Identifying Requirements for Effective Human-Automation Teamwork

Jeffrey C. Joe[a*], John O'Hara[b], Heather D. Medema[a], and Johanna H. Oxstrand[a]
[a] Idaho National Laboratory, Idaho Falls, ID, USA
[b] Brookhaven National Laboratory, Upton, NY, USA

Abstract: Previous studies have shown that poorly designed human-automation collaboration, such as poorly designed communication protocols, often leads to problems for the human operators, such as: lack of vigilance, complacency, and loss of skills. These problems often lead to suboptimal system performance. To address this situation, a considerable amount of research has been conducted to improve human-automation collaboration and to make automation function better as a "team player." Much of this research is based on an understanding of what it means to be a good team player from the perspective of a human team. However, the research is often based on a simplified view of human teams and teamwork. In this study, we sought to better understand the capabilities and limitations of automation from the standpoint of human teams. We first examined human teams to identify the principles for effective teamwork. We next reviewed the research on integrating automation agents and human agents into mixed agent teams to identify the limitations of automation agents to conform to teamwork principles. This research resulted in insights that can lead to more effective human-automation collaboration by enabling a more realistic set of requirements to be developed based on the strengths and limitations of all agents.

Keywords: Team Performance, Human-Automation Teams, Human Factors Issues with Automation.

1. INTRODUCTION

As the role of automation expands in new and advanced systems, one goal of research is to make automation a team player [1, 2, 3]. One motivation behind this effort is the finding that humans relate to automation in similar ways to the way they relate to human teammates [4]. However, this approach typically uses incomplete models of human teamwork in the design of human-automation teams. Furthermore, we assert that automation agents cannot currently behave and interact with humans in the same way other humans can. These assertions are logical inferences from past research that has found that human-automation teams, that are designed based on an incomplete understanding of what makes good human teams effective, often lead to problems for operators, such as:

- Undesirable changes in the overall role of personnel
- Difficulty understanding automation
- Poor monitoring, lack of vigilance, and complacency
- Out-of-the-loop unfamiliarity and situation awareness
- Workload to interact with automation and when transitioning to greater manual control
- Loss of skills for performing tasks automation typically performs
- New types of human error

We use the term "multi-agent" teams to refer to teams having both human and automation agents who work cooperatively to accomplish tasks and plant functions. Designing automation to be a good "team players" has typically based on an implicit notion of what it means to be a team player and how members of a team should perform to function successfully. General concepts of team characteristics and behavior are employed, such as trust, goal and intention sharing, cooperation, and redundant responsibilities (especially in the case of adaptive automation where shifting of responsibility is a hallmark of the approach). These concepts are based loosely on a sense of how human teams perform.

[*]Corresponding Author: Jeffrey.Joe@inl.gov

However, the concept of multi-agent teamwork relies considerably on simplifications and popular notions about what is needed to foster teamwork, which do not always transfer well to human-automated agent teams. That is, the work to define how automation should behave (be designed) to be a team player has not been based on the recognition that there are some fundamental differences between human behavior and how automation has been programmed to behave, and that there are different models of human teamwork, each with its own set of member responsibilities and behaviors [5]. As a result, prior work on identifying how automation can be a team player is fragmented and incomplete at best.

Further, even if a belief about how human teams behave is at the core of this research, it does not sufficiently address the fact that automation agents are not humans, and cannot completely fulfill the role of a crewmember. Nor can we expect that it will fully behave as a human member of a team will. For example, automation agents cannot assume responsibility. Automation can be given the authority to act, but humans always maintain responsibility [6]. As another example, automation is not "concerned" about the consequences of its actions, nor is it as able to innovate as human crews will do when things do not go as planned.

The objective of this research is to identify the principles for effective human-automation collaboration that are appropriate in a commercial nuclear power environment involving a team of agents, specifically one where the anticipated use and/or level of automation is high. Once identified, we can elaborate on the specific requirements of human and automation team members, and can use those requirements to develop guidance for what automation characteristics and attributes are needed to be good "team players."

This paper will present some of our findings related to identifying the important characteristics of human teams, identifying the approaches taken to create multi-agent teams, and our preliminary approach to the formulation of general principles for effective human-automation teams.

2. METHODOLOGY

Our methodology is based primarily on an analysis of the literature. We first examined the research on what are the characteristics of effective teams working in complex systems, with an emphasis on the commercial the nuclear domain. Based on this research, we identified the general principles for effective teamwork. We next examined the research on mixed-agent team, specifically human-automation teams. We examined this literature from two perspectives: First we reviewed the research examining the characteristics of mixed-agent teams. Then we examine the research focused on defining principles of human-automation teamwork. This literature provided us with some understanding of the necessary characteristics of mixed agent team as well as an understanding of the limitations and issues. Then we examined how human-automation teams are different from human teams. To do this we compared the principles of effective human teams to what we know about the characteristics and issues associated with mixed-agent teams. Finally we began the process of identifying the principles for successful automation agent participation in teams. We are currently in the process of developing these principles, which will be further developed into design requirements as the research progresses.

3. RESULTS

3.1. Characteristics of Human Teams

3.1.1. Research on Human Team Effectiveness

The literature on teams composed only of human members is varied in terms of the genres that have studied this subject, and in terms of the depth of analysis that has been on performed to understand how to improve human team performance. Following a brief discussion about research on teamwork in general, we will focus on teamwork in nuclear power plant (NPP) crews.

A considerable amount of research in the business leadership and human factors literatures has been conducted on teams and team performance. The human factors literature has studied team cognition and team collaboration in particular. One popular book from the business leadership literature [7], presents "five dysfunctions" that can undermine team performance: (1) absence of trust, (2) fear of conflict, (3) lack of commitment, (4) avoidance of accountability, and (5) inattention to results. All of these factors make intuitive sense and are backed by other research studies. If team members do not trust one another, are afraid of conflict, lack commitment, etc., the likely outcome is suboptimal team performance. The human factors research on cognitive models of team and teamwork is based on the premise that in addition to the factors that [7] identified, how the team collectively thinks and coordinates cognitive and physical efforts are important elements to team success, but that it is also challenging for humans to do this without guidance and practice. A number of models of team have been developed from the human factors perspective to help elucidate this premise. Perhaps the most well known model of teams is the Crew Resource Management model [8], but there have been a number of other important and insightful contributions including the Teamwork Model [9], work on a meta-cognitive and macro-cognitive model of team collaboration [10], team sensemaking [11], and the Mutual Belief Model [12]. One key insight this literature has revealed is that it is important for team members to know what other people on the team are thinking and doing (i.e., situation awareness about team processes), along with knowing what is going on in the system and surrounding context (i.e., global situation awareness). This research also generally shows that people on teams tend to think that maintaining good global situation awareness is sufficient to also having good situation awareness about team processes. That is, individuals on teams tend to discount the importance of having good situation awareness specifically focused on what other teammates are thinking and doing, and believe that their awareness of team processes can be adequately maintained by having good global situation awareness. Furthermore, research findings have shown that people's habits, assumptions, complacency, and reliance on established conduct of operations standards tend to contribute to people losing situation awareness of other team members. The result of this typically leads to sub-optimal team performance, and sometimes leads specifically to teamwork errors, such as failing to do peer checks or independent verifications, which can subsequently lead to system-wide failures. Helpful reviews this line of human factors research can be found in [13] and [14].

In today's commercial NPPs, several reasons dictate the need for teams, including the distribution of workload, and the physical layout of the control room where tasks are completed via dozens of control boards. Thus, the successful operations of commercial NPPs are accomplished through the coordinated activity of multi-person teams or crews, which has lead to specialized training for NPP crews, and increased focus and rigor on the requirements for becoming a licensed operator. The Institute of Nuclear Power Operations (INPO) is one of the main industry drivers for NPP control room operator training and licensing in that they provide input on the requirements and standards for operator training and licensing. According to [15], along with the expected requirements that operators have the fundamental knowledge of technical topics such as nuclear physics and reactor theory, electrical science, and chemistry, there are guidelines on teamwork, and related interpersonal issues such as control room supervision and operational decision-making [16]. These INPO reports and guidelines cover a variety of teamwork issues including, but not limited to, communications, interaction among team members of different personality types, self- and peer-checking, and constructive conflict management [15]. INPO expects licensed reactor operators to be well versed in these aspects of teamwork as they recognize that teamwork is an important facet of the ability for the crew to function effectively, particularly in novel or emergency situations. According to [15], INPO recommends that teamwork training not only be taught in the classroom, but "should be continually reinforced during day-to-day work and training activities." (pg. 28).

There are also a number of studies of NPP crews that provide additional insights into the specific challenges crews face. Research has shown that factors such as: task characteristics, team member characteristics, and team dynamics affect how successfully the crew performs [17]. Other researchers has shown that successful teams monitor each other's activities, back each other up, actively identify errors, and question improper procedures [18]. Another study showed that communication protocols, communication content, and communications errors in NPP crews vary depending on the philosophy underlying the control room's design and the implementation of instrumentation and controls technologies [19]. Research by [20] studied how characteristics of communication among NPP crews during simulated emergencies affected crew performance, and found that characteristics such as: (1) a tightly coupled communication structure (which is an indicator of good team cohesion), (2) increasing the amount/density of communication to increase team situation awareness (e.g., crew members speaking up when observing changes in the system state), and (3) increasing the thoroughness of communication to make shared understanding more explicit (e.g., greater adherence to three-way communication practices) improved overall crew performance during simulated emergency scenarios. Clearly, important human factors aspects of teamwork in NPPs include having common, coordinated goals, maintaining shared situation-awareness, engaging in open and effective communication, and cooperative planning. In summary, existing NPPs are highly complex systems that require teams of human operators to be well trained on both the technical aspects of operations as well as the 'soft-skills' of effective teaming. This literature reviewed showed that in order for NPP operators to perform well as a team, they need to be both technically proficient, and well trained on how to manage the coordination of both their cognitive processing and physical efforts.

3.1.2. General Principles for Effective Human Teams

Based on all of the literature reviewed, we have derived a set of general principles for effective human teamwork. They are summarized below.

Belief in the Concept of Team: Effective human teams have individuals that believe in (or they are told that their compensation depends on) the idea that there is a mutual benefit to working together. This is also referred to as people having a team orientation versus an individualistic orientation [9].

Effective Communication: High performing teams communicate effectively. Communication is the central behavior team members engage in to function as a team. It is central to exchanging information, establishing team situation awareness, coordinating and regulating individual efforts, and building trust among team members.

Team Leadership: Effective teams have good leadership. Leadership by individuals formally appointed as leaders, and informal leadership by other team members, must principled and consistent, and needs to have both transactional and transformational qualities.

Monitoring Individual and Group Performance and Providing Feedback: Effective teams monitor and provide feedback on both overall group performance and individual contributions. This is key to the systematic adjustment of coordinated team activities. In NPPs, these coordinated team activities are needed to make adjustments to system-level process parameters (that are within the span of control of operators) that are necessary to achieve the desired process outcomes (e.g., operate the NPP safely and generate electricity profitably). Monitoring individual contributions to overall performance is also important to mitigating a well-known social psychological phenomenon called social loafing.

Coordination and Assistance: Effective teams are well coordinated and provide assistance to one another. When this happens, performance above and beyond what each individual could achieve on their own can be realized. Included in the principle of coordination is the extent to which team members have overlapping capabilities and are willing and able to seek and provide assistance to one another (i.e., provide redundancy or defense in depth) and increase overall system resiliency.

Awareness of Internal and External Performance Shaping Factors That Affect Team Processes: Effective teams are aware that personality traits and cognitive abilities can vary from person to person. How the team leverages, and not just manages, individual differences to enhance team performance is important. Similarly, external performance shaping factors (PSFs), such as task complexity, time pressure, and information/knowledge uncertainty can affect the team's ability to collaborate, which in turn can affect overall system performance.

<u>*Awareness that each individual's mental model is unique and that it is difficult to create shared mental models in teams:*</u> Effective teams recognize that everyone has an idiosyncratic mental model of their environment, and that each team member must clearly communicate their mental model to others. That is, humans construct their mental models based on an attentionally constrained ability to detect and process external stimuli, and a less than perfect long-term memory system, which leads to the formation of an idiosyncratic understanding (i.e., mental model) of the situation. Furthermore, humans have difficulty communicating or sharing their mental models with other humans such that they and one or more team members are subsequently 'on the same page of the playbook.'

Generally speaking, most of the issues with human teams stem from fundamental human frailties that define a large part of what it means to be human. Humans have beliefs, motivations, and emotions that affect their performance. Human team performance is affected by numerous factors including: the quality of communication, the extent to which they trust others, the morale of the group, the quality of leadership, and how well coordinated they are in monitoring individual and team performance, providing feedback, and timely assistance. Team performance is also affected by significant between-person differences in values, and mental and physical abilities. External situational factors further affect individual and team performance. And finally, the mental models humans develop of complex situations are idiosyncratic, often inaccurate, and are difficult to communicate to others such that they all have the same understanding. These are frailties that automated agents do not necessarily share.

In summary, the literature reviewed on human teams provides some good insights on the challenges and issues with these teams, and often provides a range of practical and workable solutions. The extent to which these insights and principles translate to human-automation teams, and automated agents in particular, however, will vary depending on how applicable the principle is to multi-agent teams, and specifically to the automated agent.

3.2. Characteristics of Multi-Agent Teams

3.2.1. Research on the Integration of Automation Agents into Human Teams

Klein, Sycara, and colleagues discussed integrating agents and human in teams [21, 22]. They viewed teamwork from the human team perspective of Sycara and Lewis [22]. They identified the four critical dimensions: information exchange, communication, supporting behavior, and team initiative/leadership. They stated that the human-automation interaction requirements differ based on the role of the automation in the team. They identified three different types of contributions automation can make as part of a team: (1) to work on tasks independently of human teammates, (2) to collaborate with human teammates and support human task performance, and (3) to support teamwork processes such as facilitating communication and coordination of human team members.

Sycara and Lewis noted that the greatest obstacle to integrating machine automation agents with human teams is communicating a human's intent. Furthermore, as automation becomes more flexible and autonomous, it becomes more difficult for humans to monitor and evaluate its performance. They identified three key factors in human-automation interaction: Mutual predictability of teammates, team knowledge (shared understanding), and the ability to redirect and adapt to one another. Agent predictability and shared understanding is made more difficult because automation often does not communicate its intent. They suggest that predictability can be fostered by: (1) consistently pairing simple observable actions with inputs, (2) making the causes and rules governing an agent's behavior accessible to the human, and (3) making the purpose, capability, and reliability of the agent known to the human.

Predicitability is related to trust. In the Introduction, we noted that humans relate to automation similarly to how they relate to human teammates. It is not surprising therefore, that operators' trust in automation greatly impacts their use of it [23]. Trust in automation is based on the operator's perception of its reliability and capability. This perception may or may not be consistent with reality.

When the operator's perceptions accurately match the automation's reliability and capabilities, trust is "well-calibrated" and operators use it appropriately. Miscalibrated trust leads to either an overreliance on automation (i.e., misuse) or its underutilization (i.e., disuse). Misuse (i.e., overreliance) on automation is associated with a failure to properly monitor it. This can engender large errors in system performance, (e.g., the system deviates considerably from the desired performance before it is recognized, if it is recognized at all). Disuse (i.e., underutilization) of automation can lead operators to turn off automation or to ignore its potential benefits.

Directability and mutual adaptation are important aspects of teamwork and enable teams to be flexible in different task contexts. Sycara and Sukthankar define directability as "assigning roles and responsibilities to different team members" and mutual adaptation as "how team members alter their roles to fulfill the requirements of the team" [21]. Researchers acknowledge that the most effective agents change their level of initiative, or exhibit adjustable autonomy, in response to the situation. For agents to appropriately exercise initiative, they must have a clear model of the team's goals, member roles and team procedures.

Communication between agents is key to achieving mutual predictability, shared understanding, and an ability to redirect and adapt. Communication structure between automated and human team members, however, is significantly different than typical human communication. While human-to-human communication includes face-to-face communication and nonverbal cues, automation is limited to transmitting and receiving information via specific formats within an array of displays [14]. Suchman attributed problems with modern automated system to the limits of syntax and semantics in automation [24]. The automated system is at the mercy of its programming, specifically, the types of information the system can provide and request. These constraints limit the human teammate, who must anticipate how to acquire and format the information exchanges with the automation agent. This sharing of information is critically different than the "spontaneous, free-flowing nature of human communication" [14]. Good communication, therefore, is a cornerstone of achieving effective human-automation interaction.

Woods et al. also noted that designers often fail to appreciate the need for humans and automaton to interact even when automation is designed to be mostly autonomous [25]. One reason for this is that designers often underestimate the complexities of operational environments.

Fiore et al. suggested that human-agent team interaction introduces a number of issues with respect to both team and individual cognition such as humans having to deal with an increased level of abstraction that may place unique demands on their information processing abilities, and increased difficulties coordinating team members given that differing communication patterns may be necessary to share cues [26]. Fan and Yen made a similar observation. They note that human teams rely on global situation awareness and shared mental models [27]. When automation agents are involved, developing this awareness and shared mental models can create additional costs attributed to communication, resolution of conflict, and social acceptance.

Pritchett's and colleagues identified ways in which automation agents are different from human agents [28, 29, 30]. Human team members will continue to attempt effective performance in unfamiliar circumstances, while automation generally cannot (e.g., behavior when outside boundary conditions). Good teammates anticipate each other's information needs and provide information (e.g., anticipating the needs of a teammate). Automation is limited in this capability. Humans time their interactions based on an awareness of the current situation, such as another teammate's workload (e.g., managing interruptions). Automation can be "clumsy" in this regard and interrupt human teammates at inopportune times. Finally, humans have a sense of responsibility. Automation does not have motivation, or a sense of responsibility.

Steinberg identified two significant limitations in multi-agent, or human-automation teamwork. First, automation cannot fully capture the human operator's intention in performing tasks [3]. Second, automation cannot flexibly adapt to situations that have not been considered by the designer.

Similarly, human operators often do not have a good understanding of how automation functions and find its interactions disruptive. Thus, Steinberg concludes that it may not be realistic, in the near-term at least, to think automation can be a teammate in the sense that another human operator is. Steinberg suggests that to create more effective multi-agent teams, researchers need to think more broadly about human-machine relationships. Other researchers have made similar observations [26].

Automation agents also affect one person's performance with another person. Roth and O'Hara observed the introduction of a computer-based emergency operating procedures (EOPs) system as part of one utility's digital instrumentation and control upgrade of a NPP [31]. This system automated the functions of information acquisition and analysis and gave some support to decision-making, such as retrieving data and assessing its quality, resolving step logic, and tracking the location in the procedure. The crews handled disturbances on a training simulator. Following each scenario, their interviews focused on the impact of the procedures on operations. They found that by introducing the new system, the procedure management workload was reduced to the point that procedure use became a one-person activity. The board operators were far less engaged in this, except to take occasional control actions at the request of the supervisor. Consequently, the operators felt they were out-of-the-loop, had lost situation awareness of EOP activities, and were unsure what to do. Thus, the introduction of the automation impacted teamwork, a finding that was unanticipated by the designers and plant staff.

Wright and Kaber evaluated the effects of automation on the performance and coordination of teams in a complex decision making task when it was applied to different stages of information processing [32]. Two-person crews performed a simulated mission to protect a home base from enemy attack. They found the effects on teamwork differed based on the generic tasks that were automated. The findings suggested that automation of early and intermediate stages of processing may have benefits with respect to teamwork, while automation of decision selection may be more limited, in that its benefit depends on the context in which it is used.

In summary, this research suggests that a strict human teamwork model may not be appropriate when designing human-automation teams. The limitation of current and near term automation agent capabilities and automation-human interaction include:

- Difficulty establishing shared mental models (i.e., shared understanding) of the situation when human and automation agents are part of the same team
- Failure of both agents to know or anticipate the other's intentions, actions, and overall team goals
- Limited flexibility of automation agents to redirect activities and adapt to shifting needs of the team and novel situations (outside the boundary conditions of its programming)
- Limited interaction between human and automation agents when the level of automation is high
- Clumsy and potentially disruptive effects on teamwork, roles, and responsibilities of the introduction of automation into human teams
- Automation agents have no sense of responsibility, conflict resolution, or social acceptance
- Increase in human workload when the requirements of interacting with automation are high
- Poor communication between agents and difficult constraints on the ease with which humans can communicate with automation

3.2.2. General Principles for Effective Human-Automation Teamwork

Many authors used human teamwork as a model to identify the general characteristics of desired human-automation interaction [4, 33, 34, 35]. Based on a review of studies of multi-agent teamwork, O'Hara and Higgins derived several general principles for human-automation interaction applicable to the commercial nuclear industry [23]. These principles are presented in Table 1.

Table 1: General Principles for Supporting Teamwork with Machine Agents

Principle	Definition
Define the purpose of automation	Automation should have a clear purpose, meet an operational need, be well integrated into overall work practices, and be sufficiently flexible to handle anticipated situational variations and adapt to changing personnel needs.
The designer should establish locus of authority	In general, personnel should be in charge of the automation, i.e., be able to redirect, be able to stop it, and assume control when necessary. This does not preclude the automation from initiating actions. Some actions are allocated to automation because they cannot be reliably performed by personnel within time or performance requirements. Further, there may be situations where automation initiates a critical action because personnel have failed to do so.
The designer should optimize the performance of the human-machine team	The allocation of responsibilities between humans and automation should seek to optimize overall integrated team performance. This may involve defining various levels of automation, each with specific responsibilities for all agents and each. It also may involve flexible allocations that change in response to situational demands. Personnel's interactions with automation should support their development of a good understanding of the automation, and the maintenance of their skills needed if automation fails. The HSIs should support a clear mutual understanding of the roles and responsibilities for both human and automation agents.
Personnel should understand the automation	Personnel should understand automation's goals, abilities, and limitations; and be able to predict its actions within various contexts of use. Minimizing automation's complexity will support this objective. While their understanding largely will come from training and experience, the HSI should support that understanding by reinforcing the operators' mental model through the information provided in automation displays. That is, the HSI should accurately represent how automation performs and how it interacts with the plant's functions, systems, and components.
Personnel should trust the automation	Personnel should have appropriately calibrated trust in automation that involves knowing the situations when the automation can be relied on, those which require increased oversight by personnel, and those for which automation's performance is not acceptable. The HSIs should support trust calibration by providing automation's reliability in various contexts of use.
Personnel should maintain situation awareness	The HSIs should provide sufficient information for personnel to monitor and maintain awareness of automation's goals, current status, progress, processes (logic/algorithms, reasoning bases), difficulties, and the current responsibilities. Special attention should be given to changing levels of automation where the responsibilities of agents may change.
The HSI should support interaction and control	Personnel interaction with automation should support their supervisory role: • HSIs should support personnel's interaction with automation at a level commensurate with the automation's characterization, e.g., level, function, flexibility, and its reliability. • Communication functions should enable personnel to access information about automation's processes. Automation should communicate with personnel when necessary, such as when it encounters an obstacle to meeting a goal, or when information is needed from personnel (e.g., information not accessible to automation). Communications from automation should be graded for importance, so as not to be overly intrusive. • Personnel should be able to redirect automation to achieve operational goals and should be able to override automation and assume manual control of all or part of the system.
The HSI should minimize workload	The HSI design should minimize the workload required to interact with automation, e.g., to configure automation and to communicate with it.
The team should manage failures	The multi-agent team should support error tolerance and failure management: • Personnel should monitor the activities of automation to detect automation errors, and to be sufficiently informed to assume control when automation fails. • Automation displays should support operators in determining the locus of failures as being either the automation or the systems with which the automation interfaces. • Automation should monitor personnel activities to minimize human error by informing personnel of potential error-likely situations. • When situations change sufficiently to render automation's performance unacceptable, it should communicate the situation to personnel in a timely way to enable them to become more engaged in automation's current goals and responsibilities.

Note: From O'Hara and Higgins [23]

3.3. Comparison of Human and Multi-Agent Teams

The well-known dynamics of human teams fundamentally change when teams are comprised of human and automated agents. That is not to say that none of the literature on human teams is applicable to the new human-automation team dynamic. Rather, the change in team dynamics means that the applicability of the insights from this literature on the new problem space defined by the pairing of humans and automated agents is not simple and straightforward. A corollary to this is that the problems associated with the differences between humans and automated agents, and the problems these differences further create in teams, are not directly addressed by this literature's insights and solutions. Additional thought must be applied if the valuable insights from the literature are going to add any value or meaningful guidance to designers contemplating the use of human-automated teams to operate their systems.

This section compares how human-automation teams are different from human teams. The approach for this comparison was to use the principles of effective human teams from Section 3.1.2 above as the standard or set of principles for successful team performance, since much of the automation literature presumes that this is what automation should be like in order to be a good "team player."

We can gain additional insight by comparing the seven principles of effective human teams with the limitations we identified in human-automation teams. This is shown in Table 2.

Table 2: Comparison of Human Teaming Principles and Human-Automation Limitations

Human Team Effectiveness Principles	Limitations of Human-Automation Teams
Belief in the concept of team	• An automation agent's behavior or performance does not change with an appeal to a belief in the concept of team. It cannot work more or less than its predetermined programming • Automation agents have no sense of responsibility, conflict resolution, or social acceptance • Potentially disruptive effects on teamwork, roles, and responsibilities of the introduction of automation into human teams
Effective communication	• Poor communication between agents and difficult constraints on the ease with which humans can communicate with automation • Clumsy and disruptive interactions between agents • Limited interaction between human and automation agents when the level of automation is high
Team leadership	• There is no concept of leadership in automation agents. They are indifferent to the quality and style of team leadership • If automation is ever in a leadership position, it is unlikely that the automated leader will effectively convey these leadership 'soft skills' that would improve the morale and confidence of the human subordinates
Monitoring individual and group performance and providing feedback	• Limited ability to both human and automated agents to know or anticipate the other's intentions, actions, and overall team goals
Coordination and assistance	• Limited flexibility on automation agents to knowing when the human needs additional assistance, and redirect activities and adapt to shifting needs of the team and novel situations (outside the boundary conditions of automation)

Human Team Effectiveness Principles	Limitations of Human-Automation Teams
Awareness of internal and external performance shaping factors that affect team processes	• Automation agents do not feel anxiety (i.e., internal PSF), time pressure (i.e., external PSF), or the interaction of PSFs in the same way humans do • Increase in human workload when the requirements of interacting with automation are high
Awareness that each individual's mental model is unique and that it is difficult to create shared mental models in teams	• Difficulty establishing shared mental models (shared understanding) of the situation when human and automation agents are part of the same team

What this comparison shows is that these principles do not translate well to human-automation teams, primarily because automation agents do not have many of the same key 'soft' characteristics that humans do. This is further exacerbated by the fact that humans know that automation does not have these characteristics, frequently try to compensate for this shortcoming, but nevertheless have trouble adopting new and effective ways of teaming with automated agents. In short, this comparison shows that when viewed from the perspective of what constitutes effective human teams, automation may find it difficult to be a "good teammate" in the same sense as a human teammate can be.

3.4. Developing Requirements for Multi-Agent Teams

Prior research has suggested some general principles for multi-agent teams (see Table 1) that can be developed further into preliminary requirements. However, these principles are based on a simplified model of teamwork that does not fully address the complexity of human teams, their processes, or the effects of automation agents on human team processes. Our analysis was based on a comprehensive consideration of the principles for effective human teamwork. This analysis shows that, at least as far a near-term technology is concerned, significant limitations exist in automations capabilities to fulfil the role of a good teammate.

Such a conclusion is not a showstopper. Instead, we view our findings as an opportunity to define reasonable requirements for integrating humans and automation to accomplish work cooperatively and in a manner that recognizes and capitalizes on the strengths and weaknesses/limitations of all agents in the team. The design of automation teammates should address these differences. Where agent capabilities can mimic those of a human teammate, principles for doing so can be developed. For those teammate characteristics that agents cannot mimic, principles for designing alternative approaches for accomplishing the characteristic need to be developed. In short, automation agents affect human team performance and this has to be addressed in the design. Further, operator training and experimentation will need to establish and reinforce a realistic view of automation's role in NPP monitoring and control as part of a multi-agent team.

Defining reasonable requirements is the next phase of our research. Using the general principles in Table 1 as a start, we will develop requirements for human-automation collaboration that realistically considers the capabilities and limitations of all agents on the team.

4. CONCLUSIONS

The human factors literature on human-automation teams has had the implicit belief that the way to make automated agents better team players is to translate the best practices and principles for effective human teams and use them to design human-automation interactions. This analysis shows, however, that many key principles for effective human teams do not translate well to automated agents, primarily because there are inherent differences between humans and automated agents. These differences have had significant consequences on human-automation team performance. Clearly, additional research should be directed to determine how insights from the human team literature can help improve human-automation team performance. The recognition of an automation agent's strengths and weaknesses should also be factored into the design of more effective collaborations.

Furthermore, given the human operators' tendency to interact with automation in ways similar to their human teammates, more emphasis is needed on calibrating the expectations and behaviors of human operators toward a more realistic means of collaborating with their automation teammates.

Acknowledgements

INL is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517. This work of authorship was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights. The United States Government retains, and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof. The INL issued document number for this paper is: INL/CON-14-31340.

References

[1] K. Christoffersen, and D. Woods, *"How to Make Automated Systems Team Players,"* In E. Salas (Ed.) *Advances in Human Performance and Cognitive Engineering Research*, Elsevier, 2002, New York, NY.
[2] S. Land, J. Malin, C. Thronesberry, and D. Schreckenghost, *"Making Intelligent Systems Team Players: A Guide to Developing Intelligent Monitoring Systems,"* (NASA Technical Memorandum 104807), 1995, National Aeronautics and Space Administration, Houston, TX.
[3] M. Steinberg, *"Moving from Supervisory Control of Autonomous Systems to Human-Machine Teaming,"* 4th Annual Human-Agent-Robot Teamwork Workshop, (2012).
[4] J. Lee, and K. See, *"Trust in Automation: Designing for Appropriate Reliance,"* Human Factors, 46(1), 50-80, (2004).
[5] E. Salas, N. Cooke, and R. Rosen, *"On Teams, Teamwork, and Team Performance: Discoveries and Developments,"* Human Factors: The Journal of the Human Factors and Ergonomics Society, 50(3), 540-547, (2008).
[6] A. Pritchett, *"Reviewing the Role of Cockpit Alerting Systems,"* Human Factors and Aerospace Safety, 1, 5–38, (2001).
[7] P. Lencioni, *"The Five Dysfunctions of a Team,"* Josey-Bass, 2002, San Francisco, CA.
[8] R. Helmreich and H. Foushee, *"Why Crew Resource Management? Empirical and Theoretical Bases of Human Factors Training in Aviation,"* In E. Wiener, B. Kanki & R. Helmreich (Eds.) *Cockpit Resource Management*. Academic Press, 1993, San Diego, CA.
[9] T. Dickinson, and R. McIntyre, *"A Conceptual Framework for Teamwork Measurement,"* In M. Brannick, E. Salas, & Prince, C. (Eds.) *Team Performance Assessment and Measurement*. Lawrence Erlbaum Associates, 1997, Mahwah, NJ.
[10] M. Letsky, N. Warner, S, Fiore, and C. Smith, *"Macrocognition in Teams: Theories and Methodologies,"* Ashgate Publishing Limited, 2008, Hampshire, UK.
[11] G. Klein, S. Wiggins, and C. Dominguez, *"Team Sensemaking,"* Theoretical Issues in Ergonomics Science, 11(4), 304-320, (2010).
[12] Y. Soraji, K. Furuta, T. Kanno, H. Aoyama, S Inoue, D. Karikawa, and M. Takahashi, *"Cognitive Model of Team Cooperation in En-Route Air Traffic Control,"* Cognition, Technology & Work, 14(2), 93-105, (2012).
[13] E. Salas and S. Fiore, *"Team Cognition: Understanding the Factors that Drive Process and Performance,"* American Psychological Association, 2004, Washington, DC.
[14] C. Bowers, E. Salas, E. and F. Jentsch, *"Creating High-Tech Teams: Practical Guidance on Work Performance and Technology,"* American Psychological Association, 2005, Washington, DC.

[15] Institute for Nuclear Power Operations. *"Guideline for Training and Qualification of Licensed Operators,"* (INPO ACAD 10-001), 2010, Atlanta, GA.
[16] Institute for Nuclear Power Operations. *"Control Room Supervision, Operational Decision-Making, and Teamwork,"* (INPO SOER 96-1), 1996, Atlanta, GA.
[17] J. Toquam, J. Macaulay, C. Westra, Y. Fujita, and S. Murphy, *"Assessment of Nuclear Power Plant Crew Performance Variability,"* In M. Brannick, E. Salas, & Prince, C. (Eds.) *Team Performance Assessment and Measurement*. Lawrence Erlbaum Associates, 1997, Mahwah, NJ.
[18] J. O'Hara and E. Roth. *"Operational Concepts, Teamwork, and Technology in Commercial Nuclear Power Stations,*" In C. Bowers, E. Salas, & F. Jentsch (Eds.) *Creating High-Tech Teams: Practical Guidance on Work Performance and Technology*. American Psychological Association, 2005, Washington, DC.
[19] Y. Chung, W. Yoon, and D. Min, *"A Model-Based Framework for the Analysis of Team Communication in Nuclear Power Plants,"* Reliability Engineering & System Safety, 94(6), 1030-1040, (2009).
[20] J. Park, W. Jung, and J. Yang, *"Investigating the Effect of Communication Characteristics on Crew Performance Under the Simulated Emergency Condition of Nuclear Power Plants,"* Reliability Engineering & System Safety, 101, 1-13, (2012).
[21] K. Sycara, and G. Suktghankar, *"Literature Review of Teamwork Models,"* (CMU-RI-TR-06-50), 2006, Carnegie Mellon University, Pittsburgh, PA.
[22] K. Sycara, and M. Lewis, *"Integrating Intelligent Agents into Human Teams,"* In E. Salas & S. Fiore (Eds.) *Team Cognition: Process and Performance at the Inter and Intra-Individual Level*. American Psychological Association, 2004, Washington, DC.
[23] J. O'Hara, and J. Higgins, *"Human-System Interfaces to Automatic Systems: Review Guidance and Technical Basis,"* (BNL Technical Report 91017-2010), 2010, Brookhaven National Laboratory, Upton, NY.
[24] L. Suchman, *"What is Human-Machine Interaction,"* In W. Zacahary, S. Parasuraman, & J. Black (Eds) *Cognition, Computing, and Cooperation*. Ablex Publishers, 1990, New York: NY.
[25] D. Woods, J. Tittle, M. Feil, and A. Roesler, *"Envisioning Human–Robot Coordination in Future Operations,"* IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, 34(2), 210-218, (2004).
[26] S. Fiore, F. Jentsch, I. Becerra-Fernandez, E. Salas, and N. Finkelstein, *"Integrating Field Data with Laboratory Training Research to Improve the Understanding of Expert Human-Agent Teamwork,"* Proceedings of the 38th Hawaii International Conference on System Sciences, (2005).
[27] X. Fan, and J. Yen, *"Modeling Cognitive Loads for Evolving Shared Mental Models in Human–Agent Collaboration,"* IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, 41(2), 354-377, (2011).
[28] K. Feigh, and A. Pritchett, *"Requirements for Effective Function Allocation,"* Journal of Cognitive Engineering and Decision Making, 8(1), 23-32, (2014).
[29] A. Pritchett, S. Kim, and K. Feigh, *"Measuring Human-Automation Function Allocation,"* Journal of Cognitive Engineering and Decision Making, 8(1), 52-77, (2014).
[30] A. Pritchett, S. Kim, and K. Feigh, *"Modeling Human-Automation Function Allocation,"* Journal of Cognitive Engineering and Decision Making, 8(1), 33-51, (2014).
[31] E. Roth, and J. O'Hara, *"Integrating Digital and Conventional Human System Interface Technology: Lessons Learned from a Control Room Modernization Program,"* (NUREG/CR-6749), U.S. Nuclear Regulatory Commission, 2002, Washington, DC.
[32] M. Wright, and D. Kaber, *"Effects of Automation of Information-Processing Functions on Teamwork,"* Human Factors, 47(1), 50-66, (2005).
[33] C. Billings, *"Aviation Automation: The Search for a Human-Centered Approach,"* Lawrence Erlbaum Associates, Inc., 1997, Mahwah, NJ.
[34] R. Parasuraman, and V. Riley, *"Humans and Automation: Use, Misuse, Disuse, Abuse,"* Human Factors, 39(2), 230-253. (1997).
[35] D. Woods, and E. Hollnagel, *"Joint Cognitive Systems,"* CRC Press (Taylor & Francis), Boca Raton, FL.