTI distinguishes teamwork processes from performance outcomes, it enables evaluators to identify not only whether a team performed effectively, but which behavioural, cognitive, organisational, or technological factors contributed to success or failure.

CONCLUSION

The proposed CTI offers a comprehensive framework for assessing teamwork in military operations involving human, hybrid human–AI, and future autonomous teams. By extending established teamwork theory with

autonomy-specific constructs, it addresses the limitations of frameworks developed primarily for human-only teams, integrating behavioural, cognitive, physiological, and technological indicators within a unified Input–Process–Output architecture that distinguishes collaborative processes from operational outcomes. A key strength is scalability: the framework supports both high-technology implementations incorporating multimodal sensing and AI-assisted analytics, and low-technology implementations based on structured observation, mission performance metrics, and validated rating instruments, facilitating adoption across diverse military environments while maintaining a consistent conceptual foundation. Future work will focus on empirical validation through laboratory studies, simulation, field exercises, and operational evaluations, establishing the reliability, construct validity, and predictive value of individual indicators, refining composite scoring methodologies, and evaluating relationships between teamwork processes and mission effectiveness. Beyond assessment, the CTI has the potential to inform training, support adaptive team design, and enhance mission effectiveness by providing objective, actionable insight into the processes that underpin effective teamwork.

REFERENCES

Bakeman, R., & Haynes, S. (2015). Behavioral observation. In R. Cautin & S. Lilienfeld (Eds.), The Encyclopedia of Clinical Psychology. John Wiley & Sons.

Bales, R. F. (1950). Interaction Process Analysis. Addison-Wesley.

Belbin, R. M. (1993). Team Roles at Work. Butterworth-Heinemann.

Benesch, D., Lafond, D., Lemaire, N., MacMullin, M., Pynn, R., & Tay, H. H. (2027, in press). AI-based trainee proficiency assessment in marine training simulations. In Adaptive Instructional Systems, Proceedings of HCII 2026 (Part 1, LNCS 16737). Springer.

Benne, K. D., & Sheats, P. (1948). Functional roles of group members. Journal of Social Issues, 4(2), 41–49.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. Science, 323(5916), 892–895. https://doi.org/10.1126/science.1165821

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2012). Measuring neurophysiological signals in team collaboration. Neuroscience Letters, 529(2), 81–86.

Britton, N., Stubbe, D., Blaha, J. T., & Davis, D. (2017). Development and validation of a behavioral marker tool for assessing individual teamwork skills. Military Medicine, 182(Suppl. 1), 282–289. https://doi.org/10.7205/MILMED-D-15-00528

Cooke, N. J., Kiekel, P. A., & Helm, E. E. (2001). Comparing and validating measures of team knowledge. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 45(4), 361–365.

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. Cognitive Science, 37(2), 255–285. https://doi.org/10.1111/cogs.12009

Cooke, N. J., & Shope, S. M. (2017). Designing a synthetic task environment. In Scaled Worlds: Development, Validation and Applications (pp. 273–288). Routledge.

Demir, M., & Cooke, N. J. (2014). Human teaming changes driven by expectations of a synthetic teammate. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58(1), 16–20.

Dirks, K. T., & Ferrin, D. L. (2001). Trust in leadership: Meta-analytic findings and implications for research and practice. Journal of Applied Psychology, 87(4), 611–628.

Driskell, J. E., Salas, E., & Hughes, A. M. (2010). Collective orientation and team performance: Development of an individual differences measure. Human Factors, 52(2), 316–328. https://doi.org/10.1177/0018720810376059

Essens, P. J. M. D., Vogelaar, A. L. W., Mylle, J. J. C., Blendell, C., Paris, C., Halpin, S. M., & Baranski, J. V. (2005). Military Command Team Effectiveness: Model and Instrument for Assessment and Improvement (NATO RTO HFM-087 TP/59). NATO Research and Technology Organization.

Fernandez, R., Shah, S., Rosenman, E. D., Kozlowski, S. W., Parker, S. H., & Grand, J. A. (2017). Developing team cognition: A role for simulation. Simulation in Healthcare, 12(2), 96–103.

Ferrari, V., Gagnon, J.-F., Camachon, C., & Kopf, M. (2018). Psycho-physiological evaluation of the pilot: A study conducted with pilots of the French Air Force. In Engineering Psychology and Cognitive Ergonomics (pp. 285–295). Springer. https://doi.org/10.1007/978-3-319-91122-9_24

Fiore, S. M., Wiltshire, T. J., Oglesby, J. M., O'Keefe, W. S., & Salas, E. (2014). Complex collaborative problem-solving processes in mission control. Aviation, Space, and Environmental Medicine, 85(4), 456–461.

Gorman, J. C. (2014). Team coordination and dynamics: Two central issues. Current Directions in Psychological Science, 23(5), 355–360. https://doi.org/10.1177/0963721414545215

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. Human Factors, 53(5), 517–527.

Idalski Carcone, A., Hasan, M., Alexander, G. L., Dong, M., Eggly, S., Brogan Hartlieb, K., & Kotov, A. (2019). Developing machine learning models for behavioral coding. Journal of Pediatric Psychology, 44(3), 289–299.

Johnson, M., Vera, A., & Johnson, B. (2011). Coactive design: Designing support for interdependence in human–robot teams. Journal of Human-Robot Interaction, 1(1), 43–69. https://doi.org/10.5898/JHRI.1.1.Johnson

Johnson, M., Vera, A., & Knox, E. (2014). Principles of observability, predictability, and directability in collaborative human-robot systems. Human Factors, 56(5), 936–948. https://doi.org/10.1177/0018720814527050

Jones, M. (2021). Fidelity and knowledge transfer in simulation-based learning. Educational Technology Research and Development, 69, 1203–1221.

Kaupp, T., & Makarenko, A. (2008). Measuring human-robot team effectiveness to determine an appropriate autonomy level. Proceedings of the 2008 IEEE International Conference on Robotics and Automation, 2146–2151. https://doi.org/10.1109/ROBOT.2008.4543524

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. IEEE Intelligent Systems, 18(3), 81–85. https://doi.org/10.1109/MIS.2003.1200735

Konak, A., Clark, T. K., & Nasereddin, M. (2014). Using Kolb's experiential learning cycle to improve student teamwork in engineering education. Journal of Engineering Education, 103(3), 449–466.

Lavelle, M., Reedy, G. B., Cross, S., Jaye, P., Simpson, T., & Anderson, J. E. (2020). An evidence-based framework for the Temporal Observational Analysis of Teamwork in healthcare settings. Applied Ergonomics, 82, 102915. https://doi.org/10.1016/j.apergo.2019.102915

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming with a synthetic teammate: Insights into human-autonomy teaming. Human Factors, 60(2), 262–273.

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

Paul, T. S., & Lafond, D. (2024). AI agents learning human decision policies for collaborative situation assessment in NORAD C2 operations. Proceedings of the IEEE 4th International Conference on Human-Machine Systems (ICHMS 2024). https://doi.org/10.1109/ICHMS59971.2024.10555797

Paul, T. S., Lafond, D., et al. (2024). NeuroTeaming: Using power spectral density for adjusting teaming dynamics in pilot-AI task allocation. Proceedings of AHFE 2024. https://doi.org/10.54941/ahfe1004735

Paul, T. S., Durrmeyer, C., Bliss, J., Ferrari, V., & Lafond, D. (2026). Adaptive autonomy in the air force: Testbed for human-AI collaboration. Proceedings of the 16th International Conference on Human Interaction & Emerging Technologies: Artificial Intelligence & Future Applications (HAET 2026). https://doi.org/10.54941/ahfe1007175

Robinson, T., Smith, J., & Lee, K. (2024). AI-based measures of closed-loop communication in teams. Journal of Artificial Intelligence Research, 71, 233–250.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), Teams: Their Training and Performance (pp. 3–29). Ablex.

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "Big Five" in teamwork? Small Group Research, 36(5), 555–599. https://doi.org/10.1177/1046496405277134

Scholtz, J. (2003, January). Theory and evaluation of human-robot interactions. Proceedings of the 36th Hawaii International Conference on System Sciences (HICSS 36).

Sciaraffa, N., Borghini, G., Di Flumeri, G., & Aricò, P. (2021). Neurophysiological indicators of cooperation in team tasks. Journal of Neural Engineering, 18(4), 046002.

Singh, R., Chanel, L., & Roy, R. (2021). Adaptive autonomy in human-robot teams: Concepts, frameworks, and evaluation methods. Autonomous Robots, 45, 123–141. https://doi.org/10.1007/s10514-020-09932-7

Smy, V., Cahillane, M., MacLean, P., Hilton, M., & Humphreys, L. (2021). Evaluating teamwork development in combat training settings: An exploratory case study utilising the Junior Leaders' Field Gun competition. Applied Ergonomics, 95, 103459.

Spain, R., Min, W., Kumaran, V., Pande, J., Saville, J., & Lester, J. (2025). Applying large language models to enhance dialogue and communication analysis for adaptive team training. International Journal of Artificial Intelligence in Education, 1–35.

Sreenu, G., & Durai, S. (2019). Intelligent video surveillance: A review through deep learning techniques for crowd analysis. Journal of Big Data, 6(1), 1–27.

Torkjazi, M., & Raz, G. (2022). Taxonomies for autonomous unmanned aerial systems: Frameworks and operational considerations. Journal of Defense Technology, 18(2), 88–102. https://doi.org/10.1016/j.jdt.2022.03.005

Turnage, J. T., Houser, T. L., & Hofmann, D. A. (1990). Assessment of Performance Measurement Methodologies for Collective Military Training (ARI-RN-90-126). U.S. Army Research Institute for the Behavioral and Social Sciences.

Waller, M. J., & Kaplan, S. A. (2018). Systematic behavioral observation for emergent team phenomena: Key considerations for quantitative video-based approaches. Organizational Research Methods, 21(2), 500–515.

Wang, S. K., Hashemi, G., Zhang, Y., & Zheng, B. (2025). Exploring a novel approach to assessing surgical team collaboration: Evidence of brain activity synchronization. Laparoscopic, Endoscopic and Robotic Surgery, 8(1), 6–12.

White, A., & Richards, D. (2006, June 21–22). Building operator trust in semi-autonomous UCAV systems. Proceedings of the 1st "Moving Autonomy Forward" Conference, Grantham, UK.

Zhao, Y., Niu, J., & Shen, X. (2020). Cognitive load and stress in human-UAV teaming: Implications for adaptive autonomy. Human Factors, 62(7), 1035–1050. https://doi.org/10.1177/0018720819893210