



Exploring human-autonomy teaming methods in challenging environments: the case of fighter pilots and loyal wingmen

Rune Stensrud¹ · Bjørn Mikkelsen¹ · Sigmund Valaker¹

Received: 10 April 2023 / Accepted: 24 January 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

The changes in the security environment run parallel to changes in humans and artificial cognitive systems to meet these challenges. In this article, the purpose is to discuss some human-autonomy teaming (HAT) design approaches (mechanisms for coordination): levels of automation (LOA), mixed initiative (MI), and coactive design (COAD). Specifically, we discuss how humans and artificial cognitive systems, exemplified as loyal wingmen, can be orchestrated to enable the handling of complexity and dynamics of an environment, e.g., handling military threats, and how different design trade-offs are affecting mission solutions. We also discuss some consequences of various AI/ML modes used by LW, on the utility of the HAT design approaches. Ways of using these HAT designs in a complementary way are suggested to support coordination through both plan and feedback, such as by integrating external and internal feedback in prediction of future action. We illustrate our suggestions through a use case, which provide additional nuance to our theoretical discussion. Lastly, we provide directions for future research, in particular through experimental design and the use of simulation, and provide practical implications.

Keywords Levels of automation · Mixed-initiative design · Coactive design · Human Systems Integration · Systems engineering · Environmental characteristics · Coordination

1 Introduction

Handling an external environment is key to many types of human-machine teams: e.g., the use of unmanned surveillance. The changes in the security environment run parallel to changes in humans and artificial cognitive systems to meet these challenges. In a military setting, changes in the environment are exemplified by novel technologies of adversaries such as in high-speed and high-precision missiles that can be deployed to ensure anti-access area denial (A2AD) capabilities, i.e., ability to control access to and within an operating environment. While novel humans and artificial cognitive systems may be important to handle such situations, it is important to enable the use of the technologies so that they will actually have the effect of reducing threats. In this paper, we discuss some human-autonomy teaming (HAT) design approaches (mechanisms for coordination) to handle a changing environment. Specifically, we

discuss levels of automation (LOA), mixed initiative (MI), and coactive design (COAD). We discuss how humans and artificial cognitive systems can be orchestrated to enable the handling of complexity and dynamics of an environment, e.g., handling military threats, and how different designs are affecting mission solutions. Specifically, we suggest that there are trade-offs between the HAT designs so that LOA and, to some degree, MI provide better coordination in low complexity and low dynamics environment, while COAD could support coordination in high complexity and high dynamics. LOA and MI could be relatively less costly in low complexity and dynamics while the opposite holds for COAD.

Ways of using these HAT designs in a complementary way are suggested to support coordination through both prescribed route planning and feedback, such as by integrating external and internal feedback in prediction of future action. We illustrate our suggestions through a conceptual use case of a fighter jet and loyal wingman drone collaboration, which provide additional nuance to our theoretical discussion. We consider how different artificial intelligence (AI) and machine learning (ML) modes of the LW may influence the utility of different coordination mechanisms.

✉ Rune Stensrud
Rune.stensrud@ffi.no

¹ The Norwegian Defence Research Establishment,
Instituttveien 20, Postbox 25, N-2027 Kjeller, Norway

Lastly, we provide directions for future research and practical implications for drones in difficult rescue operations and the widespread use of such drones in a military setting to ensure endurance and reach.

In the use of loyal wingmen, there are often requirements for “tightly coupled interaction between humans and machines” (Lyons et al. 2021, p. 2). This underscores the need for human-autonomy teams (HATs) defined as “at least one human working cooperatively with at least one autonomous agent (McNeese et al. 2018), where an autonomous agent is a computer entity with a partial or high degree of self-governance with respect to decision-making, adaptation, and communication (Demir et al. 2018; Mercado et al. 2016; Myers et al. 2019)” (O’Neill et al. 2022, p. 904). We further define autonomous agents as any “computer-based entity that is individually recognized as occupying a distinct team member role” (O’Neill et al. 2022, p. 907). They are “capable of making decisions independent of human control (Vagia et al. 2016)” (Lyons et al. 2021, p. 3), and, as suggested by O’Neill et al. (2022), autonomy is a matter of degree. Related, but not similar, notions to HAT such as human-automation-interaction (HAI) have also been coined to cover parts of this collaboration (Kaber 2018a; O’Neill et al. 2022). A problem facing HAT is how to be able to adjust own actions, both human and machine, taking into account environmental characteristics, in such a way that the HAT is continued and able to perform the joint tasks despite various environmental conditions. Prior research has suggested that when complexity of the environment is low, more decisions may be delegated to the computer entity, but when the complexity increases human control is needed (Abbink et al. 2018). Furthermore, prior research indicates that task resolution, of both human-to-human and human-machine teams, is impaired, in general, when complexity increases (O’Neill et al. 2022).

However, the degree to which more complex situations can be handled by machine and humans may depend on the different designs of HATs ranging from planned levels of automation (LOA) and adjusted automation (AA) via mixed initiative (MI) to enacted coactive designs (COAD) (Bekier 2013; Jiang and Arkin 2015; Kaber 2018a). We define and discuss below these concepts and theories. Complicating the way one designs collaboration is that both machines and humans vary as to their capabilities both between entities and over time in an entity. Some research is, for example, more optimistic concerning the ability of technology to handle unforeseen situations (Goodrich et al. 2022; Lundberg et al. 2021). However, with the aid of an Adjusted Automation Acceptance Model, the type of user, in this case, a professional user, can impact on automation acceptance (Bekier 2013, p.181). A concrete example is the development of drones that over time have developed to more adaptable entities. In our case, in the development of LW, the use

of a priori and reactive AI/ML modes is an example of the versatility of the autonomous entities. Kaber (2018a) also points out that humans vary according to how they handle workload and the way they make decisions, sometimes trying to satisfy rather than optimizing.

We suggest that environmental characteristics, specifically complexity and dynamics (defined below) and their influence on coordination and performance (Mouloua et al. 2020), will vary according to the HAT design used and the capabilities and behaviors of both human and machine. Prior research have pointed out that HAT models are in need of further description and of taking into account the actual variability of human and machine capability (Kaber 2018a; Endsley 2023). Lundberg and Johansson (2021, p. 382) also pointed out that there are “pitfalls associated with dividing control between humans and machines according to unidimensional and categorical models.” O’Neill et al. (2022) in a similar vein developed further a LOA model by extending the unidimensional levels of autonomy to the second dimension of levels of autonomy, thus problematizing and extending a much used model of HAT.

However, much of this research, to our knowledge, is still focused on developing further the LOA model and, to a less extent, discussing in depth other models. We broaden this outlook and suggest that MI and COAD are particularly relevant as ways of conceptualizing HAT taking into account more dynamic interactions between man and machine than the LOA perspective. By comparing and contrasting these approaches to HAT, we aim to contribute to the need for developing more unified theories of HAT as suggested by O’Neill et al. (2022, p. 927). Specifically, we discuss LOA, MI, and COAD, in the face of environmental complexity and dynamism, and suggest that the three HAT designs have strengths and weaknesses, and to overcome such trade-offs, one may use different HAT in concert.

For example, delegating tasks using a LOA framework may reduce communication costs, while COAD may increase synergies among human and computer entities. Thus, interfaces (Rico et al. 2018) need to be refined in parallel to advancements in the capabilities of autonomy, especially with the vision of a human operator and intelligent agent dynamically collaborating to solve problems and share task completion in a manner similar to effective human teams, e.g., supporting peer-to-peer type communication between human-autonomy team members (Schraagen et al. 2022). We ground these suggestions on dynamic decision theory (DDM) and organization theory, which points to the complementarity of feedforward (planned coordination), and feedback (coordination through adjustments) (Brehmer 1992; Brehmer 2010; Chanel et al. 2020; Simon 1957) in order to handle different problems, e.g., problems of scheduling and problems of unforeseen events. We argue that it is important to use these fundamental theoretical works in

organizational research as a basis for the use of more specific theories, such as phase theories of team collaboration as suggested by O'Neill et al. (2022). The broad notions of feedforward and feedback underlie in many ways in effective teamwork as indicated by Lyons et al. (2021, p. 5) and would also apply to HAT through their suggested aspects of HAT: signaling intent, promoting shared cognition, and monitoring performance (Frazier 2022; Johnson et al. 2011; Mathiasen and Nummedal 2022; Minos-Stensrud et al. 2021; Nummedal 2021; Sheridan and Ferrell 1974).

In less complex situations, LOA may be preferred because less feedback is required. A dominant feedforward (command-based) coordination mechanism based on a high degree of pre-programmed and role-based assignment of authority. On the other hand, it requires that formal roles and capabilities (of human and machine) must be defined prior to task resolution, something that could be increasingly difficult to foresee when complexity and dynamics increase. MI and, to an even larger extent, COAD may be (at least initially) too reactive to fully function in time-critical tasks.

On the other hand, in highly complex situations, more elements need to be observed, processed, and potentially communicated about in order to coordinate between man and machine. There may be a likelihood that the tasks cannot be as easily divided in high complexity and high dynamic situations as in low complex and low dynamic situations. We may posit that increased interdependence among coordinating elements in an extended socio-technical multiteam (Luciano et al. 2020) that includes non-human actors leads to increased coordination requirements on the boundary of this arrangement. As the degree of interdependence between entities increases, the need to process more information increases. In such situations, HAT designs that rely on feedback should function better, and MI and, perhaps, even more COAD could be chosen rather than LOA. A dominant feedback (e.g., threat-based model) coordination mechanism based on a high degree of reactive feedback and operator-based assignment of authority will promote this type of arrangement (informal control).

In other words, there may be trade-offs between the HAT designs. To compensate for the trade-offs using the different HAT designs in concert could be a solution. An integrated information model of fully integrating feedforward and feedback with a goal generator (e.g., an adaptive route planner) will promote this evolution (formal and informal) adjusting who coordinates based on both (pre-programmed) role and (individually and situated selection of) operator most suitable to mitigate any gap of authority (Johnson et al. 2018). For example, MI (and/or COAD) could be used together with dynamic LOA levels (Petousakis et al. 2021; Schmitt et al. 2018; Lindner et al. 2022). Comment on Miller and Parasuraman (2007) could be used to ensure predictability in

a COAD design. In this way, planned and emergent activities could be joined.

On this background, the purpose of this article is to elucidate the following research question: How do different HAT designs contribute to support the coordination of task under various environmental characteristics? In this article, we thus discuss how HAT designs, specifically levels of automation, mixed initiative, and coactive design, may support changes to the workflow between man and machine in a military mission (Fitts and Jones 1947; Sheridan and Verplank 1978; Parasuraman et al. 2000), due to different environmental characteristics. This article thus explores some parts of a research gap identified by O'Neill et al. (2022) who call for investigating the role of different task conditions for HAT designs. We do so by focusing on certain characteristics of the task environment. We also build on and extend the prior work that we have done on HAT designs. In a 2021 paper, we discussed the prospect of HAT teaming in the context of unmanned combat aircrafts collaborating with fighter jets. Stensrud et al. (2021) indicated that the dynamic of tasks would influence the type of coordination between human and non-human entities requiring a mix of formal and informal mechanisms, but here, we add the influence of environmental complexity and look upon a less controllable empirical setting. We discuss a use case building on prior empirical and conceptual work that we and others (Hamstra et al. 2019; Frey et al. 2018) have done regarding F-35 and loyal wingman, employing different AI/ML modes, an unmanned drone (e.g., Stensrud et al. 2020; Stensrud et al. 2021; Stensrud and Valaker 2022). Finally, we discuss future research and practical implications. In particular, we consider experimental designs as well as potential for simulation to support the analysis of the coordination schemes.

2 Use case

We are illustrating the three different HAT designs with the collaboration between fighter aircraft (e.g., F-35) that collaborate with loyal wingmen (e.g., surveillance drones) to solve missions (Rebensky et al. 2022). There is a need for an interaction mechanisms, i.e., HAT design, in order to ensure coordination. These tactical man-machine systems are embedded in a larger organization (multinational military organization) which directs the activities of the fighter aircraft and loyal wingmen, illustrated in Figs. 1 and 2. Figure 1 illustrates a prescribed socio human capability that includes non-human actors and limitations that are considered, and Fig. 2 illustrates an emergent and dynamic handover take-over event with a changed socio human capability that expands the ability to include non-human actors. The lower part of Figs. 1 and 2 illustrates the fighter aircraft

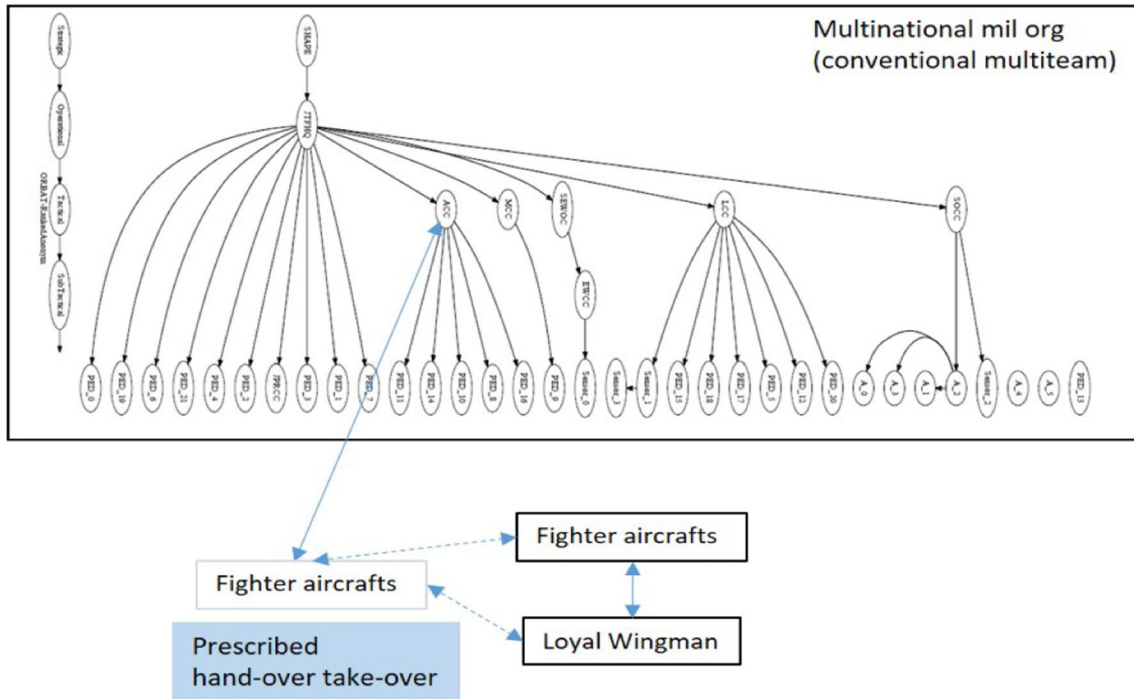


Fig. 1 Prescribed socio-technical multiteam that includes non-human actors (adapted from Stensrud et al. 2020)

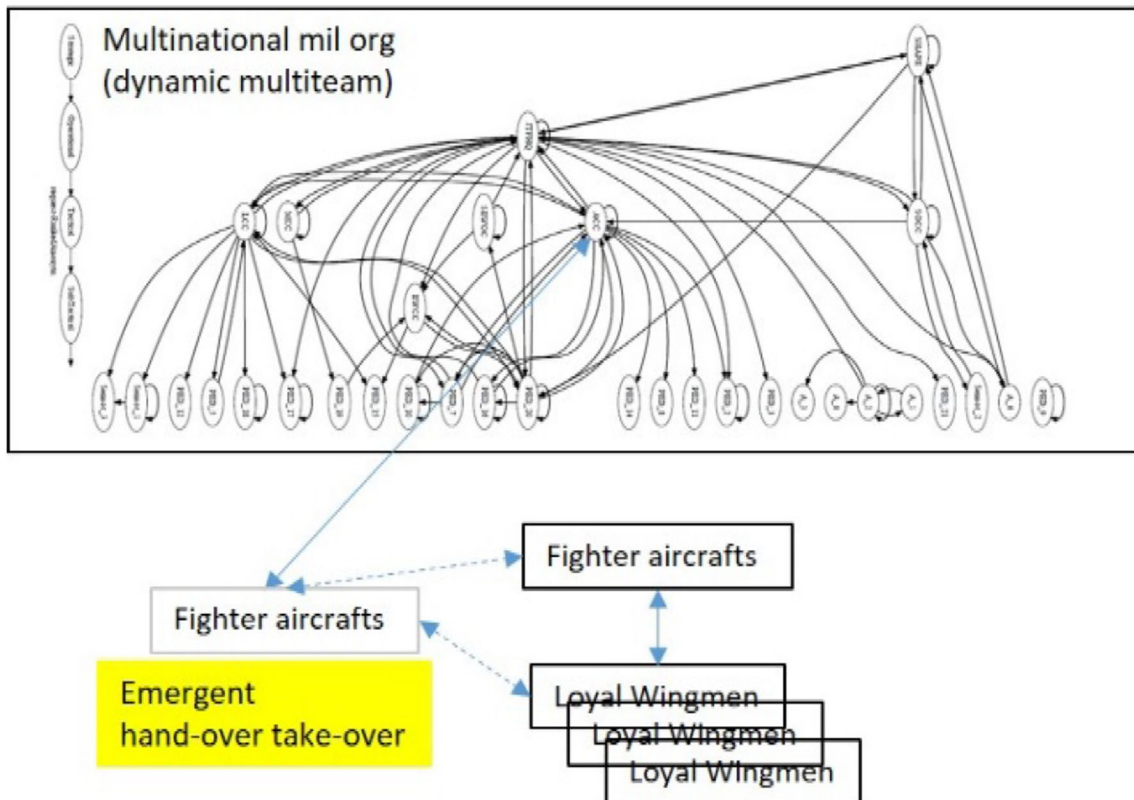


Fig. 2 Emergent socio-technical multiteam that includes non-human actors (adapted from Stensrud et al. 2020)

and loyal wingman system and its relation to the larger organization.

We draw on prior unclassified reports by RAND to form a use case:

“ (...) low-observable, multirole F-35 Lightning II Joint Strike fighter could be used as both a sensor and a shooter in a SEAD campaign.⁷ Still, air planners will likely want to reduce the amount of time that F-35 aircraft spend in highly contested airspace by leveraging space-based ISR to help locate adversary SAMs [surface-to-air-missiles].⁸ Using long-range precision ground fires would also increase the firepower available to strike targets, offer a redundant capability to strike SAMs if aircraft need to leave the area, and complicate an enemy’s defense planning (...)” (Priebe and Douglas 2020: p. 37). We extend this case to include loyal wingmen (Stensrud et al. 2020) that augment and are complementary to the F-35 (Frey et al. 2018). The loyal wingmen are drones that could but used in a forward position both to surveil and attack an enemy air defense system.

3 Effects of different system design approaches on coordination under different environmental characteristics

We now discuss how LOA, MI, and COAD may support coordination in the two types of environmental characteristics: complexity and dynamism. We define coordination as the integration of interests, understanding, and activities to reach a common goal (Mathieu et al. 2018; Van de Ven et al. 1976; Kouchaki et al. 2012; Grote et al. 2018). We define complexity as the number of elements and number of relations among elements in an environment (Schneider et al. 2017) and dynamics as the rate of change in elements in the environment (Dess and Beard 1984). We specify these dimensions in the case to the following: the number of elements in the enemy-integrated air defense (e.g., number of radars, missile launchers, and C2 nodes), which is the complexity of the environment, and the variability of an additional air threat (e.g., number of incoming enemy fighter aircraft) which is the dynamic characteristic. We assume that in all conditions, a set number of elements in the SAMs should be stroked (even though the total number of SAMs increases), and we keep the number of friendly airframes constant. The low complexity and low dynamic conditions are illustrated in the planned mode in Fig. 1, while the high complexity and high dynamics are illustrated in the emergent mode in Fig. 2. We now summarize key definitions of the three HAT designs and illustrate how they may be utilized

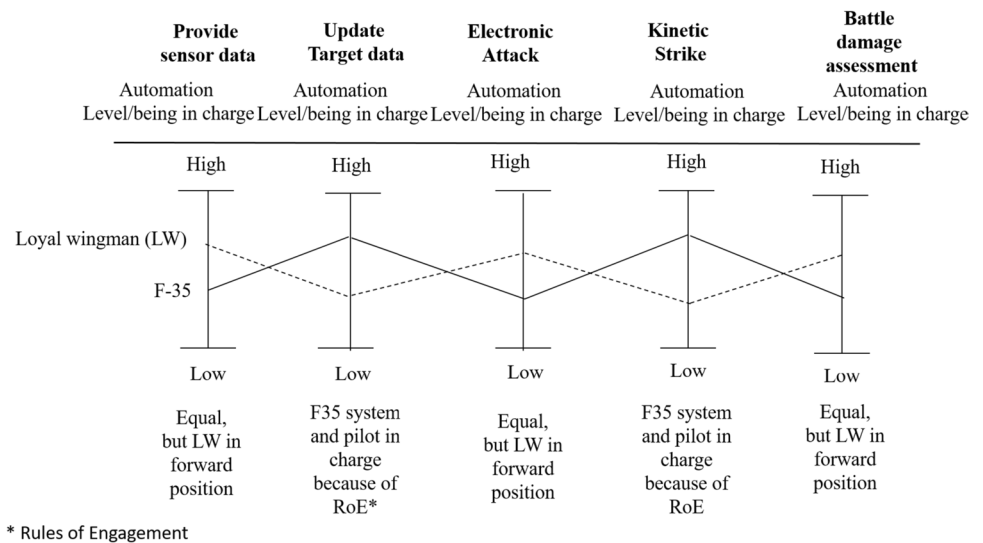
in the case of F-35 and loyal wingman collaboration in a SEAD mission that we described above. Before doing so, we introduce the principal AI/ML modes employed by the developmental LW of our case and discuss how these modes may handle various environmental contingencies.

The LW systems can be differentiated according to not only the “payload” that they carry but also the AI/machine learning architecture that they use. The LW system in its current form is using a reinforced learning approach for route planning with respect to use the LW for sensor tasks. This consists of algorithms for an efficient detection process which allow for optimizing the search patterns of the LW area of operation. Two different ways are used to accomplish “optimization.” Random and “perfect” search are two modes that the LW system uses. Random search is dependent on the “a priori” search pattern, while the “perfect” search is based on uniform search, which basically means searching the whole area. The human operator can actively decide what pattern is used by the LW. Similar algorithms may be thought for other uses of the LW such as electronic attack and thus represent a random and perfect mode.

The next step in development is the combination of the two modes. The two modes utilize different starting values, and if both types of variables are available, then the human operator may orchestrate the different modes.

Another way of conceptualizing the different AI/machine learning schemes employed by the potential LW is the deliberate and reactive architecture and the combination of these two modes. In general, one may foresee that more elaborate deliberate modes of AI/machine learning may lead the human operator to delegate more tasks and decisions to the LW. On the other hand, the reactive mode may require more specific instructions as the LW tries to solve tasks. On a general note, if the environmental complexity and dynamics is low, one may prefer to use the deliberate AI/machine learning architecture because it is easier to foresee the inputs and rules for efficient use of the LW since the environment is less complex and dynamic. In these circumstances, there may be less need for time-critical man–machine teaming, but rather some extensive involvement from humans in planning and assessing the LW tasks. Conversely, when the environmental complexity and dynamics is high, one may use the reactive AI/machine learning mode because one wants to inform the LW decisions in a more updated way about information not foreseen initially in the task resolution. Such situations may lead to a higher degree of teaming between the human operator and the LWs as the LW tasks are carried out. In tasks which contain a mix of the two, a parallel use of both modes may be very much desired, e.g., if one area that must be surveilled is stable while another is highly dynamic, or if there are areas of low complexity and areas of high complexity simultaneously.

Fig. 3 Levels of automation and human activity along a timeline divided in subtasks of a SEAD mission



We now go through the relation between LOA, MI, and COAD and how they may be dynamically used as modes of collaboration between F-35 and loyal wingmen. Moreover, as generally discussed above, one may foresee different implications of the AI/machine learning modes for autonomy and teaming, and we will remark on such implications to the end of our discussion.

LOA concern prescribed “levels” where the human does “everything to a level where the computer does everything” (Sheridan and Verplank 1978, p. 8–5), and the research have been concerned with how LOA changes the working conditions for human operators (Fitts and Jones 1947; Sheridan and Verplank 1978; Parasuraman et al. 2000). LOA is focused on the task itself and what subtasks man and machine respectively should carry out. Several extensions and modifications to the original framework have been made (Vagia et al. 2016); a recent example is Cabrall et al. (2018), and it has been used in several practical applications (Hopkins and Schwanen 2018). One key observation is that automation may hamper visibility and as a consequence situation awareness of humans if used extensively. Miller and Parasuraman (2007), for example, suggest that LOA could be changed if the environment demanded and that automating decision-making functions “may reduce human operators’ awareness of system and environmental dynamics” (Miller and Parasuraman 2007). On the other hand, extensive automation and delegation may help perform tasks in difficult environments and when speed is essential (Kaber 2018a, b). The research to date indicates that flexible use of LOA could be used to handle some environmental demands.

With respect to LOA, and to systematically visualize the different ways, levels of automation could be foreseen in the collaboration between F-35 and loyal wingman; we use the profiling of automation levels by Parasuraman et al. (2000) which is cited in Miller and Parasuraman (2007). In

the profiling of levels of automation for the loyal wingman, we add the being in charge of which is the human actors’ discretion to do specific tasks. We focus on five critical tasks: provide sensor data, update target data, perform electronic attack, and kinetically strike parts of SAM sites as well as initial battle damage assessment (BDA).

Adjustment of levels of automation and human activity may be illustrated along a timeline of subtasks of a SEAD mission. Figure 3 illustrates a proposal for a selection of levels of automation for loyal wingman and for what the F-35 fighter pilot/aircraft can do with respect to critical tasks in SEAD (cf. Priebe and Douglas 2020). With respect to providing sensor data, we foresee that it is possible that the loyal wingman takes an active part and provides sensor data on par or better than the F-35. With regard to updating target data (i.e., SAM sites position, etc.), the loyal wingman has less discretion as we suggest that the F-35 system must be involved for the quality check of the data, as well as we want the F-35 pilot to be in charge so that rules of engagement (RoE) is taken care of, something that we do not want to delegate to the loyal wingman. Then, moving on to electronic attack (jamming), the F-35-system and the loyal wingman equal LOA but LW in a forward position. Kinetic strike F-35 system and pilot will be in charge because of RoE. However, from the point of view of the drone system kinetic attack, e.g., acting as a platform to deliver a weapon, this may, however, be relatively easier to achieve than more advanced types of electronic attack.

Although specific tasks could be delegated beforehand to the LW, the actual technological capability and capacity of the LW may influence to what degree the F-35 pilot actually delegates the tasks.

Another constraint is the degree to which the LW is able to communicate with the F-35 and its specific datalink requirements. Another approach that could be chosen is to

relax the requirement for the LW to solve whole subtasks, or substantial part of the subtasks. Rather one may choose the mixed-initiative approach in which the LW contribution is more situation dependent. The LW may, for example, provide sensor data that the F-35 system may or may not choose to incorporate into its task resolution. In the MI collaboration, there may be high requirements for protocols of dialogue, but the F-35 reliance on the LW completing subtasks may be less. The LW may be an option that could help the F-35 pilot in their tasks. This may also “solve” an additional issue that is salient in pilot tasks, that of limited time and span of attention by the human pilot. Somewhat paradoxically in the LOA framework, one may be more dependent on the LW than in the MI framework. In the LOA framework, there may be more attention needed to track the LW contribution to the overall SEAD mission.

The MI approach distinguishes itself from LOA by positioning a more equal role for the machine and has been described as “a flexible interaction strategy in which each agent (human or computer) contributes what it is best suited at the most appropriate time” (Allen et al. 1999, p. 14). Gombolay et al. (2017) investigated key gaps in prior literature by assessing how situational awareness is affected by the level of autonomy in mixed-initiative scheduling for human–robot teams, the effects of increased or decreased workload in human–robot team fluency, and the role of workflow preferences in robotic scheduling.

With respect to MI, we make some examples of what the F-35 can contribute to in a situation, and what the LW can contribute to in Table 2. For example, in a situation, the LW is better placed to do a specific part of the task (e.g., electronic attack), and at the same time, the F-35 is best at kinetic, while in another situation, it is more suited to do general sensor coverage because it can be closer to enemy forces, versus F-35 will do the targeting process due to its better sensor-fusion and due to legal requirements that a human in the loop should do final clearance, etc. Based on Allen et al. (1999), we propose that joint activity is about interaction and negotiation, i.e., mixed initiative within a dialogue framework between the F-35 system and loyal wingman (Table 1), and becomes more adaptable to the situation than the prescribed LOA framework.

A mixed-initiative system offers applications (e.g., an interactive planning application) where both the user (e.g., a pilot) and agent (e.g., a (pre-configured) loyal wingman) are notified if the situation or plan changes. At this basic level, the mixed-initiative system does not then coordinate the subsequent interaction on this first step allowing *unsolicited reporting*. The next level (table) involves *subdialogue initiation*, and the mixed-initiative system asks for authorizations (pre-clarification and post-assessments) that might take several *fixed subtask* interactions between F-35 system (pilot) and loyal wingman (agent). Mixed-initiative system

Table 1 Mixed initiative within a dialogue framework (Allen et al. 1999) (p.15)

| Mixed-initiative levels | Capabilities | F-35 system | Loyal wingman |
|---|--|---|--|
| Unsolicited reporting on sensing | Agent may notify others of critical information as it arises | Are notified if the situation or plan changes | Are notified if the situation or plan changes |
| Subdialogue initiation on targeting | Agent may initiate subdialogues to clarify, correct, and so on | Coordinate the subsequent interaction by <i>subdialogue initiation</i> to clarify status | Answer the subsequent interaction by <i>subdialogue initiation</i> feedback on clarifying status |
| Fixed subtask initiative on kinetic attack (optionally electric attack) | Agent takes initiative to solve predefined subtasks | Initiates clearance or request that might take several <i>fixed subtask</i> interactions | Initiates or requests for authorizations (pre-clarification) that might take several <i>fixed subtask</i> interactions |
| Negotiated mixed initiative on battle damage assessment (BDA) | Agents coordinate and negotiate with other agents to determine | At the final <i>negotiated mixed-initiative</i> level, there are no fixed assignments of responsibility or initiative | At the final <i>negotiated mixed-initiative</i> level (e.g., post-assessments), there are no fixed assignments of responsibility or initiative |

Table 2 Evaluation of coordination through HAT designs under different environmental characteristics

| Environmental characteristics | Levels of automation | Mixed initiative | Coactive design |
|-------------------------------|---|---|---|
| Low complexity | High coordination is possible because it is easy to define what can be done by man and machine respectively. An a priori AI/ML would be useful | High coordination, but may be an overly complex method of teaming since complementarity may be less needed | High coordination, but may incur a cost as too much emphasis is put on feedback, when it is not really needed to have feedback |
| High complexity | Difficult to coordinate because the number of elements and their interrelations are many, and therefore, it is hard to decompose and delegate tasks | Good coordination, but relatively few mechanisms for feedback and adjustments may hamper operations | Good coordination because feedback is processed concerning the interrelations. However, dependent on the capability to observe, predict, and handle concurrent tasks. A reactive AI/ML would be useful here |
| Low dynamics | Good coordination because it is easy to predict the upcoming tasks and hence to predict to whom to delegate to. An a priori AI/ML would be useful | Good coordination and easy to predict the complementarity (e.g., sequencing of human and machine actions is easy) | Good coordination, but may incur a cost as to how much feedback is needed, because the environment is predictable and hence feedback is not needed frequently |
| High dynamics | Difficult to coordinate because of the many changes in the environment that makes it hard to predict what to delegate | Good coordination, but few mechanisms to adjust for sudden changes in who does what | Good coordination because ongoing feedback can aid prediction even in high dynamics. A reactive AI/ML would be useful here |

is responsible for choosing routes, communication (emission control), and refueling for each action along the pre-planned timeline of subtasks of a SEAD mission. At the final *negotiated mixed-initiative* level, there are no fixed assignments of responsibility or initiative. According to Allen et al. (1999), the mixed-initiative system supporting the F-35 system (a pilot-squad-lead or more operators) and loyal wingman will help to monitor the current task solution evaluating whether it should take the initiative in the interaction with the F-35 system and loyal wingman, basing this decision on many factors (demands and workload of the agent(s) (loyal wingmen) and risk of the F-35 system (pilot)) to interfere. An assumption in the MI design is that the one entity best able to carry out a subtask should take the initiative to do so, agnostic of whether it is man and machine. In this way, it, in a sense, incorporates in its mechanism some idea of whether an entity is fit to do the particular subtask, thus requiring at least a basic understanding of the current status of an entity in relation to the task and the environment. Later on, Jiang and Arkin (2015) have developed this definition to encompass robots. Jiang and Arkin (2015) suggest that feedback from an external environment or inferred state of environment can trigger initiative in a reactive or deliberate way; however uncertainties of the environment can make initiative reasoning challenging (Kirlik et al. 1993).

Moving to COAD, one of the key components of COAD is to emphasize collaboration rather than delegation as principles of man-machine interaction. COAD emphasizes the observation of and sharing of knowledge of own status and knowledge of internal interdependencies and external environment, as well as the predictability of own actions in order for others to rely on them and direct behavior and be directed (Johnson et al. 2014). Johnson and Bradshaw (2021) explain the problem of people becoming unaware of changes in the environment and system states when changes are under control by other agents, by the idea that the system is opaque to the users. They claim that support for interdependence could mitigate such issues. In general, the COAD extend the notion of human-machine teaming to a more systemic view than what LOA and MI do, as it emphasizes the information sharing (on OPD) between entities, something that is more restricted in the LOA and MI. COAD entails three essential interdependence relations—observability, predictability, and directability (OPD) (Johnson et al. 2014)—by defining roles and individual requirements (Johnson et al. 2011). Johnson et al. (2014) claim that it is an inverted relationship between automation and interaction, and COAD is emphasizing the interaction. One of the many advantages of COAD is the focus on core interdependence relations that can provide a formative tool for designers called Interdependency Analysis (IA). Interdependency Analysis (IA) supports what could be automated, and as a fundamental principle of COAD, interdependence must shape automation.

With respect to COAD and Continuum of Task Control (Frame et al. 2020) in human-autonomy teaming, the requirements for a future loyal wingman coactively collaborating with F-35 pilot need to be further investigated, e.g., concerning observability and controllability. This is partly discussed in Frame et al. (2020) about the Continuum of Task Control in human-autonomy teaming with different levels of adaptive automation, and concerning an applied surveillance, due to workload and human factors. The case example mentioned in the article of Frame et al. (2020) is the Traffic Collision Avoidance System (TCAS). Though a silently operating loyal wingman may succeed with its task, e.g., operating solely alone providing sensor data, updating target data, whether choosing soft or hard kill solutions, still, automation level and/or being in charge are more than a technical question due to law of armed conflict (LOAC) and rules of engagement (ROE) (i.e., despite of a sophisticated and costly LW solution). Frame et al. (2020) are exploring Task Allocation in Operational Surveillance and adaptive automation and problematize that supervisory control may consist of restriction for the human in the loop, i.e., introducing dilemmas with a highly automated system with the most direct control over a few, but specific task operations. However, it will possibly be dependent upon updates from its autonomous mates when in danger.

A loyal wingman which is able to autonomously engage with its environment in direct interaction, involvement, and/or interdependency with the pilot/squad lead (mission commander) and other artificial and autonomous entities in order to meet a certain objective may cooperate most effectively when continuously under task control (Sycara et al. 2020) and updating its world status. This is exemplified by archetypes in Stensrud et al. (2020) and in the human-autonomy teamwork of Sycara et al. (2020) supporting coordinated team activity. Tasks delegated to the LW under strong emission control (radio silence) may be needed. Though a radio-silent, highly dedicated LW will be less costly for the human in the loop but require a pre-programming effort up-front. However, a smart HAT approach is needed when a LW are to cooperate with other entities, besides deciding and acting on an individual basis (according to LoA), and we should believe that both the pilot and LW complement each other's decision-making process and actions perhaps best supported by a MI approach. However, in order to do so, a LW must be able to "understand" and sense the environment that is to "understand" complex rules of engagement (relative to the activity), i.e., to be both controllable and observable, adapt effectively to the changes in the environment, and to combine tasks (i.e., to be adapted by COAD approach of HAT).

A concrete example of observability of F-35 and LW may involve a third party. The interdependencies to a third party (like a mission commander/battle captain) monitoring and following a task resolution, e.g., a squad of F-35s that tries to do the suppression enemy air defense (SEAD) mission

as usual, are "followed" by a LW that learns the (potential) interdependencies and that tries to predict what to do, over time directability, and learns how to make contributions to the SEAD mission resolution, not necessarily as a "forward" observer or -shooter, but also in rear positions, based on the observed needs for the F-35. This could be a defensive unit protecting a high-value unit from inadvertent attack by a group of agents using defending robots (Grover et al. 2022). Solutions are not easy to foresee or derive from a top-down decomposition of the SEAD task.

One way to move forward with the COAD as a strategy for collaboration is to use simulation where one may observe the interactions between the F-35 and LW in an easy and controlled way. This may give ground for trying out specific protocols of collaboration live in later stages of development.

The different AI/machine learning architectures by the LW system could imply more refined implications for autonomy level and teaming. In situations of low complexity and low dynamics, we foresee the deliberate "a priori" architecture to be chosen, and for this type of AI/machine learning to be efficient, we assume that delegating decisions extensively will lead to better performance of the LW. However if one employs less delegation of tasks, this may lead to less optimal use of the LW in these circumstances. The opposite may hold in highly complex and dynamic situations where employing deliberate architectures may lead to less performance by the LW. Rather using the reactive mode and being able to utilize mixed initiative or COAD type of coordination mechanism may lead to better performance by the LW system. Being able to select the appropriate AI/machine learning schemes according to cues of the level of environmental complexity and dynamics may thus be central. An important feature for the human operator, as well as the LW system itself, is to be able to sense and foresee changes in such levels and thus proactively select an appropriate AI/machine learning architecture as well as an appropriate approach of collaboration: LOA, MI, or COAD.

We evaluate coordination, as in our discussion above, but also add considerations of costs of coordination in the different HAT designs. The results of our analysis based on experience from workshops with military officers are presented in Table 2. We do not present results for the low dynamics as we focused on the critical issue of handling high dynamics. Largely, our analysis of the empirical use case conformed to the theoretical discussions, although it provided some granularity to the theory. Overall, similar to the discussion of the case, it was highlighted that LOA were preferred in less complex and less dynamic situations, and MI and COAD were preferred when complexity and dynamics increased. With respect to LOA, it highlighted the relatively low to medium cost of planning (i.e., deciding beforehand who does what and delegate) even in low complexity. With respect to MI, the cost of making interfaces

between man and machine to accommodate transactions was highlighted. In high dynamics situations, it also suggested that the humans in the MI mode need to prioritize tasks on a high level rather than do detailed interaction with technology. Prioritizing tasks would require advanced technology. Regarding COAD, the case illustrated the cost of using this method for low complexity and low dynamics because of the lack of delegated and planned (initially) task allocation. It should be noted that these suggestions are preliminary and need further refinement through subject-matter expert input as well as experimentation (Table 3).

4 Discussion

Our preliminary theoretical analysis and analysis of a use case indicate that LOA, MI, and COAD could be complementary in terms of supporting coordination under high and low complexity and high and low dynamics. This largely confirms prior research, but our study may open up the discussion of how the different HAT designs could be used in concert to support difficult missions. It also described the potential role of different AI/ML modes for handling environmental contingencies and selection of coordination mechanisms. We now point to some avenues for future research and practical implications.

4.1 Future research

A key extension to this work is to define more clearly what the different HAT designs could offer in the particular case discussed and in other relevant cases. As pointed out by Kaber (2018a), there is a need to generalize beyond particular contexts of collaboration between man and machine. Defining the more general requirements for the collaboration and the capabilities of the autonomous agent may thus be crucial. In this respect, the dimensions of autonomy, perception of the agent as humanlike, and interdependence (Lyons et al. 2021) could be explored as dimensions that characterize LOA, MI, and COAD and influence their utility in different environmental conditions influenced by different LW with different configurations of AI/ML modes (e.g., a priori and/or reactive modes). Such an analysis could, for instance, detail the communication requirements, the rules for delegation, and the mechanisms for updating among entities offered by the designs.

Based on our findings, we can provide design guidance for roboticists developing intelligent collaborative robots that, for example, engage in mixed-initiative decision-making with human participants, e.g., a blueprint for further experimentation. More work is needed on defining how the trade-offs between HAT designs could be overcome, e.g., through developing COAD and its interface to LOA and MI. We conducted only a very general discussion of the costs of

Table 3 Preliminary evaluation of coordination and communication cost with different HAT designs under different environmental characteristics

| Environmental characteristics | F-35 with loyal wingman | | |
|-------------------------------|--|---|---|
| | Levels of automation | Mixed-initiative | Coactive design |
| Low complexity | Coordination: Low Cost: Medium. Due to some planning needed to define task allocation and delegation authority throughout the mission in detail | Coordination: High Cost: High because of the requirement to interface during task resolution | Coordination: Medium. Require a supervisor role for the human Cost: High. Need to mitigate the lack of planning through discovering interdependencies |
| High complexity | Coordination: High Cost: High. Need to define in much detail what to delegate | Coordination: High. But more likely to result in transaction overload Cost: High. Need to define “smart” interaction mechanisms to reduce transaction overload | Coordination: Medium. But better than in low complexity since the system is able to process the high number of elements Cost: Medium. Scalability is challenging |
| High dynamics | Coordination: Medium Cost: High. Very detailed updating and definition of task allocation mechanisms needed | Coordination: High Cost: High. Requires detailed transaction mechanisms and well-known procedures that allow the human operator to prioritize between tasks (in particular, in low complexity and high dynamics) | Coordination: Medium Cost: Medium. Needed to have a repertoire of playbooks and select among these (in particular, when high complexity is also present (Perkins 2017)). Support team collaboration by modern technology (AI) (Seeber et al. 2020) |

Table 4 Experimental design examining the effect of using LOA, MI, or COAD in the various AI/machine learning architectures and levels of environmental complexity

| | Deliberate AI/machine learning | Reactive AI/machine learning |
|-----------------------------------|---|---|
| Low complexity Low dynamics | Measure appropriateness of LOA, MI, or COAD | Measure appropriateness of LOA, MI, or COAD |
| High complexity and low dynamics | Measure appropriateness of LOA, MI, or COAD | Measure appropriateness of LOA, MI, or COAD |
| High complexity and high dynamics | Measure appropriateness of LOA, MI, or COAD | Measure appropriateness of LOA, MI, or COAD |
| Low complexity and high dynamics | Measure appropriateness of LOA, MI, or COAD | Measure appropriateness of LOA, MI, or COAD |

different designs in different situations. It is likely that the cost will also change over time as the actors learn how to operate in the environment.

Using the model of cost in the aviation industry developed by de Pasquale and Savill (2022), one may point at direct operating costs that are under the control of the manufacturer such as flying cost and maintenance cost, in addition to those not under direct control of the manufacturer such as financial cost. We point to some of the flying costs that could be relevant and some costs related to the airframes (maintenance costs). Such costs could be very relevant to consider in combination with the costs of teaming between the human and the LW system. Future research could examine, for example, the attention required from the human pilot to monitor and provide tasks to the LW system. This may be considered a flying cost and cockpit crew cost in terms of requiring, for example, different crews that are needed if the demand is high on interacting with the LW system. Perhaps, one may need to use more traditional ways of operating drone systems such as remotely piloting the drones in addition to the fighter jet pilot providing some orders, if there is a high task load perceived by the fighter jet pilot. In addition to these costs, the development cost of the LW system, such as developing advanced AI/ML modes, is important to factor in. All these costs could be examined in more detail and should inform whether a LW system is feasible and yields the desired benefits relative to costs. What could be seen as a cost could however also be seen as a critical condition to implement the LW system, and a consideration of balancing costs and benefits is therefore crucial.

In the case of fighter aircraft and the loyal wingman drone collaboration, future research may examine the value of introducing the loyal wingman. One particular issue is to examine what are the task load for the fighter pilot of using different collaboration designs, and the value for overall mission resolution. One may compare the use of fighter aircraft only with the use of loyal wingman. Furthermore, one may compare the utility of fighter aircraft and loyal wingman with different capabilities and capacities. Finally, one may use the loyal wingman in either a distinct LOA framework or more interactive designs such as MI and COAD. Moderating these different experimental conditions could be the

variables mentioned as particularly important to tackle in future research by Lyons et al. (2021) such as (a) communication of intent (back and forth between the fighter pilot and the loyal wingman) and specifically related to the issue of feedback and (b) mutual monitoring and facilitation of joint attention. Further training and implementation of HAT designs in live missions may need to be carefully designed.

Several experimental research designs may be used to help test the assertions we make. For example, the comparison of the different collaboration designs in differing levels of environmental complexity and dynamics could be compared. Added to this, the type of AI/machine learning architecture used by the LW system should be considered. To simplify, one may systematically consider the appropriateness of LOA, MI, and COAD in a 2×4 experimental design. The various conditions are either the (1) deliberate or (2) reactive architecture for AI/machine learning in four different types of environmental conditions, namely (1) low complexity and low dynamics, (2) high complexity and low dynamics, (3) high complexity and high dynamics, and (4) low complexity and high dynamics. The design is shown principally in Table 4.

In addition to the AI/machine learning schemes, there are, as already mentioned, other important features of the LW system to consider. One particularly important feature to consider is the type (or mix) of different payloads and functionalities onboard the LW such as sensor and effectors (kinetic and non-kinetic). Other important things to consider are the range, speed, and endurance of the LWs. This naturally gives potentially many different variants of systems to examine. In our particular development projects, some of the most prominent aspects are the sensor capabilities and their endurance and range.

Simulation, an artificial or synthetic environment created to manage human experiences of reality (Salas et al. 2008), may be used to facilitate experiments. The specific systems noted here are particularly amenable to be included in computer-supported simulations because they already are computer systems. The simulation may thus be purely computer-supported. On the other hand, through the simulation network, one may also connect live loyal wingmen that operate in live environments. More specifically, one can use the constructive simulation, i.e., computer-generated LW, to examine the

role of various AI/machine learning architectures as well as the influence of different scales of LW networks as well as varying other features that are not yet implemented in a live LW. Such future features may be extreme endurance LWs and/or LWs that can hold multiple traditional functionalities (sensors) with novel functionality (e.g., effectors employing novel techniques such as laser weapons). It will be important in a simulation/testbed to be able to tune such features. Moreover, it will be important to leverage the synthetic virtual and constructive environment to systematically vary levels of environmental complexity and dynamics. The number of entities that are undetected can be varied to illustrate levels of complexity, and how often certain entities appear or disappear can be varied to illustrate dynamics. A challenge to be tackled in simulating is for the live LWs to react on such synthetic stimuli in their sensors and effector systems. Hence, some simulation capability may need to be incorporated into the live LW platforms as well, in order for the entities to be part of a simulation environment.

Crucially, other characteristics of environments need to be discussed in more detail such as the different interpretations that could arise regarding an environment. Related to this, the filtering and transmission of environmental cues could impact the operation of both human and machine and thus determine their interdependencies and coordination. On a general note, the interaction of different characteristics needs to be spelled out in more detail, e.g., what is the consequence of a situation with high complexity and low dynamics versus a situation with high complexity and high dynamics? And how may the HAT system change from an a priori to a reactive mode in order to handle such changes in the environment?

The technical interface that could support the different HAT designs under different environmental characteristics is also in need of discussion, and related to this is the trust among entities. In particular, how to achieve coordination, between F-35s and LW, in congested and contested environments is in need of more study.

4.2 Practical implications

A suggested practical approach to implement different HAT designs is to use a supportive framework for design and modeling interfaces to evaluate the consequences of different HAT designs in a practical context (Endsley 2023; Park et al. 2020). In the development of the collaboration between man and machine, it seems to be important to support different needs and that not one HAT design (that we know of) seems to support all needs. Thus, in designing and supporting such collaboration, one may need to orchestrate different designs, such as making elaborate ways of delegating to the machine, yet also be able to get feedback from the machine and its environment when needed to adjust and integrate on the fly.

5 Conclusion

While complexity may suggest that more control is to be held by the human, different HAT designs may ensure that control and goal achievement can be achieved also in complex and dynamic environments. Tailoring and developing further AI/ML modes may also help the autonomous systems perform tasks in ways that can be adapted to changing environmental demands. A starting point for integration of teammates (e.g., loyal wingmen) and existing human-controlled capabilities (e.g., fighter aircraft) could be to leveling up automation (building adapters and interfaces) in order to design approaches that utilize the best from LOA, MI, and COAD. The article presented a set of environmental characteristics that could be important to consider, e.g., complexity and dynamics, and pointed to trade-offs between the designs in supporting human-machine coordination in varying environmental conditions. We hope that our preliminary discussion will enable teams to collaborate better by providing a common language and process to distribute models and share information about complementarities among the HAT designs. Future experiments, supported by simulation, could aid in developing HAT.

Acknowledgements The authors would like to acknowledge other colleagues at the Norwegian Defence Research Establishment that have contributed in discussions and reading of the paper, as well as officers who have provided us with insight into military problems. Especially, we would like to extend a big thanks to fighter jet projects involved in the development of the loyal wingman concept in a Norwegian context, in particular, Stian Betten and his team as well as the autonomy research projects supervised by Rikke Amilde Seehus and Jonas Moen. We also benefitted from expert input by F-35 pilots from the Tactics, test and training squadron of the Royal Norwegian Air Force, specifically from pilot Ltcol Trond Haugen in early discussions of the topic as well as input on the implications of various HAT designs for F-35 LW collaboration by pilot Major Morten Hanche.

Data Availability The data that support the findings of this study are available on request from the corresponding author, Stensrud et al.. The data are not publicly available due to [restrictions e.g. their containing information that could compromise confidentiality].

Declarations

Competing interests The authors declare no competing interests.

References

- Abbink David A, Carlson Tom, Mulder Mark, De Winter Joost C, F, Aminravan Farzad, Gibo Tricia L, Boer Erwin R (2018) A topology of shared control systems-finding common ground in diversity. *IEEE Trans Hum-Mach Syst* 48(5):509–25
- Allen JE, Guinn CI, Horvitz E (1999) Mixed-initiative interaction. *IEEE Intell Syst* 14(5):14–23. <https://doi.org/10.1109/5254.796083>
- Bekier M (2013) Automation acceptance in air traffic management [UNSW Sydney]. <https://doi.org/10.26190/UNSWORKS/16253>
- Brehmer B (2010) Command and control as design

- Brehmer B (1992) Dynamic decision making: human control of complex systems. *Acta Psychologica* 81(3):211–41. [https://doi.org/10.1016/0001-6918\(92\)90019-A](https://doi.org/10.1016/0001-6918(92)90019-A)
- Cabrall Christopher DD, Thomas B. Sheridan, Thomas Prevot, Joost CF De Winter, Riender Happee (2018) The 4D LINT model of function allocation: spatial-temporal arrangement and levels of automation. In *Intelligent human systems integration*, edited by Waldemar Karwowski and Tareq Ahram, 722:29–34. Cham: Springer International Publishing, https://doi.org/10.1007/978-3-319-73888-8_6
- Chanel Caroline P. C, Roy Raphaëlle N, Dehais Frédéric, Drougard Nicolas (2020) Towards mixed-initiative human–robot interaction: assessment of discriminative physiological and behavioral features for performance prediction. *Sensors* 20(1):296. <https://doi.org/10.3390/s20010296>
- Demir M, Cooke NJ, Amazeen PG (2018) A conceptual model of team dynamical behaviors and performance in human-autonomy teaming. *Cogn Syst Res* 52:497–507. <https://doi.org/10.1016/j.cogsys.2018.07.029>
- Dess GG, Beard DW (1984) Dimensions of organizational task environments. *Adm Sci Q* 29(1):52. <https://doi.org/10.2307/2393080>
- Endsley Mica R (2023) Supporting human-AI teams: transparency, explainability, and situation awareness. *Comp Human Behavior* 140:107574. <https://doi.org/10.1016/j.chb.2022.107574>
- Fitts PM, Jones RE (1947) Analysis of factors contributing to 460 ‘pilot error’ experiences in operating aircraft controls. Aero Medical Laboratory, Air Material Command, Wright-Patterson Air Force Base, Dayton, OH
- Frame Mary E, Alan S. Boydston, Jennifer S. Lopez (2020) Development of an autonomous manager for dynamic human-machine task allocation in operational surveillance.” In *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*, 1–4. Rome, Italy: IEEE, <https://doi.org/10.1109/ICHMS49158.2020.9209414>.
- Frazier SN (2022) Measuring cognitive workload in automated knowledge work environments. A dissertation submitted to the Faculty of Purdue University In Partial Fulfillment of the Requirements for the degree of Doctor of Philosophy. Faculty of Purdue University
- Frey TL, Aguilar C, Engebretson K, Faulk D, Lenning LG (2018) F-35 information fusion. In: 2018 aviation technology, integration, and operations conference, Atlanta, Georgia. <https://doi.org/10.2514/6.2018-3520>
- Gombolay M, Bair A, Huang C, Shah J (2017) Computational design of mixed-initiative human–robot teaming that considers human factors: situational awareness, workload, and workflow preferences. *Int J Rob Res* 36(5–7):597–617. <https://doi.org/10.1177/0278364916688255>
- Goodrich Michael A, Julie A. Adams, Matthias Scheutz (2022) Autonomy reconsidered: towards developing multi-agent systems. In *Intelligent systems and applications*, edited by Kohei Arai, 294:573–92. Cham: Springer International Publishing https://doi.org/10.1007/978-3-030-82193-7_38
- Grote G, Kolbe M, Waller MJ (2018) The dual nature of adaptive coordination in teams: balancing demands for flexibility and stability. *Organ Psychol Rev* 8(2–3):125–148. <https://doi.org/10.1177/2041386618790112>
- Grover, Jaskaran Nishant Mohanty, Changliu Liu, Wenhao Luo, Katia Sycara (2022) Noncooperative herding with control barrier functions: theory and experiments. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, 80–86. Cancun, Mexico: IEEE. <https://doi.org/10.1109/CDC51059.2022.9992986>
- Hamstra Jeffrey W (2019) The F-35 Lightning II: from concept to cockpit. American Institute of Aeronautics and Astronautics, Inc., Reston, VA. <https://doi.org/10.2514/4.105678>
- Hopkins D, Schwanen T (2018) Governing the race to automation. In: Marsden G, Reardon L (eds) *Emerald points*. Emerald, pp 65–84
- Jiang Shu, Ronald C. Arkin (2015) Mixed-initiative human-robot interaction: definition, taxonomy, and survey. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 954–61. Kowloon Tong, Hong Kong: IEEE. <https://doi.org/10.1109/SMC.2015.174>.
- Johnson M, Bradshaw JM, Feltovich PJ (2018) Tomorrow’s human–machine design tools: from levels of automation to interdependencies. *J Cogn Eng Decis Mak* 12(1):77–82. <https://doi.org/10.1177/1555343417736462>
- Johnson M, Bradshaw JM, Feltovich PJ, Hoffman RR, Jonker C, van Riemsdijk B, Sierhuis M (2011) Beyond Cooperative Robotics: The Central Role of Interdependence in Coactive Design. *IEEE Intell Syst* 26(3):81–88. <https://doi.org/10.1109/MIS.2011.47>
- Johnson M, Bradshaw JM (2021) How Interdependence Explains the World of Teamwork. In: Lawless WF, Llinas J, Sofge DA, Mittu R (eds) *Engineering Artificially Intelligent Systems*, vol 13000. Springer International Publishing, pp 122–146. https://doi.org/10.1007/978-3-030-89385-9_8
- Johnson Matthew, Bradshaw Jeffrey M, Feltovich Paul J, Jonker Catholijn M, Birna Van Riemsdijk M, Sierhuis Maarten (2014) Coactive design: designing support for interdependence in joint activity. *J Hum Robot Interact* 3(1):43. <https://doi.org/10.5898/JHRI.3.1.Johnson>
- Kaber DB (2018a) a conceptual framework of autonomous and automated agents. *Theor Issues Ergon Sci* 19(4):406–430. <https://doi.org/10.1080/1463922X.2017.1363314>
- Kaber DB (2018b) Issues in human–automation interaction modeling: presumptive aspects of frameworks of types and levels of automation. *J Cogn Eng Decis Mak* 12(1):7–24. <https://doi.org/10.1177/1555343417737203>
- Kirlik A, Miller RA, Jagacinski RJ (1993) Supervisory control in a dynamic and uncertain environment: a process model of skilled human-environment interaction. *IEEE Trans Syst Man Cybern* 23(4):929–952. <https://doi.org/10.1109/21.247880>
- Kouchaki M, Okhuysen GA, Waller MJ, Tajeddin G (2012) The treatment of the relationship between groups and their environments: a review and critical examination of common assumptions in research. *Group Org Manag* 37(2):171–203. <https://doi.org/10.1177/1059601112443850>
- Lindner Sebastian, Dennis Mund, Axel Schulte (2022) How human-autonomy teams change the role of future fighter pilots: an experimental assessment. In *AIAA SCITECH 2022 Forum*. San Diego, CA & Virtual: American Institute of Aeronautics and Astronautics. <https://doi.org/10.2514/6.2022-2551>.
- Luciano MM, Nahrang JD, Shropshire C (2020) Strategic leadership systems: viewing top management teams and boards of directors from a multiteam systems perspective. *Acad Manag Rev* 45(3):675–701. <https://doi.org/10.5465/amr.2017.0485>
- Lundberg Jonas, Johansson Björn J. E (2021) A framework for describing interaction between human operators and autonomous, automated, and manual control systems. *Cogn Technol Work* 23(3):381–401. <https://doi.org/10.1007/s10111-020-00637-w>
- Lyons JB, Sycara K, Lewis M, Capiola A (2021) Human–autonomy teaming: definitions, debates, and directions. *Front Psychol* 12:589585. <https://doi.org/10.3389/fpsyg.2021.589585>
- Maarten Schraagen Jan, Barnhoorn Jonathan Sebastiaan, Van Schendel Jef, Van Vught Willeke (2022) Supporting teamwork in hybrid multi-team systems. *Theor Issues Ergon Sci* 23(2):199–220. <https://doi.org/10.1080/1463922X.2021.1936277>
- Mathiassen K, Nummedal OR (2022) LandX20 experiment report – experiment for future land warfare capabilities with focus on increased situational awareness and unmanned systems (FFI-Report 22/00274). FFI:978-82-464-3390-5. <https://www.ffi.no/publikasjoner/arkiv/landx20-experiment-report-experiment-for-future-landwarfare-capabilities-with-focus-on-increased-situational-awareness-and-unmanned-systems>

- McNeese NJ, Demir M, Cooke NJ, Myers C (2018) Teaming with a synthetic teammate: insights into human-autonomy teaming. *Hum Factors: J Hum Factors Ergon Soc* 60(2):262–273. <https://doi.org/10.1177/0018720817743223>
- Mathieu JE, Luciano MM, DeChurch LA (2018) Multiteam systems: The next chapter. *The SAGE Handbook of Industrial, Work & Organizational Psychology* 2:333–353
- Mercado Joseph E, Rupp Michael A, Chen Jessie Y. C, Barnes Michael J, Barber Daniel, Procci Katelyn (2016) Intelligent agent transparency in human-agent teaming for multi-UxV management. *Hum Factors: J Hum Factors Ergon Soc* 58(3):401–15. <https://doi.org/10.1177/0018720815621206>
- Miller CA, Parasuraman R (2007) Designing for flexible interaction between humans and automation: delegation interfaces for supervisory control. *Hum Factors: J Hum Factors Ergon Soc* 49(1):57–75. <https://doi.org/10.1518/001872007779598037>
- Minos-Stensrud Mathias, Hans Jonas Fossum Moen, Jan Dyre Bjerknæs (2021) Information sharing in multi-agent search and task allocation problems.” In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–7. Orlando, FL, USA: IEEE. <https://doi.org/10.1109/SSCI50451.2021.9660121>
- Mouloua MP, Hancock A, Ferraro J (2020) Human performance in automated and autonomous systems. *Current Theory and Methods*. CRC Press/Taylor & Francis Group
- Myers C, Ball J, Cooke N, Freiman M, Caisse M, Rodgers S, Demir M, McNeese N (2019) Autonomous intelligent agents for team training. *IEEE Intell Syst* 34(2):3–14. <https://doi.org/10.1109/MIS.2018.2886670>
- Nummedal OR (2021) Flamingo - a UAV for autonomy research (FFI-Report 21/00318). FFI:978-82-464-3307-3. <https://www.ffi.no/publikasjoner/arkiv/flamingo-a-uav-for-autonomy-research>
- O’Neill T, McNeese N, Barron A, Schelble B (2022) Human-autonomy teaming: a review and analysis of the empirical literature. *Hum Factors: J Hum Factors Ergon Soc* 64(5):904–938. <https://doi.org/10.1177/0018720820960865>
- Parasuraman R, Sheridan TB, Wickens CD (2000) A model for types and levels of human interaction with automation. *IEEE Trans Syst Man Cybern- Part a: Syst Hum* 30(3):286–297. <https://doi.org/10.1109/3468.844354>
- Park S, Grosser TJ, Roebuck AA, Mathieu JE (2020) Understanding work teams from a network perspective: a review and future research directions. *J Manag* 46(6):1002–1028. <https://doi.org/10.1177/0149206320901573>
- Pasquale Davide Di, Savill Mark (2022) The importance of coupling aerodynamic and cost analysis in aircraft design. *CEAS Aeronaut J* 13(4):1085–1100. <https://doi.org/10.1007/s13272-022-00600-7>
- Perkins DG (2017) Multi-domain battle driving change to win in the future. *Military Review*. <https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/July-August-2017/Perkins-Multi-Domain-Battle/>
- Petousakis Giannis, Manolis Chiou, Grigoris Nikolaou, Rustam Stolkin (2021) Human operator cognitive availability aware mixed-initiative control. <https://doi.org/10.48550/ARXIV.2108.11885>
- Priebe M, Douglas C (2020) Multiple dilemmas: challenges and options for all-domain command and control. *RAND Corp*. <https://doi.org/10.7249/RR4381-1>
- Rebensky S, Carmody K, Ficke C, Carroll M, Bennett W (2022) Teammates instead of tools: the impacts of level of autonomy on mission performance and human-agent teaming dynamics in multi-agent distributed teams. *Front Robot AI* 9:782134. <https://doi.org/10.3389/frobt.2022.782134>
- Rico R, Hinsz VB, Davison RB, Salas E (2018) Structural influences upon coordination and performance in multiteam systems. *Hum Resour Manag Rev* 28(4):332–346. <https://doi.org/10.1016/j.hrmr.2017.02.001>
- Salas E, DiazGranados D, Klein C, Burke CS, Stagl KC, Goodwin GF, Halpin SM (2008) Does Team Training Improve Team Performance? A Meta-Analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50(6):903–933. <https://doi.org/10.1518/001872008X375009>
- Schmitt Fabian, Gunar Roth, Daniel Barber, Jessie Chen, Axel Schulte (2018) Experimental validation of pilot situation awareness enhancement through transparency design of a scalable mixed-initiative mission planner. In *Intelligent Human Systems Integration*, edited by Waldemar Karwowski and Tareq Ahram, 722:209–15. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-73888-8_33
- Schneider A, Wickert C, Marti E (2017) Reducing complexity by creating complexity: a systems theory perspective on how organizations respond to their environments. *J Manage Stud* 54(2):182–208. <https://doi.org/10.1111/joms.12206>
- Seeber Isabella, Bittner Eva, Briggs Robert O, De Vreede Triparna, De Vreede Gert-Jan, Elkins Aaron, Maier Ronald et al (2020) Machines as teammates: a research agenda on AI in team collaboration. *Inf Manag* 57(2):103174. <https://doi.org/10.1016/j.im.2019.103174>
- Sheridan TB, Verplank WL (1978) Human and computer control of under sea teleoperators. In: Technical Report, man-machine systems laboratory, department of mechanical engineering, Massachusetts institute of technology, Cambridge, Mass (PDF) Levels of Automation in Process Control. Available from: https://www.researchgate.net/publication/270724475_Levels_of_Automation_in_Process_Control. Accessed 1 Feb 2024
- Sheridan TB, Ferrell WR (1974) Man-machine systems: information, control, and decision models of human performance. MIT Press
- Simon HA (1957) Models of man, social and rational. Wiley, Oxford, England
- Stensrud R, Mikkelsen B, Betten S, Valaker S (2021) A proposal for a simple evaluation method in support of the initial concept phase assessing a future unmanned Loyal Wingman for Royal Norwegian Air Force (RNoAF)
- Stensrud R, Mikkelsen B, Valaker S (2020) 5th generation fighter aircraft and future command and control: A note on some possible changes. 25th international command and control research and technology symposium
- Stensrud R, Valaker S (2022) Methods to meet changes in the security environment, ICCRTS 2022
- Stensrud R, Valaker S, Haugen T (2020) Interdependence as an Element of the Design of a Federated Work Process. 2020 IEEE International Conference on Human-Machine Systems (ICHMS) 1–6. <https://doi.org/10.1109/ICHMS49158.2020.9209423>
- Sycara Katia, Dana Hughes, Huao Li, Michael Lewis, Nina Lauharatanahirun (2020) Adaptation in human-autonomy teamwork. In *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*, 1–4. Rome, Italy: IEEE. <https://doi.org/10.1109/ICHMS49158.2020.9209410>
- Vagia M, Transeth AA, Fjerdingen SA (2016) A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? *Appl Ergon* 53:190–202. <https://doi.org/10.1016/j.apergo.2015.09.013>
- Van De Ven AH, Delbecq AL, Koenig R (1976) Determinants of coordination modes within organizations. *Am Sociol Rev* 41(2):322. <https://doi.org/10.2307/2094477>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.