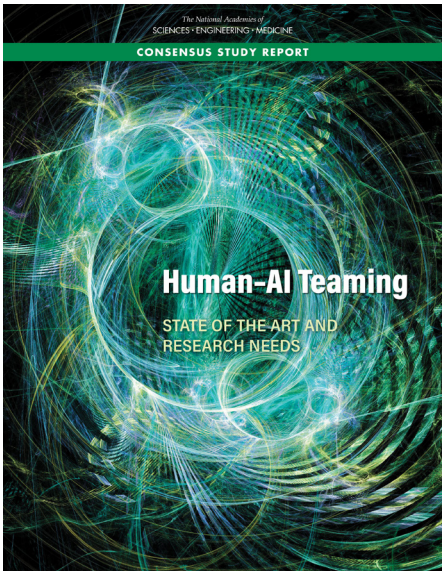
Consensus Study Report

December 2021

HIGHLIGHTS

# Human-AI Teaming: State-of-the-Art and Research Needs




The Department of Defense has recognized the growing need to move toward multi-domain operations (MDO), a shift that requires that major systems and crucial data sources are integrated across the traditionally separate air, land, maritime, space, and cyberspace domains. The ability to rapidly collect and analyze dynamic and timely data as well as understand relationships between these data is essential to support MDO. Artificial intelligence (AI) is being considered as a potential tool to support command and control for future MDO by significantly increasing the ability to analyze data, boost the processing rate of a wide variety of data inputs, automate mission planning, and create faster predictive targeting and systems maintenance.

To achieve these potential benefits, however, AI possesses a number of significant technical limitations or barriers that must be addressed, including perceptual limitations, hidden biases, and lack of a model of causation that is important for understanding and predicting future events, among others. Given these challenges, it is envisioned that AI systems will be implemented in ways that augment, but do not fully replace human decision makers in MDO. Therefore, the effect that AI has on human decision makers within MDO must be carefully considered and designed to accommodate.

Humans are significantly challenged in monitoring and understanding complex automation, including AI, resulting in poor situation awareness (SA). AI software may also introduce hidden biases that can negatively affect decision making. To be effective in supporting future military operations, AI must be reliable and robust across a wide variety of potential future missions and must work well as a team with humans. The development of effective human-AI teams capable of taking advantage of the unique abilities of both humans and AI, while overcoming the known challenges of each team member, are needed.

Recognizing the challenge that AI has for human performance, the Air Force Research Laboratory (AFRL) asked the National Academies of Sciences, Engineering, and Medicine to convene an expert committee to examine and recommend research to guide the appropriate use of this technology in future operations. The resulting committee, which included experts in human factors, cognitive engineering, human-computer interaction, industrial-organizational psychology, and AI, as well as experts in military operations related to human-autonomy teaming, assessed the state of research on human-AI teaming and determined gaps and future research

*The National Academies of*
SCIENCES • ENGINEERING • MEDICINE

priorities. The committee's report, *Human-AI Teaming: State-of-the-Art and Research Needs*, examines the factors that are relevant to the design and implementation of AI systems with respect to human operations and recommends needed research for achieving successful performance across teams of AI and human decision makers.

## KEY RESEARCH PRIORITIES: ADVANCING EFFECTIVE HUMAN-AI TEAMS

Through its analysis of currently available research, the committee developed key research priorities designed to strengthen the effectiveness of human-AI teams within future MDO. These priorities, outlined below, address key areas including human-AI team models, team effectiveness and processes; SA; AI transparency and explainability; human-AI interaction approaches; trust; reduction of human and AI bias; training needs; and an examination of human-systems integration (HSI) processes.

**Improving Human-AI Team Models.** A number of models of human-AI teams exist that provide an indication of relevant factors for team performance. Improved computational models of human-AI teams are needed that consider interrelated, dynamically evolving, distributed, and adaptive collaborative tasks common in MDO. Improved metrics of human-AI team performance need to be developed that consider team interdependencies and dynamic role assignments, and the reduction of uncertainty, as well as the ability of the AI to meet the expectations of warfighters.

**Strengthening Team Effectiveness and Processes.** Research is needed to improve team effectiveness in long-term, distributed, and agile human-AI teams through improved team assembly, goal alignment, communication, coordination, social intelligence, and the development of a new human-AI language. While this research can leverage the large body of existing work on human-human teaming, the committee recognizes that new research is needed to better understand and support effective team processes between humans and AI.

**Increasing Situation Awareness.** Human SA is critical for effective MDO performance, including for the oversight of AI. As such, methodologies for supporting individual and team SA in command and control operations need to be extended to MDO, and methods for using AI to support information integration, prioritization, and routing across the joint battle space are needed, as well as for improving resilience to adversarial attacks on SA. Methods for improving human SA of AI systems are needed, as well as research to create shared SA within the human-AI team. The degree to which AI systems should have both self-awareness and awareness of their human teammates needs to be explored to determine the benefit for overall team performance. Future AI systems will need to possess integrated situation models to appropriately understand the current situation and project future situations for decision making.

**Improving Transparency and Explainability.** Improved AI transparency and explainability are critical to achieving improved human SA of AI systems and to build trust. Research is needed to better define the information requirements and methods for achieving transparency of machine learning (ML)-based AI systems. Further research is also needed to examine the relationship between AI explainability and trust. Effective mechanisms to adapt explanations to receivers' needs, prior knowledge and assumptions, and cognitive and emotional states need to be developed. Research should also be conducted to determine whether explanations of human reasoning could likewise improve AI and team performance.

**Strengthening Interaction Mechanisms and Strategies.** Research to determine improved methods for supporting collaboration between humans and AI in shared functions, to support human operators working with AI at multiple levels of automation (LOAs), and to determine methods for maintaining or regaining SA when working with AI at high LOAs (i.e., on-the-loop control) is needed. Future research should also determine new requirements to support dynamic functional assignments across human-AI teams and the best methods for supporting dynamic transitions in LOAs over time. Research to develop a better understanding and prediction of emergent human-AI interactions, and a better understanding of the effects of interaction design decisions on human-AI team resilience would be beneficial.

**Improving Trust in Human-AI Teams.** Trust in AI is essential. Better documentation of the decision context and goals involved in the teaming environment to advance understanding of how broader sociotechnical factors affect trust in human-AI teams is important. Also, improved measures of trust are needed that draw on the importance of cooperation, and that separate the concept of distrust from trust. Dynamic models of trust are also needed that capture how trust evolves and affects performance outcomes in various human-AI team contexts.

**Reducing Human and AI Bias.** The potential for bias in AI systems can be introduced through the development of its algorithms and systemic biases in training sets among other factors. Humans can also suffer from decision biases and can be directly affected by the accuracy of the AI system, creating a human-AI team bias. Research is needed to better understand the interdependencies between human and AI decision-making biases, how these evolve over time, and methods for detecting and preventing bias with ML-based AI.

**Developing Training to Support Human-AI Environment.** Training of the human-AI team is needed to develop the appropriate team constructs and skills necessary for effective performance. In particular, research should examine what, when, why, and how to best train human-AI teams, taking into consideration various team compositions and sizes. In addition, training may be needed to better calibrate human expectations of AI teammates and to foster appropriate levels of trust.

**Building HSI Processes, Measures, and Testing.** Achieving the successful development of an AI system that can function as a good teammate requires advances in HSI processes and measures. Good HSI practice will be key to the design, development, and testing of new AI systems. New HSI design and testing methods for effective human-AI teams will also be needed. Improved approaches for multidisciplinary AI development teams should include human factors engineers, socio-technical researchers, systems engineers, and computer scientists. Improved test-beds and metrics for human-AI teaming may also be needed, as well as the development of AI lifecycle testing and auditability.

Action on the above research priorities will support the military in its move toward MDO by significantly advancing human-AI teaming competence. This research provides an essential foundation for improving performance across the joint human-AI team, ultimately strengthening the use of AI in the military context, as well as in many other environments in which people will need to work closely with AI systems.

---


## COMMITTEE ON HUMAN-SYSTEM INTEGRATION RESEARCH TOPICS FOR THE 711TH HUMAN PERFORMANCE WING OF THE AIR FORCE RESEARCH LABORATORY

**MICA R. ENDSLEY (*Chair)***, SA Technologies; **BARRETT S. CALDWELL**, Purdue University; **ERIN K. CHIOU**, Arizona State University; **NANCY J. COOKE**, Arizona State University; **MARY L. CUMMINGS**, Duke University; **CLEOTILDE GONZALEZ**, Carnegie Mellon University; **JOHN D. LEE**, University of Wisconsin-Madison; **NATHAN J. MCNEESE**, Clemson University; **CHRISTOPHER MILLER**, Smart Information Flow Technologies; **EMILIE ROTH**, Roth Cognitive Engineering; **WILLIAM B. ROUSE** (NAE), Georgetown University; **DANIEL TALMAGE**, *Study Director*.


---

**For More Information . . .** This Consensus Study Report Highlights was prepared by the Board on Human-Systems Integration based on the Consensus Study Report, *Human-AI Teaming: State of the Art and Research Needs* (2021). The study was sponsored by the Wright Brothers Institute as a subcontract to the U.S. Air Force Research Laboratory. Any opinions, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect the views of any organization or agency that provided support for the project. Copies of the Consensus Study Report are available from the National Academies Press, (800) 624-6242 or https://www.nap.edu/26355.

Division of Behavioral and Social Sciences and Education

*The National Academies of*
SCIENCES • ENGINEERING • MEDICINE

The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people's lives worldwide.
**www.national-academies.org**