AI-Enabled Decision Support Systems: Tool or Teammate?

Myke C. COHEN [a,1], Mickey V. MANCENIDO [b]
[a] *Arizona State University, Mesa, AZ, USA*
[b] *Arizona State University, Glendale, AZ, USA*

Abstract. In this talk, we discuss perspectives on the deployment of artificial intelligence technologies as decision support systems in high-stakes scenarios, such as national defense. We will present a project proposed to the US Department of Defense (DoD) that asks questions regarding the benefits of perceiving AI as a tool or as teammate in light of trust and the teamness of human-AI systems.

Keywords. Decision support systems, Trust, Human-AI teaming, Teamness

1. Background

1.1. People and Decision Support Systems

Decision support systems (DSSs) are critical components of the U.S. Department of Defense (DoD)'s multi-domain operations (MDO) in warfare [1]. However, the adoption of faulty DSS recommendations can have unintended effects that can be lethal or catastrophic. An infamous example is from 2004, when US military personnel failed to veto automated engagements made by the Patriot missile system, ultimately causing the fratricide of British and American pilots [2]. Thus, for ethical and legal motivations, keeping humans in the decision-making loop remains the status quo in the design of DSSs for high-stakes domains such as national security, defense, and warfare [3].

DSSs have traditionally supplemented human cognitive capabilities with computerized information processing efficiencies to improve decision-making quality and speed [4]. Initial applications from the 1970s-1990s were largely tools that collated and presented information to support human decision-making, such as in military housing occupancy assignment, officer manpower planning, and aircraft design compendiums [5, 6]. In the 2000s, DSS design philosophy shifted towards prosthetic functionalities that recommended decisions and actions altogether [2, 4]. The hyperactive acceleration and democratization of machine and deep learning methods in recent years have introduced artificially intelligent DSSs (AI-DSS; [7]) that are capable of processing highly complex information and executing actions autonomously, often at the cost of transparency and explainability [8]. An example of an AI-DSS central to the DoD's MDO roadmap includes the navigation and engagement mechanisms of unmanned combat vehicles [9,10].

[1] Corresponding Author: Myke C. Cohen, Arizona State University, Mesa, AZ, USA; E-mail: myke.cohen@asu.edu.

People typically occupy supervisory roles over DSSs to perform checks and balances, especially over high-stakes decision domains [11]. However, AI-DSSs can be found in contexts that require rapid decision-making, in which human supervision over AI-recommended decisions may result in suboptimal system performance. Consider automated border control gates at airports, for example, where AI-DSSs are used to verify if a traveler's identification document matches their identity [12]. Under human supervisory control setups, each face matching task is first performed by the AI-DSS to generate a recommended decision (e.g., "identity verified" or "needs further verification") for a border security agent's final approval. But with automated face matching capabilities already at par with or surpassing human accuracy in most circumstances [13], it is unclear whether the presence of human supervisors truly increases the accuracy of automated border control systems. Furthermore, people and AI have different sets of strengths and weaknesses in face matching, likely resulting from differences in how each process facial images [14]. Border control agents may thus fail to distinguish their own face-matching performance relative to that of their AI counterpart as they make a final face-matching decision. This is likely also the case for other high-stakes domains in which human decision-making speed and accuracy tend to lag behind AI capabilities.

Nevertheless, it has been argued that people remain vital even where the presence of a human supervisor may not significantly improve system performance [15]. The presence of people helps satisfy legal and ethical considerations in automated decision-making settings [16], particularly in critical situations that require human supervisors to innovate non-algorithmic solutions [2, 15, 17]. Note, however, that people's abilities to effectively monitor and intervene in automated decisions in such instances are subject to numerous cognitive biases and limitations. For example, people often become complacent in the presence of AI, resulting in the ineffectual detection of errors or irregularities in AI outputs [18]. People are also prone towards *automation bias*, or the human tendency to blindly adopt an automation-recommended decision even while faced with evidence that the recommendation is faulty [19]. These cognitive phenomena have been linked to various inappropriate uses of DSSs, such as overreliance on automated decisions, aversion to algorithm-heavy decision processes, and the relegation of human operators to passive decision-making roles [20, 21].

1.2. AI-DSSs as Human-AI Teams

The inherent limitations of human decision-making—both at the individual level as people perform decision-making tasks, as well as the interaction level as people consider decisions recommended by DSS counterparts—raise questions about the role of people in the high-stakes domains wherein AI-DSSs are likely to be found. For instance, can the involvement of people strengthen decision recommendations made by AI-DSSs? In 2005, amateur chess players interacting with AI-DSSs in so-called "centaur" teams successfully defeated grandmaster-level individuals and AI algorithms to demonstrate that this may be true in some instances [22]. How is it that collaborations between non-expert human decision-makers and expertly-trained AI algorithms can result in even greater performance than the AI algorithms performing alone?

An explanation may be found in the theory that DSS-assisted decisions can be more meaningfully understood as outcomes of joint cognitive systems that are formed as people interact with automation to achieve a common goal [23]. This is akin to how teams comprising only people interacting with each other have been understood to achieve team-level decision-making outcomes beyond the sum of individual team member inputs [24]. With rapid advancements in AI capabilities, the idea that humans can form teams with highly autonomous forms of automation to perform similar feats has recently gained traction [25], though not without controversy (*cf.* [26, 27]). Current definitions of human-AI teams stipulate that people and AI teammates must interact while performing unique but interdependent tasks towards common goals [25]. However, we posit with [28] that when interdependent interactions between people and AI result in emergent cognitive phenomena that cannot be broken down into individual member contributions, “teamness” is exhibited and interactions must also be considered at the system level—regardless of whether the system fits traditional definitions of teaming.

Considering AI-DSSs and the cognitive outputs of their interactions with people in light of human-AI teaming is particularly relevant in light of design guidelines that recommend providing AI-embedded systems with human-like characteristics for the purpose of being perceived, trusted, or interacted with anthropomorphically—or further still, to be accepted as a teammate (e.g., [29–31]). Arguments in [26] refute the claim that AI should be treated in systems design as human-like agents, and provide justification for instead framing AI-DSSs as tools or “supertools”. However, the proliferation of AI products and interfaces like the Amazon Echo and ChatGPT that can interact with people in humanlike manners have been observed to increase people’s propensity towards personifying and socializing with certain types of AI as though they were teammates [32]—a trend that is only projected to grow with the rise of generative AI algorithms [33]. Theories of AI teammate-likeness (e.g., [34]) support the idea that such features may make people more likely to engage with AI counterparts in ways that support effective team cognitive processes [28]. However, there is also evidence that people might still interact with highly autonomous AI-DSSs as though they were interacting with another person but without perceiving it as if it were human or a teammate. For instance, people may withhold criticism, use verbal politeness cues (“please”, “excuse me”, etc.), or talk about an AI using gendered language, yet retrospectively view it simply as an inanimate object [35–38]. These interactive behaviors and expectations not only relate to the team-like performance of human-AI teams, but also to sociocognitive phenomena that drive human interactions with AI-DSSs, particularly trust [39].

1.3. Trust in AI-DSSs

Trust has been related to biased decision-making in DSSs, stemming from mismatches between human expectations about the performance of automated decision-making systems versus the actual levels thereof [4, 40]. In human supervisory control settings, trust has been defined as a person’s willingness to rely on automation as an aid to achieve specific goals [40]. The relationship between human trust in automation and automation performance has thus been a focal point of interest in the design of semi-autonomous or autonomous technology [20, 40, 41].

The decision-making demands of the contexts in which AI-DSS “centaurs” and other human-AI teaming settings are likely to be found make it such that people may not necessarily be able to effectively monitor and veto all decisions recommended (and in some cases, automatically executed) by their AI counterparts [42]. Thus, interactions with AI-DSSs can be characterized by a heightened sense of vulnerability, making trust central to maintaining effective system interactions in the long run [41]. It has been noted that people are increasingly taking on the role of collaborators rather than supervisors when teaming with advanced AI that is capable of adaptively interacting with them [43]. In such cases, people may demonstrate higher propensities to adopt relational trusting expectations and behaviors as they interact with AI [41]. Trust in AI-DSSs may therefore manifest and be measured not only in how people may respond to AI recommendations, but also in how they respond to lateral exchanges with AI. This is in contrast with how trust in automation has been largely measured through unidirectional perceptual or behavioral measures, such as reflective questionnaires that ask people how they perceive it (e.g., [44, 45]) and observational metrics on how people adopt AI-recommended decisions (i.e., compliance) or rely on automation to perform actions in the absence of indications that interventions are needed [46]. Techniques to measure trust in reciprocal human-AI team relationships have been proposed (e.g., distributed dynamic team trust; [47]). However, these typically rely on network metrics that limit their applicability to human-AI dyadic contexts in which most human-DSS interactions presumably take place [43].

This proposed project addresses several important questions in light of the current state of trust and team conceptualizations for AI-DSSs: how important is it for AI-DSS to be viewed as a teammate? Is it desirable for AI-DSS to be perceived as teammate-like, considering the cognitive biases surrounding trust that already plague human decision-making in traditional DSS contexts? Are teammate-like perceptions of AI-DSS necessary to harness the benefits of emergent team cognitive processing?

2. Research Questions

In this proposed research, we ask three questions that partially address how people import human social expectations surrounding trust into human-AI decision-making contexts where the AI *may* be interacted with or considered as a teammate. The proposed project aims to answer the following questions:

1. Do people evaluate the trustworthiness of an AI agent differently depending on whether they are prompted to consider it as a teammate or a tool?
2. Do patterns of human-AI decision-making interactions over time correlate with trustworthiness perceptions? Do these patterns and correlations change depending on whether people are prompted to consider AI-DSS as a teammate or a tool?
3. How does an AI agent's teammate-likeness and perceived trustworthiness affect the teamness of human-AI decision-making interactions?

References

[1] National Academies of Sciences E. Human-AI Teaming: State-of-the-Art and Research Needs; 2021.

[2] Cummings ML. Automation and Accountability in Decision Support System Interface Design. JOTS. 2006 Sep;32(1).

[3] De-Arteaga M, Fogliato R, Chouldechova A. A Case for Humans-in-the-Loop: Decisions in the Presence of Erroneous Algorithmic Scores. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20. New York, NY, USA: Association for Computing Machinery; 2020. p. 1-12.

[4] McGuirl JM, Sarter NB. Supporting Trust Calibration and the Effective Use of Decision Aids by Presenting Dynamic System Confidence Information. Hum Factors. 2006 Dec;48(4):656-65.
[5] Eom S, Kim E. A Survey of Decision Support System Applications (1995–2001). J Oper Res Soc. 2006 Nov;57(11):1264-78.
[6] Eom HB, Lee SM. A Survey of Decision Support System Applications (1971–April 1988). Interfaces. 1990 Jun;20(3):65-79.
[7] Gupta S, Modgil S, Bhattacharyya S, Bose I. Artificial Intelligence for Decision Support Systems in the Field of Operations Research: Review and Future Scope of Research. Ann Oper Res. 2022 Jan;308(1-2):215-74.
[8] Rudin C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. Nat Mach Intell. 2019 May;1(5):206-15.
[9] Cooke NJ, Gawron V. Human Systems Integration for Remotely Piloted Aircraft Systems. In: Remotely Piloted Aircraft Systems: A Human Systems Integration Perspective. John Wiley & Sons, Ltd; 2016. p. 1-14.
[10] Fahey KM, Miller MJ. DoD Unmanned Systems Integrated Roadmap 2017–2042. Department of Defense; 2017.
[11] Sheridan TB. Human Supervisory Control. In: Salvendy G, editor. Handbook of Human Factors and Ergonomics. 4th ed. John Wiley & Sons, Ltd; 2012. p. 990-1015.
[12] Sanchez del Rio J, Moctezuma D, Conde C, Martin de Diego I, Cabello E. Automated Border Control E-Gates and Facial Recognition Systems. Computers & Security. 2016 Sep;62:49-72.
[13] O'Toole AJ, An X, Dunlop J, Natu V, Phillips PJ. Comparing Face Recognition Algorithms to Humans on Challenging Tasks. ACM Trans Appl Percept. 2012 Oct;9(4):16:1-16:13.
[14] Phillips PJ, O'Toole AJ. Comparison of Human and Computer Performance across Face Recognition Experiments. Image and Vision Computing. 2014 Jan;32(1):74-85.
[15] Parasuraman R, Wickens CD. Humans: Still Vital After All These Years of Automation. Hum Factors. 2008 Jun;50(3):511-20.
[16] Ananny M, Crawford K. Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability. New Media & Society. 2018 Mar;20(3):973-89.
[17] Atkins E. Emergency Landing Automation Aids: An Evaluation Inspired by US Airways Flight 1549. In: AIAA Infotech@Aerospace 2010. Atlanta, Georgia: American Institute of Aeronautics and Astronautics; 2010. .
[18] Wickens CD, Clegg BA, Vieane AZ, Sebok AL. Complacency and Automation Bias in the Use of Imperfect Automation. Hum Factors. 2015 Aug;57(5):728-39.
[19] Skitka LJ, Mosier KATHLEEN, Burdick MD. Accountability and Automation Bias. International Journal of Human-Computer Studies. 2000 Apr;52(4):701-17.
[20] Parasuraman R, Riley V. Humans and Automation: Use, Misuse, Disuse, Abuse. Hum Factors. 1997 Jun;39(2):230-53.
[21] Snow T. From Satisficing to Artificing: The Evolution of Administrative Decision-Making in the Age of the Algorithm. Data & Policy. 2021/ed;3:e3.
[22] Case N. How To Become A Centaur. Journal of Design and Science. 2018 Jan.
[23] Woods DD. Cognitive Technologies: The Design of Joint Human-Machine Cognitive Systems. AI Magazine. 1985 Dec;6(4):86-6.
[24] Cooke NJ, Gorman JC, Myers CW, Duran JL. Interactive Team Cognition. Cogn Sci. 2013 Mar;37(2):255-85.
[25] O'Neill T, McNeese N, Barron A, Schelble B. Human–Autonomy Teaming: A Review and Analysis of the Empirical Literature. Hum Factors. 2020 Oct:001872082096086.
[26] Shneiderman B. Intelligent Agents and Supertools. In: Human-Centered AI. New York: Oxford University Press; 2022. p. 99-103.
[27] Groom V, Nass C. Can Robots Be Teammates?: Benchmarks in Human–Robot Teams. Interaction Studies. 2007 Jan;8(3):483-500.
[28] Cooke NJ, Cohen MC, Fazio WC, Inderberg LH, Johnson CJ, Lematta GJ, et al. From Teams to Teamness: Future Directions in the Science of Team Cognition. Hum Factors. 2023 Mar.
[29] Murphy J, Gretzel U, Pesonen J. Marketing Robot Services in Hospitality and Tourism: The Role of Anthropomorphism. Journal of Travel & Tourism Marketing. 2019 Sep;36(7):784-95.
[30] Pelau C, Dabija DC, Ene I. What Makes an AI Device Human-like? The Role of Interaction Quality, Empathy and Perceived Psychological Anthropomorphic Characteristics in the Acceptance of Artificial

Intelligence in the Service Industry. Computers in Human Behavior. 2021 Sep;122:106855.

[31] Fenwick A, Molnar G. The Importance of Humanizing AI: Using a Behavioral Lens to Bridge the Gaps between Humans and Machines. Discov Artif Intell. 2022 Aug;2(1):14.

[32] Purington A, Taft JG, Sannon S, Bazarova NN, Taylor SH. "Alexa Is My New BFF": Social Roles, User Satisfaction, and Personification of the Amazon Echo. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. CHI EA '17. New York, NY, USA: Association for Computing Machinery; 2017. p. 2853-9.

[33] Erwin S. On National Security — Analyzing Intelligence in the Age of ChatGPT; 2023.

[34] Wynne KT, Lyons JB. An Integrative Model of Autonomous Agent Teammate-Likeness. Theoretical Issues in Ergonomics Science. 2018 May;19(3):353-74.

[35] Nass C. Etiquette Equality: Exhibitions and Expectations of Computer Politeness. Commun ACM. 2004 Apr;47(4):35-7.

[36] Nass C, Moon Y. Machines and Mindlessness: Social Responses to Computers. J Social Isssues. 2000 Jan;56(1):81-103.

[37] Cohen MC, Demir M, Chiou EK, Cooke NJ. The Dynamics of Trust and Verbal Anthropomorphism in Human-Autonomy Teaming. In: 2021 IEEE 2nd International Conference on Human-Machine Systems (ICHMS). Magdeburg, Germany: IEEE; 2021. p. 1-6.

[38] Fischer K. Interpersonal Variation in Understanding Robots as Social Actors. In: 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI); 2011. p. 53-60.

[39] Parasuraman R, Miller CA. Trust and Etiquette in High-Criticality Automated Systems. Commun ACM. 2004 Apr;47(4):51-5.

[40] Lee JD, See KA. Trust in Automation: Designing for Appropriate Reliance. Hum Factors. 2004 Mar;46(1):50-80.

[41] Chiou EK, Lee JD. Trusting Automation: Designing for Responsivity and Resilience. Human Factors: The Journal of the Human Factors and Ergonomics Society. 2021 Apr:001872082110099.

[42] Zieba S, Polet P, Vanderhaegen F, Debernard S. Principles of Adjustable Autonomy: A Framework for Resilient Human–Machine Cooperation. Cogn Tech Work. 2010 Sep;12(3):193-203.

[43] Chiou EK, Lee JD. Cooperation in Human-Agent Systems to Support Resilience: A Microworld Experiment. Hum Factors. 2016 Sep;58(6):846-63.

[44] Jian JY, Bisantz AM, Drury CG. Foundations for an Empirically Determined Scale of Trust in Automated Systems. International Journal of Cognitive Ergonomics. 2000 Mar;4(1):53-71.

[45] Chancey ET, Bliss JP, Yamani Y, Handley HAH. Trust and the Compliance–Reliance Paradigm: The Effects of Risk, Error Bias, and Reliability on Trust and Dependence. Hum Factors. 2017 May;59(3):333-45.

[46] Meyer J, Lee JD. Trust, Reliance, and Compliance. In: Lee JD, Kirlik A, editors. The Oxford Handbook of Cognitive Engineering. Oxford Library of Psychology. New York, NY, US: Oxford University Press; 2013. p. 109-24.

[47] Huang L, Cooke NJ, Gutzwiller R, Berman S, Chiou E, Demir M, et al. Distributed Dynamic Team Trust in Human, Artificial Intelligence, and Robot Teaming. In: Trust in Human-Robot Interaction. 1st ed. Academic Press; 2020. p. 301-19.