Quenching the Thirst for Human-Machine Teaming Guidance: Helping Military Systems Acquisition Leverage Cognitive Engineering Research

Patricia L. McDermott[a], Katherine E. Walker[a], Cynthia O. Dominguez, Ph.D.[a], Alex Nelson[b] and Nicholas Kasdaglis, Ph.D.[a]
[a]*The MITRE Corporation*
[b] *Air Force Research Laboratory*

ABSTRACT

There is a need in systems acquisition that is not currently being met. The insights from cognitive engineering research in human-automation interaction are not being systematically applied to acquisition processes associated with operational military systems. To address this gap, we synthesized guidance from the literature and translated it into a set of general cognitive interface requirements for human-machine teaming. By presenting the guidance as requirements, we are attempting to remove barriers from effective insights being used in implementation. This paper describes ten themes of human-machine teaming that need to be supported: Observability, Predictability, Directing Attention, Exploring the Solution Space, Directability, Adaptability, Common Ground, Calibrated Trust, Design Process, and Information Presentation. Example requirements are provided for Exploring the Solution Space. The general set of requirements can be tailored to specific systems as needed. To support this tailoring, we are developing and piloting cognitive task analysis techniques focused on human-machine teaming.

KEYWORDS

Coordination; military; cognitive engineering; common ground; systems development; human automation interaction

INTRODUCTION

The cognitive engineering (CE) communitity has produced useful research findings and related insights for designing effective collaboration between automation and humans. These findings and insights are not making their way into the design of operational systems. We are finding a general lack of familiarity and application of these findings among program managers, systems engineers and developers. There is no systematic process for applying this human-automation research when engineering complex human-machine systems. Users and designers talk about concepts like “transparency” but it is not clear what this means in terms of design requirements for a system. The systems acquisition community may not be familiar with the CE literature or they may not have the necessary background to apply the findings to envisioned or upgraded systems. Regardless, a gap exists between the CE research and the system accquistion process for fielding new systems.

The lack of CE insight in highly autonomous systems is apparent. So-called “clumsy” automation has many negative consequences, such as brittle performance, miscalibrated trust in the system, and lack of user acceptance (Wiener, 1989; Sarter, Woods, & Billings, 1997; Lee & See, 2004; Parasuraman & Riley, 1997). Although the research community has answered the call to provide guidance on how to improve interactions between humans and advanced technological systems (Lyons, 2013), surprisingly little work has been done to translate the literature into materials that non-researchers can use as guidance for design of highly automated and autonomous systems.

Our analysis has aggregated and packaged the human-machine teaming (HMT) literature into several interconnected themes for the design of human-autonomy interaction.

METHODS

Researchers completed three main tasks: literature review, analysis, and creation of general cognitive interface requirements. First, to leverage and package existing research for system acquision personnel and program managers, we reviewed the large applied literature on human-automation interaction, human-machine trust, human-robot teaming, flight deck automation surprises, and other related work. Seventy-six papers were triaged to identify the papers most likely to yield citable guidance for HMT. We summarized 39 of these papers, extracting key information on HMT philosophy, principles, key terms, and requirements and considerations (The authors can be contacted for a list and summary of papers reviewed). The research findings were analyzed for common themes and elements of cognitive support. The analysis yielded ten themes that capture unique aspects relevant for HMT design across the literature and domains. Guidance was sorted into these themes and then translated into requirements language.

RESULTS

Themes for Design of Effective HMT

Below are the ten themes that emerged from our bottom-up analysis of HMT guidance in the literature. The first eight reflect guidance on design content and thus are ripe for tailoring through cognitive task analysis (CTA). The last two, design process and information presentation, are of a different nature. They capture insights about how design teams work together and processes they use for analysis and validation, as well as how to apply basic information presentation principles. These processes are consistent regardless of domain and therefore are not directly addressed in a CTA.

Design Content Themes to be Addressed through CTA

1. *Common Ground*. Support for common ground means pertinent beliefs, assumptions, and intentions are shared among team members. Common ground is constantly and actively updated and maintained as team members maintain a shared picture of what's happening in the world and the status of the overall plan (Bradshaw et al., 2004; Klein, Feltovich, Bradshaw, & Woods, 2005). Teammates engage in proactive backup behavior and proactive communications (Johnson, Bradshaw, Hoffman, Feltovich, & Woods, 2014).
2. *Observability*. Observability is defined as transparency into what an automated partner is doing relative to task progress. Observability can be best achieved by supporting shared understanding of the problem to be solved and progress towards goals (Bradshaw et al., 2004; Joe, OHara, Medema, & Oxstrand, 2014). This understanding must be provided early in use in a way that is comprehensible to operators, and it must be informative enough to enable effective intervention (Hoff & Bashir, 2015; Parasuraman & Riley, 1997).
3. *Predictability*. Future intentions and activities of the partner are discernible and understandable. It is a subset of observability, and based on an understanding of the automation's goals, abilities, and limitations (Klein et al., 2005). Supporting predictability aids the user in anticipating changes in the system and gives them an idea of where to look next (Bradshaw et al., 2004; Lyons, 2013).
4. *Directability*. Directability is supported when humans are enabled to direct and redirect an automated partner's resources, activities, and priorities. The people on the scene will inevitably know about information associated with the particular situation that the automated system will have no knowledge of (Roth et al., 1987). Ultimately then, the human must be in control and be able to stop the process, toggle between levels of autonomy, or override and manually control the automation (Joe et al., 2014; Lyons, 2013).
5. *Directing Attention*. Each partner must be able to direct the attention of the other partners to critical problem features, cues, indications, and warnings. Automation must be able to flag additional and relevant information it cannot process so the human can redirect. It should also communicate proactively when information it has becomes relevant, such as when goal obstacles are encountered (Joe et al., 2014).
6. *Exploring the Solution Space*. Exploring the Solution Space is supported when partners can leverage multiple views, knowledge, and solutions to jointly understand the problem space. Automation should be able to rapidly generate multiple distinct courses of action and give people ways to rapidly compare those solutions. Both humans and technology should be able to expand or constrict the solution considerations and shift perspectives (Woods & Hollnagel, 2006).
7. *Adaptability*. Adaptability is defined as partners being able to recognize and adapt fluidly to unexpected characteristics of the situation. Supporting this requires multiple options to both recognize an unexpected situation and to address and recover from it (Johnson et al., 2015).
8. *Calibrated Trust*. Calibrated trust is supported when human operators have a strong understanding of when and how much to trust an automated partner in context. Users should understand clearly when automation can be relied upon, when more oversight is needed, and when performance is unacceptable (Joe et al., 2014). Calibrating trust requires providing information on the source of the diagnostic information, and the credibility of the alert (Johnson et al., 2014).

Design Process Themes (not addressed through CTA)

9. *Information Presentation*: This category was created to capture published recommendations on how to present information to support simplicity and understandability.
10. *Design Process*. Elements of the cognitive engineering-driven design process of systems engineering that must be incorporated to enable effective HMT.

Cognitive Interface Requirements

Cognitive Interface Requirements were drafted for each theme. These themes go beyond informaiton presentation to address how the automation and the human(s) can work together to achieve task and mission goals.

The requirements are purposefully general in nature. The goal is to provide a comprehensive general set of requirements that can be tailored to specific domains. Toward this end, the requirements contain rationale - italicized statements that convey the purpose behind the requirement – so the rationale is not lost when the requirements are tailored. Each requirement is tagged with a numbers that link it to the relevant research article(s) that inspired it. The requirements have hyperlinks (shown in blue underlined text) which provide an example or more details about the requirement.

Observability Example:

- The system shall provide information about how automation activities may evolve in the future. [18] Example: Auto GCAS merging chevrons

Calibrated Trust Example:

- The system shall convey system reliability in different contexts *to promote calibrated trust*:
    - situations in which the automation can be relied upon;
    - situations that require increased oversight by personnel;
    - situations in which the automation's performance is not acceptable. [3]

The set of general requirements can be tailored for specific systems. Designers can identify which requirements in the general set are applicable to a given system. For example, if the envisioned automation is for a commander's virtual assistant, the requirements on remote perception may not be applicable. Likewise, the requirements relating to what-if analysis may not be applicable for small autonomous vehicles. Another way to tailor the requirements is to rewrite them using more specific language. Consider this requirement for Exporing the Solution Space:

- The automation shall allow a human to input information that wasn't initially taken into account by the algorithms *so the algorithm can incorporate new information that the human has or expertise of the human* [74].

For a new automated planning system with a logistics tool for calculating battery and fuel expenditures, it could be rewritten as follows:

- The automation shall allow a logistician or commander to input their knowledge of troop speed, given the expected impact of casualties and the weight of supplies to be transported. This information will be used in the calculation of the fuel expenditure.

[3] Joe, J., O'Hara, J., Medema, H., & Oxstrand, J. (2014, June).

[18] Christoffersen, K., & Woods, D. D. (2002). How to make automated systems team players. Advances in human performance and cognitive engineering research, 2, 1-12.

[74] Scott, R., Roth, E., Truxler, R., Ostwald, J., & Wampler, J. (2009, October). Techniques for effective collaborative automation for air mission replanning. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 53, No. 4, pp. 202-206). Sage CA: Los Angeles, CA: SAGE Publications.

CTA Methods focused on HMT

The first eight HMT themes are ripe for for exploration via CTA with subject matter experts (SMEs). We conducted preliminary work to develop a CTA methodology to address HMT. We identified building blocks for knowledge elicitation techniques and developed a technique to tailor the scope and focus of a knowledge elicitation to elicit critical HMT requirements.

After a survey of CTA methods, we concluded that two knowledge elicitation techniques are ripe for modifications that would enable addressing the eight HMT requirements themes which require SME input. The Critical Decision Method (Hoffman, Crandall, & Shadbolt, 1998) can capture expert insights on how they and their teammates performed tasks in past situations (either with or without automation) and can be used to follow the thread of how a situation would be different with automation. The Knowledge Audit (Militello & Hutton, 1998), a survey method of elicitation, could be used to probe specific cognitive elements or themes and may be well-suited to envisioned systems. We plan to use the Critical Decision Method and the Knowledge Audit as starting points but to tailor them to address the HMT themes. For exmaple, the Big Picture probe in the Knowledge Audit could be modified. Instead of asking about the major elements of the situation that you need to know and keep track of, the question could ask about the major elements of "what the automation is doing," that need to be known to understand the situation. This would address Observability and could lead to follow-up questions on Calibrated Trust. A modified Knowledge Audit paired with the Critical Decision Method could allow researchers to elicit information to tailor the requirements and characterize the HMT interactions and the challenges.

Addressing eight HMT themes may be too large a scope for a single CTA interview. To help focus the interview, we are developing a HMT Triage technique. The idea is to focus on the themes that that warrant the most attention. A theme may warrant attention because it is prevalent (e.g., the automation performs numerous calculations and assessments and there is a question about how many of the assumptions and information sources should be observable to the human), it is challenging (e.g., to enhance common ground, the automation needs just in time information from the human but system developers are reluctant to distract the human from his or her main task), or because it is critical to success (e.g., at different points in a mission, the unmanned aerial vehicle operator needs to direct his or her attention to different pieces of information in order to meet dynamically changing surveillance goals). Some of the triage questions are technology based. For example, "Is the automation in motion (actively controlling something in the environment, such as where a telescope is pointing) or at rest (operating virtually, such as a planning system)?" The implication is that autonomy in motion may warrant more focus on Observability so the human can see the effects on the object being controlled and Directability because the human may need to intervene. Other triage questions are designed to help the interviewee determine if the theme is likely to be applicable and a challenge. For Common Ground, the interviewee could ask if beliefs and assumptions are likely to change during a mission. The goal of the HMT Triage is narrow the focus to 3-5 HMT themes. This is not to say that other themes are not important, it simply acts as a screening technique to focus on the areas most in need of cognitive support. The questions in the HMT Triage can be used prior to an interview if the interviewing team has access to background material, SMEs, or system developers than can help answer questions. Otherwise, it can be used with a SME at the beginning of an interview.

CONCLUSION

As the prevalence of autonomous and automated components of military systems increases, issues with HMT often surface and can impact trust in technology and create a barrier to operational acceptance. This effort bridges the gap from research to technology development to better enable systems that synergize the joint strengths of humans and technology towards asymmetric advantages. These research efforts are critical for complex systems to go beyond decision aiding to achieve true teaming in which the joint performance of human and machine is greater than either alone.

Our next step is to develop a set of tailored requirements in two domains. To do this, we will tailor CTA methods to specifically address the themes in HMT. We will document the CTA methods used and develop a "how to" guide for knowledge elicitiation with experts and using the results to tailor cognitive interface requirements for HMT. The result will be a repeatable method that development teams can apply to ensure their systems support human-automation teaming.

REFERENCES

Bradshaw, J. M., Acquisti, A., Allen, J., Breedy, M. R., Bunch, L., Chambers, N., & Johnson, M. (2004, March). Teamwork-centered autonomy for extended human-agent interaction in space applications. In *AAAI 2004 Spring Symposium*, 22-24.

Hoff, K. A., & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, 57(3), 407–434.

Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998). Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(2), 254-276.

Joe, J., O'Hara, J., Medema, H., & Oxstrand, J. (2014, June). Identifying requirements for effective human-automation teamwork. In *Proceedings of the 12th international conference on probabilistic safety assessment and management* (PSAM 12, Paper# 371),(INL/CON-14-31340).

Johnson, M., Bradshaw, J. M., Hoffman, R. R., Feltovich, P. J., & Woods, D. D. (2014). Seven Cardinal Virtues of Human-Machine Teamwork: Examples from the DARPA Robotic Challenge. *IEEE Intelligent Systems,* 74-80.

Johnson, M., Shrewsbury, B., Bertrand, S., Wu, T., Duran, D., Floyd, M., et al. (2015). Team IHMC's Lessons Learned from the DARPA Robotics Challenge Trials. *Journal of Field Robotics*, 32(2), 192–208.

Klein, G., Feltovich, P. J., Bradshaw, J. M., & Woods, D. D. (2005). Common ground and coordination in joint activity. *Organizational simulation*, 53.

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(1), 50–80.

Lyons, J. B. (2013). Being Transparent about Transparency: A Model for Human-Robot Interaction. In *AAAI Spring Symposium Series*.

Militello, L. G., & Hutton, R. J. (1998). Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, 41(11), 1618-1641.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, 39(2), 230–253.

Sarter, NB, Woods, DD, & Billings, C. (1997). Automation Surprises. In G. Salvendy (Ed), *Handbook of human factors and ergonomics,* 2nd Edition. Wiley.

Woods, D., & Hollnagel, E. (2006). *Joint Cognitive Systems*. CRC Press.

Wiener, E.L. (1989). Human factors of advanced technology ("glass cockpit") transport aircraft. (NASA Contractor Report No. 177528). Moffett Field, CA: NASA Ames Research Center.