ORIGINAL ARTICLE

Reduced autonomy workspace (RAW)—an interaction design approach for human-automation cooperation

Magnus Nylin[1,2] · Jimmy Johansson Westberg[1] · Jonas Lundberg[1]

Abstract

Lack of support for handling a reduction of autonomy in a highly autonomous automation may lead to a stressful situation for a human when forced to take over. We present a design approach, the Reduced Autonomy Workspace, to address this. The starting point is that the human and the automation work together in parallel control processes, but at different levels of autonomy cognitive control, such as setting goals or implementing plans, which is different from levels of automation. When autonomy is reduced, the automation should consult the human by providing information that has been aligned to the level at which the human is working, and the timing of the provision should be adapted to suit the human's work situation. This is made possible by allowing the automation to monitor the human in a separate process. The combination of these processes, information level alignment and timing of the presentation, are the key characteristics of the Reduced Autonomy Workspace. The Reduced Autonomy Workspace consists of four phases: Identification of the need; evaluation of whether, and, if so, when, and how to present information; perception and response by the human; implementation of a solution by the automation. The timing of the information presentation should be adapted in real-time to provide flexibility, while the level of the information provided should be tuned offline and kept constant to provide predictability. Use of the Reduced Autonomy Workspace can reduce the risk for surprising, stressful hand-over situations, and the need to monitor the automation to avoid them.

1 Introduction

When an autonomous automation that is working side-by-side with a human reaches a limit for its autonomy, there must be a way for the human to cope with a reduction in autonomy. When a human takes over in a stressful situation, he or she must understand not only the situation but also how it was handled by the automation before the autonomy was reduced. It may take more time to reach this understanding than handling it without the automation. Furthermore, the human must maintain knowledge about the fallback procedures that are available to cope with autonomy reduction

should it happen, which may be both costly and time-consuming. In the worst case, the human has not performed the tasks without the automation for a long time, and as a result of this has lost skills. This may cause the situation to quickly become critical. The use of automation in this case leads to more work, a need for higher skills, and potentially new dangers—contrary to the probable reasons that the automation was introduced. Bainbridge (1983) described these issues as "automation ironies". Much effort has been invested to mitigate these problems, but they are still with us, decades later (Baxter et al. 2012; Strauch 2018).

This article introduces an interaction design approach known as the Reduced Autonomy Workspace (RAW) that can mitigate some of these issues. The workspace is especially intended to be used with highly autonomous systems. When autonomy decreases and cooperation between automation and human is needed, the RAW should enable the automation to initiate communication that relates to the situation by consulting the human in a well-timed way that puts

✉ Magnus Nylin
    magnus.nylin@liu.se

1   Department of Science and Technology, Linköping University, Norrköping, Sweden

2   Department of Systems and Development, Air Navigation Services of Sweden (LFV), Norrköping, Sweden

as little additional cognitive load as possible onto the human. This allows the human to decide whether, and, if so, how and when to respond. This should be possible to apply whether it is a major reduction in autonomy calling for the human to take over some of the work, as described in the opening paragraph, or a less dramatic one when both the automation and the human soon can resume autonomous work. The key idea is to allow the automation to initiate communication using information that has been transformed to a format and level that are appropriate for the role and responsibilities of the human. A further aspect of the communication is that its timing has been adapted to the human's situation. The RAW was designed using the Joint Control Framework (JCF) (Lundberg and Johansson 2020), and a case-driven approach using a case from the domain of Air Traffic Management (ATM). The JCF was chosen primarily for its ability to describe temporal aspects of control processes and the interactions between these processes, properties not covered by most other frameworks.

This article first presents related work and the ATM domain with a concept design case based on a future ATM system in which a highly autonomous automation and a human work together. The case is then modelled using the RAW approach. Finally, different aspects of the RAW are discussed and how RAW can be implemented. Some directions for future research are suggested, and conclusions drawn from the modelled case are presented.

2 Related work

Humans working with automated systems is not a new phenomenon, but some issues related to human-automation cooperation remain to a large extent unsolved.

2.1 Still struggling to team up with automation

Baxter et al. (2012) reviewed work that had been carried out during the 30 years since Bainbridge (1983) formulated the now well-known ironies (e.g. that the human's task shifts from working with a process to monitoring an automation that now works with the process; that this is a shift of tasks rather than only a reduction in workload; that humans do not perform well in passive monitoring tasks). Baxter showed that some of these ironies remained to be solved. More recently, Strauch (2018) revisited the ironies and came to largely the same conclusions—they are still valid and, in many cases, unresolved. Strauch described several incidents, mainly from aviation, in which the classical automation ironies were found to be just as valid as previously. He pointed out, however, that safety has increased in the aviation sector, as knowledge about how to deal with the ironies has increased. He also suggested that the development of new types of automation, some of which have greater autonomy, has introduced new ironies or issues. One of those is the autonomy conundrum described by Endsley (2017), which relates to the operator's decreased Situational Awareness (SA) caused by high Levels of Automation (LOA) and reduced possibility to revert to manual control when needed. Endsley considers this to be a major challenge that must be addressed with good design. Trapsilawati et al. (2016) emphasize that it is important to engage the operator in the work and suggests that the reduction in workload for the human, made possible by the automation, may be utilized by the human to maintain SA. Endsley also emphasizes that it is important that SA is shared between automation and human if the former is highly autonomous and works in a team with the human (Endsley 2017). In this case, Endsley states that the LOA is probably high with a low degree of adaptivity, and a low granularity of control for the human. This means that the human's direct engagement in the automation's tasks is low.

The SESAR Master Plan (SJU 2020) describes a development roadmap for European ATM. In SESAR, increased automation is considered to be a key factor to achieve the goals of a single European sky. Five different LOAs are defined, where Level 5 is full automation with no human involvement. It is not expected that this fifth level will be fully achieved for ATM, at least not in the near future. Instead, the master plan points toward adaptive automation concepts. Automation is seen as being able to initiate most tasks, and it is not expected that automation will be able to execute all of them fully. As long as full autonomy is not achieved, automation is described as a support.

Strauch (2018) suggests that some of the ironies simply cannot be solved as long as we keep the premises the same, one of which is that the human should step in when automation gives up. Hence, the question arises of whether there are other ways of tackling the problem.

2.2 Humans and automation working side-by-side

Norman (2015) suggested how to change the paradigm on a conceptual level. He describes how the thinking around automation in the control of highly autonomous cars should change from an emphasis on the assignment of control tasks based on LOA to an approach that focuses on the human-automation cooperation. This avoids the problem of the human having to constantly supervise the automation and act as backup. Instead, the human and automation continuously cooperate. Shively et al. (2018) also focus on cooperation when they show how the principles of Crew Resource Management (CRM) can be used in Human-Automation Teaming (HAT) (Cooper et al. 1980; Kanki et al. 2010; Maurino and Murray 2010). They show that bi-directional communication between

automation and human is a crucial factor. Christofferssen and Woods (2002) strongly emphasized the need for looking at human-automation cooperation as a teamwork. They identified two crucial aspects to achieve this. First, make the automation understandable in a way that requires as little effort as possible from the human, yet creating knowledge about what is going on. Second, ensure directability of the automation, so that the automation is able to adhere to new directions from the human and adapt accordingly. Directability is also one of "Ten Challenges for Making Automation a "Team Player" in Joint Human-Agent Activity" (Klein et al. 2004). Klein et al. also brought up aspects such as mutual predictability and understanding. Roth et al. (2019) focused on function allocation and emphasized that cooperation is an important aspect that must be considered. Earlier, Bradshaw et al. (2013) brought up the need for a focus on cooperation, but at the same time criticized the use of LOA for autonomous systems. They questioned whether tasks can simply be switched between human and automation with no effect on the system. Kaber responded to this criticism and suggested that levels of automation are often confused with levels of autonomy, stating that the problem does not lie with levels of automation as such (Kaber 2018). Kaber also responded to concerns that LOA may not be the right starting point from which to design human-automation coordination and cooperation, concerns raised by, for example, Christoffersen and Woods (2002). Kaber points out that the design sooner or later reaches engineers for implementation, and it is then important that roles are specified so that it is made clear who should do what in which situation, something for which an LOA framework is convenient. Jamieson and Skraaning (2018), in contrast, claimed that LOA has lost its relevance as a basis for the design of human-automation interaction as the field of automation has developed over the years. They mean that LOA has not delivered on predicting effects on important aspects such as situational awareness, task performance, and workload in complex work settings and thus does not help designing new automation. Hence, they argue, it is not meaningful to elaborate on new, more fine-grained LOA frameworks, instead new approaches should be sought for. LOA is not a concept with *one*, single definition. Many variants and taxonomies have been presented since the first definition of LOA by Sheridan and Verplank (1978), some more elaborated and some more simplified, as pointed out by e.g. Kaber (2018) and Wickens (2018). An extensive overview of LOA frameworks was presented by Vagia et al. (2016). Yet, the LOA variants share the basic idea that tasks are to be divided between a human and an automation at different levels, where tasks at each level are assigned either to the automation or to the human, and this does not encourage a focus on cooperation.

2.3 Workspaces to accommodate human-automation cooperation

Some proposed models focus on human-automation cooperation. Itoh and Pacaux-Lemoine (2018) discuss the issue from the perspective of trust. They identified a need for a common workspace in which the automation and a human agent can understand each other and the cooperation process. Pacaux-Lemoine and Flemisch (2019) also discussed this common workspace and present an example from Air Traffic Control (ATC) (previously presented in Lemoine et al. (1996)). Cooperation centers mainly around tasks and takes place within any of three layers of cooperation, which are linked together by the workspace. Flemisch (2019) discusses the same model and combines it with theories of control sharing. He also introduces a temporal aspect, by showing a snapshot of a situation with different control-sharing states at two distinct points in time. Gutzwiller et al. (2018) suggested that cooperation in human-automation teams should be organized by pre-defining task allocation under different conditions in a working agreement, and including transition points that define when to change the allocation of tasks and responsibilities. In addition to the already mentioned older work of Lemoine et al. (1996), the workspace notion in the ATC context was used in a later follow up study (Pacaux-Lemoine and Debernard 2002), and by Hoc and Debernard (2002), the main focus there were task distribution. Riera and Debernard (2001) had a similar idea around a workspace in ATC context, and also highlighted the need for a frame of common references. Jones et al. (1997) proposes a workspace to keep track of progress in coordinated and automated activities, including pre- and -post conditions. Vanderhaegen (2021) focuses on human-automation coordination (availability, possibility to act, competence).

2.4 Levels of automation and levels of cognitive autonomy

As already pointed out, LOA taxonomies may be adequate to use to describe a system with respect to how actions and responsibilities are split between human agent(s) and automation. Vagia et al. (2016) distinguish between autonomation and autonomy where autonomy is the ability to make decisions without the need for external input, referring to Albus et al. (1998), while automation only is about following a fixed set of rules. None of the LOA taxonomies contain information about the joint control processes where tasks are split up between human and automation and the interaction between those processes. For instance, is the automation autonomously carrying out a given plan or is it also setting goals and making new plans? This could all take place within the same LOA, but they are distinctively different in terms of cognitive autonomy.

Lundberg and Johansson defined a Joint Control Framework (JCF) (Lundberg and Johansson 2020) where they introduce six Levels of Autonomy on Cognitive Control to assess and describe cognitive autonomy and limits of cognitive autonomy (Table 1). The six LACC levels (Table 1) were derived from decades of scientific work in control of complex systems, and include core levels and abilities of models of control and situation awareness. With this LACC, it is possible to assess:

1. System core capability: What kinds of control tasks can the AI/Automation/Autonomous System perform? (i.e. which of the LACC levels can it work on?)
2. System performance limit: Having described the overarching system capabilities, how good is the Autonomous System at those tasks, at those levels, and what performance envelope does it have?
3. Human-automation/AI collaboration requirements: Having characterized the performance envelope, this gives us LACC levels on which humans and automation need to collaborate.

This characterizes the system in terms of what limits and strengths of the system is regarding the LACC and also characterizes the human-AI/Automation collaboration requirements.

The relation between the LACC and LOA is shown in Fig. 1 The combination of different LACC and LOA forms an automation competence scheme**.** Note that Fig. 1 is only one example, the LACC can be combined with any LOA taxonomy of choice. Also note that the interactions for any LACC can take place at any LOA. Hence, an automation can act on different LACC, at any LOA, and, at least in theory, act at any level of autonomy. If we apply this on human-automation cooperation, we can imagine a human agent that sets the goals and frames. The automation works autonomously within the frames and rules set and decides on *how* to achieve them, though the LOA may differ depending on the task to be performed. However, if the automation can make its own decisions

and cannot communicate with the human, it needs to be fully autonomous but also works within rigid frames with a loss of flexibility. To avoid this, communication with the human agent is needed when reaching the limits for the automation's autonomy. The LACC will help assessing what human and automation will collaborate on. The LACC-LOA thus gives both an assessment of division of work, and collaborative work, the latter which is of most interest for the RAW approach.

The JCF also includes an associated score notation that enables the interaction of an agent with a process to be described and analyzed over time. One example of an agent is an air-traffic controller (ATCO) who controls air traffic. The framework can be used to model processes for both a human and an automation, and the interaction between them. The model in this way describes the processes and makes it easier to understand them. The JCF can be used to analyze not only existing systems but also non-existing, first-of-a-kind systems (Lundberg et al. 2018).

The JCF combines ideas from several previous frameworks for modelling of control processes, including the Extended Control Model (Hollnagel and Woods, 2005). The JCF not only brings these ideas together into one framework; the provided score notation adds a temporal extension of the joints between the control processes and controlled processes that are modelled in the JCF.

In each process, a subject (such as an ATCO) interacts with an object (such as an aircraft). Each point of interaction is called a "cognitive joint" and may relate to either perceiving information, making decisions, or performing an action. These interactions are known in JCF as perception points (PP), decision points (DP), and action points (AP). While PPs and APs correspond to more explicit interactions with the object, DPs are more of a subject-internal nature, even though they still relate to the object in the control process and affect the object through the subsequent actions. Each process interaction can take place at any one of six Levels of Autonomy in Cognitive Control (LACC), spanning from high-level framing of the situation to physical interaction with interfaces, Table 1 (Lundberg and Johansson 2020).

Table 1 Levels of autonomy in cognitive control (LACC), adapted from (Lundberg and Johansson 2020)

Level	Description
6 Frames	Defines the situation and context, e.g. to maintain safe and efficient air traffic in a certain airspace at a certain time
5 Effects	Effect goals within the situation, e.g. keeping aircraft apart by a certain distance or striving towards the minimum delay of traffic
4 Values	Trade-offs between different criteria and effect goals
3 Generic	Making up plans and how to implement them to achieve what has been decided upon at higher levels
2 Implementations	Actual implementation of plans, setting constraints and making implementation decisions
1 Physical	Physical world object status, e.g. positions of aircraft on a radar screen

Fig. 1 This matrix shows the Levels of Autonomy in Cognitive Control (LACC) and how it relates to LOA. Different LACC can occur at any LOA. A system that can use or go between several LOAs, or that uses sub-systems with different LOAs require several columns. The LACC can be used in several ways: firstly, to set a limit on the maximum capability of the agent. This also characterizes the autonomy of the agent. An agent that can make trade-offs and select plans or plan based on that, is very different from an agent that only has one plan that is executed over and over again. Second, to describe limitations and capabilities at each level. Can the automation for instance create new plans on the "generic level"? Can it make trade-offs between effect goals at the values level, and can it compute or estimate the actual state of the values in the process? Third, the LACC can be used to describe joints between control systems, where a system requires help from another system and for instance presents information or accepts steering from the outside

3 Designing a reduced autonomy workspace

A few basic ideas were used as starting points from which we built the RAW approach. The automation is considered to be a highly autonomous, cognitive agent that works together with a human agent in different roles at different levels of cognitive control. This combines several views on humans and machines. Arguably, LOA is a prosthesis (Roth et al. 1987) progression, where control gradually shifts to the (more efficient or competent) machine. In contrast, managing breakdowns may require a tools/instrument view (Roth et al. 1987), to amplify skilled abilities in active work with a problem:

> Unanticipated situations inevitably arise in problem-solving situations and confound preplanned response strategies. (p 502).

This means that some means must exist to manage breakdowns in machine activities. Roth et al. (1987) further suggests focusing on increasing the abilities of a human–machine ensemble, with a key factor being the display of a shared frame of reference. This goes towards a more collaborative view, much in line with that of automations as agents, that, as Lee and Seppelt (2012) observe, works on the behalf of the human. With the agent view, cognitive competence becomes central. Thus, as a basis for RAW, the shift from task-sharing based on LOA to a more elaborate cooperation process based on LACC is central. Further, the temporal aspect of the cooperation between the automation and the human is a key characteristic of the work on the RAW, that stems from the focus on time-to-control by a subject versus time-of-change in a process in the Joint Control Framework (Lundberg and Johansson 2020).

The RAW was designed by using a scenario-driven approach. It was initially designed as a theoretical concept using an analysis framework, and the ideas were subsequently applied to a specific case. The case used was an ATM case, as ATM was considered to be a relevant domain. The case was constructed as a scenario in an air traffic control real-time simulator, which enabled the temporal aspects of the RAW to be studied. This case was analyzed in-depth using the JCF score notation. The scenario was deliberately kept relatively simple to make it useful for the purpose of explaining the RAW principles. Even though the RAW has been exemplified within a specific domain, its principles relate to any type of autonomous automation that needs to cooperate with a human. In the rest of the chapter, we describe the RAW principles in more detail, present the domain, including the case, and the analysis framework.

3.1 Application context: air traffic management

Air Traffic Management (ATM) is a domain in which automation has increased more or less constantly, from the introduction of radar to today's complex systems with a multitude of tools and sub-systems that assist the ATCOs. The increase in automation has been driven mainly by demands for higher efficiency, and at the same time maintain high safety standards. However, the level of automation is still relatively low and differs between systems, although higher automation is seen as a key component for the success of future ATM systems (Federal Aviation Administration 2019; SJU 2020). Even though the systems may reach a higher level of automation and become more autonomous than they are today, it is expected that ATCOs will work with the ATM systems for the foreseeable future, while it is accepted that the roles may change (SJU 2020).

3.2 The scenario

The RAW targets highly autonomous systems. As the autonomy in most currently used ATM systems is limited, a scenario in an imagined, future ATM system was created. Lundberg et al. (2018) used a similar approach to evaluate a non-existing, first-of-a-kind system, though both the technology and the traffic situation were first-of-a-kind in the work presented there. In the work presented here, the autonomous system was fictitious, while the traffic situation was not. In this scenario, a highly autonomous system works side-by-side with the ATCO. The main automation manages the traffic on a tactical level, solving conflicts and carrying out plans made by the ATCO, typically at LACC between 1 and 3. For this, we assume that the automation has the ability to create a correct and extensive world model of the world in which it is operating as well as the processes going on and tasks being performed. The ATCO plays an active role, and works also on a tactical level, although with a longer time horizon and on a higher level of cognitive control, typically at LACC between 4 and 6. The ATCO makes plans, sets up goals, and coordinates with other stakeholders.

The scenario was built around a traffic situation with a high-level crossing of two aircraft at the same altitude, in which several solutions are available: The trajectories of two aircraft conflict, Fig. 2. Aircraft SAS123 is crossing slightly behind the route of KLM456 at an acute angle. If nothing is done, the distance will become shorter than the allowed minimum separation (5 nautical miles, NM). The automation calculates that the most efficient solution is to turn SAS123 slightly to the right until it is free of traffic, i.e. KLM456. However, this would bring SAS123 too close to the adjacent sector. How close *too close* is varies, but a typical rule is that distance required to a sector border is half the distance required as separation between two aircraft. This
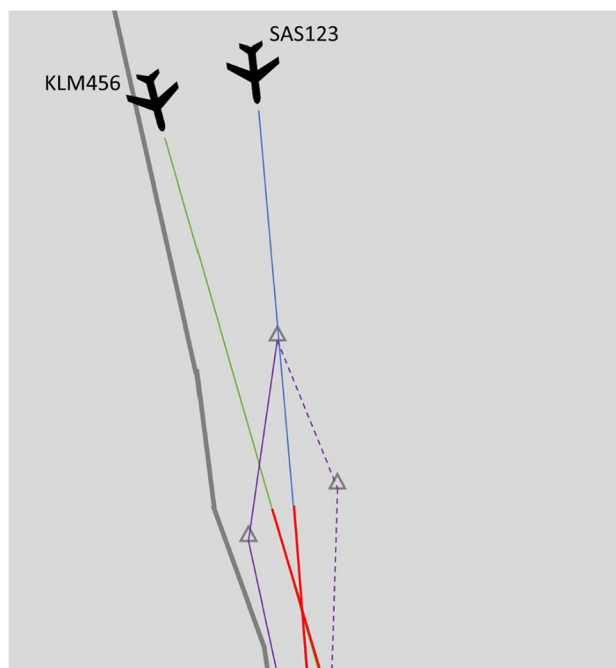

Fig. 2 The traffic scenario (not to scale). The flight paths of SAS123 (blue line) and KLM456 (green line) will cross each other, and the aircraft will come too close (red parts of the flight paths). The automation calculates that the optimal solution is to turn SAS123 slightly to the right (full purple line), but that it will bring it too close to the sector border (thick grey line). The triangles show the turning points. The second-best alternative is a left turn (dashed purple line), but this gives a longer way route for SAS123. The turning point where the purple lines begin is the latest point calculated by the automation at which it is possible to give the turn command to the aircraft to achieve the intended effect. Hence, the response from the ATCO must be given before SAS123 reaches this point. We assume if no, or negative, response is given from the ATCO, se automation proceeds with the less optimal but still safe second-best option

ensures that aircraft are always separated by more than the minimum permitted, even if they are on each side of a sector border. The second-best option is to turn SAS123 left, but that results in a longer flown distance.

In this situation, an ATCO would decide whether it is a good idea to call the ATCO in the adjacent sector to ask for permission to fly closer to the sector border than the rule prescribes. The automation establishes that it would probably not be a problem. However, the automation does not have the jurisdiction nor the means to coordinate this with the ATCO in the adjacent sector. Furthermore, the ATCO working together with the automation may possess additional knowledge about the situation. The automation concludes that it is time to consult the ATCO, to establish a RAW. This kind of trade-off is not uncommon, and a normal part of the work of an ATCO, and dealing with them is a one of the many skills that an ATCO must possess.

It is important to remember that the example in the case is an extract from a specific situation, used to illustrate the

ideas of the RAW. It is assumed that other solutions, such as changing speed or altitude, have already been rejected by the automation based on consideration of other factors such as wind conditions and other traffic, leaving changes in the lateral route as the best alternative. Furthermore, one might argue that an autonomous automation should be designed such that it can handle this kind of situation and make the necessary judgements itself. That may be possible in a predictable reality where all cases can be identified and defined. However, the world is complex and even if it is possible to solve this scenario in other ways, there will always be situations in which the automation is exposed to unanticipated situations, or situations in which its autonomy limits are exceeded.

3.3 Analysis framework

With its ability to model interactions at different LACC and how they are situated in time in, and between, different control processes, the JCF (Lundberg and Johansson 2020) was chosen as a suitable framework to use for the RAW.

The JCF provides a score notation for interactions where a process is visualized by six parallel lines. Each line represents a certain LACC, and the horizontal extension of the lines represents time. The joints are depicted as dots in the score, just as notes in a sheet-music score. The vertical position of a point in the score depicts the LACC at which the interaction takes place, and the horizontal position depicts the time at which it occurs. Johansson and Lundberg (2017) used a similar notation in which the cognitive joints were distributed in time, but their notation did not depict cognitive levels. The JCF enables not only principles to be modelled, but also specific episodes, due to its temporal extension, which is well compatible with the ideas of RAW. The score notation also makes it possible to visualize several simultaneous processes, similar to a musical score. To make them easier to read, the example scores have been enriched with arrows that depict the flow of cognitive joints (Figs. 3, 4, 5, 6).

4 Resulting designs and analyses

The scenario was implemented in a real-time Air Traffic Control (ATC) simulator. However, since the case includes a highly autonomous automation that does not exist, the different solutions were scripted, which means that the aircraft flew as if they had received the instructions from the automation. The simulator was used as a means to visualize the traffic situation that was to be modelled using the JCF score notation. Although using a scripted scenario, real-time playback showed a realistic temporal progress of the traffic situation.
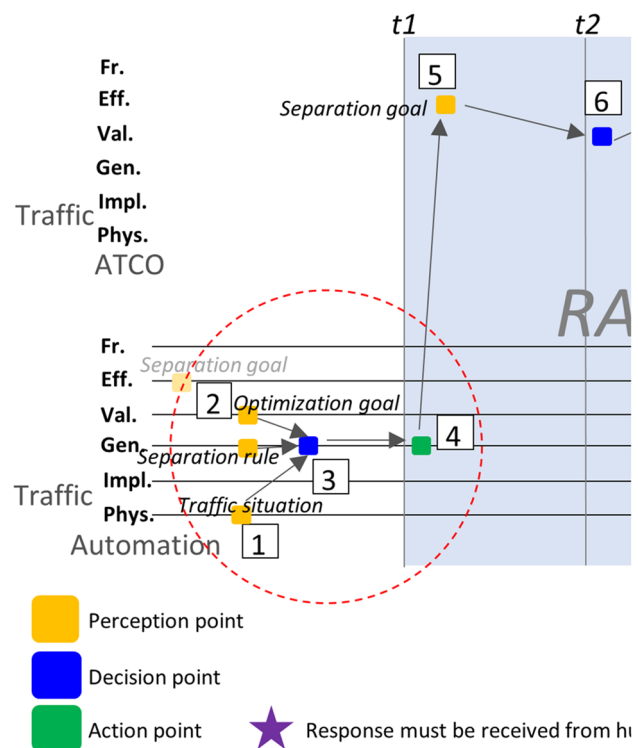

Fig. 3 Joint Control Framework (JCF) score to visualize the ATC example. The time is not to scale for readability reasons. The automation detects the conflict between the two aircraft (1). It considers the rules and goals at hand (2) and concludes that an optimal solution would require an exception from the rule stipulating how close to the sector border an aircraft is allowed to fly. It determines that it must consult the ATCO (3), which is done (4)

4.1 Aligning the LACC of the information

A key element of the RAW design is to identify the level of information (Table 1) to be provided by the automation to the human when initiating the communication. Note that we only model one human stakeholder in this case, the ATCOs working directly together with the automation. Over time, there may be other stakeholders involved who gives input through other processes. The automation and the human work in parallel processes with the same objects, and share the frames and goals set by the human. They work, however, on different LACCs. To illustrate the RAW principles, two JCF scores were initially used. One score showed the process of the ATCO controlling the air traffic on a high level by setting goals, Level 5 (Effects). A second score showed the automation working with the same process on a lower level, solving tactical problems and implementing the solutions. Figures 3, 4, 5 show step-by-step how the situation is identified by the automation, how information is transformed, sent to, and responded on by the ACTO, and finally how the automation receives the response from the ATCO and implements the solution. The work is performed mainly on
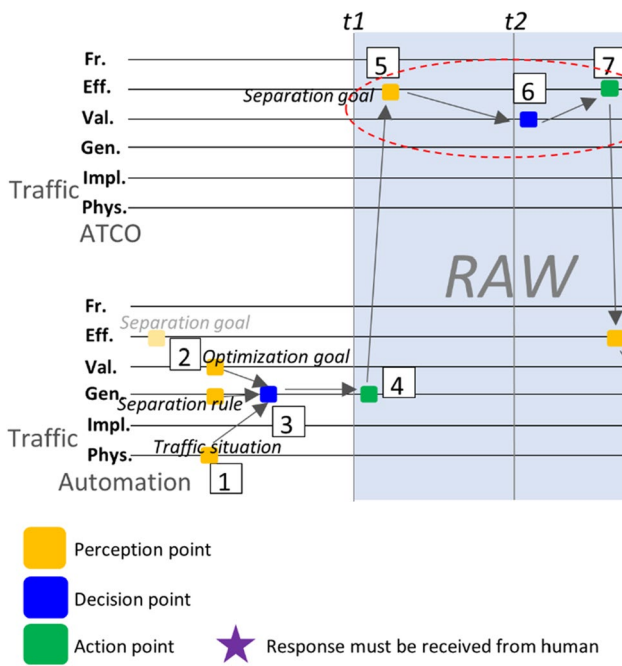
Fig. 4 The information that the automation needs to consult the ATCO is presented to and perceived by the ATCO (5). Note that it is the level of the information that is indicated in the score, not how, or from which HMI it is received. The ATCO decides on how to act (6). In the example, the ATCO decides to coordinate with the ATCO in the neighboring sector. The result of this is sent to the automation (7)
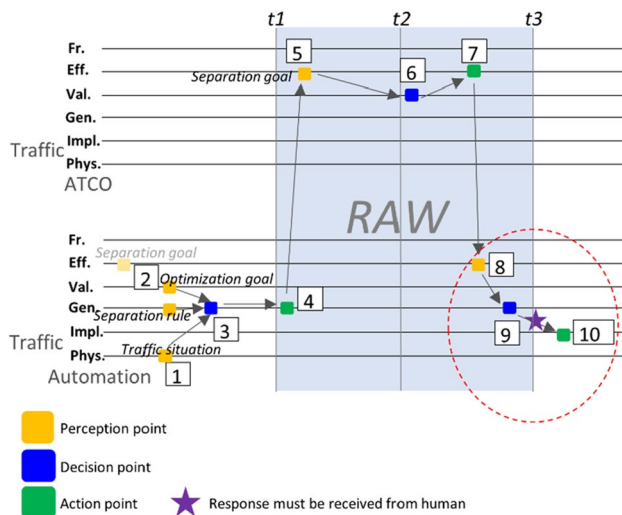

Fig. 5 The automation receives a response from the ATCO (8) and transforms this into a decision about which solution to implement (9). It then implements it (10). If the ATCO gave a positive response, the main alternative is implemented, otherwise the automation goes for the second-best alternative. The solution can be implemented well before t3 if the ATCO responds earlier
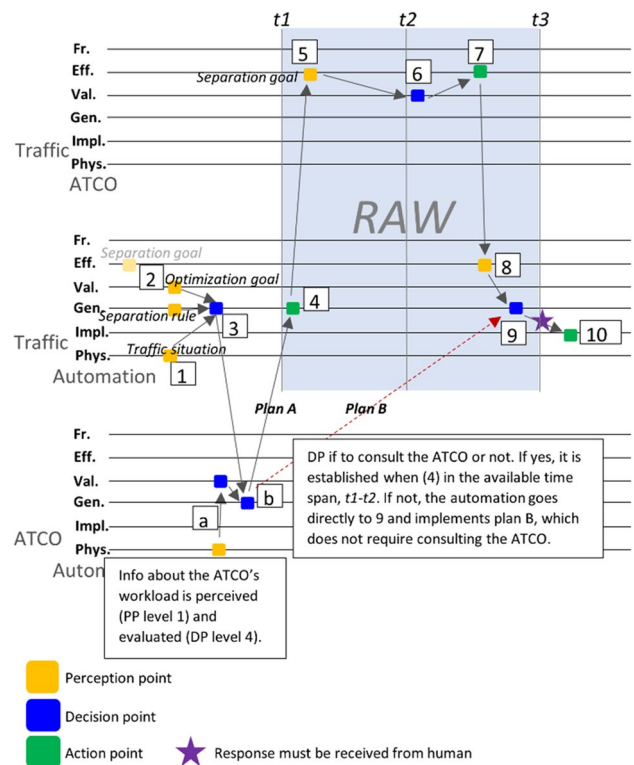

Fig. 6 Introduction of the process by which the automation monitors the ATCO to be able to adapt the RAW to the ATCO's situation. When the automation has decided that an RAW is to be established (3), it checks the situation of the ATCO (**a**) with respect to such matters as workload and attention. It combines this information with when response is needed (t3) and decides at (**b**) when to initiate the RAW (at a suitable time between (t1) and (t2)). If no convenient time can be found, the RAW should not be established, and the automation implements the second-best alternative (9) without consulting the ATCO

Levels 1, 2, and 3 (physical, implementations, and generic), though it is governed by goals set on Levels 4 and 5 (values and effects). The automation identifies the conflict between the two aircraft (Fig. 3, Point 1), compares the rules of separation and the goal of optimizing the traffic (Fig. 3, Point 2). A decision is made that the ATCO should be consulted (Fig. 3, Points 3 and 4) to decide whether if it is a good idea to ask the neighboring sector for approval to implement the most efficient solution.

When initiating the communication about the RAW, the automation must be able to transform the low-level information into an information package that can be presented and understood at the level at which the human agent is working (Fig. 4, Points 4 and 5) to avoid that the ATCO must switch cognitive levels when prompted by the automation. Consider the differences between the two descriptions of the problem with the SAS123-KLM456 crossing presented in Table 2.

The first (Table 2) is a medium–low level description of the problem, Level 3. It would force the ATCO to dig into

Table 2 Description of one same problem at different LACC

LACC	Description
3 Generic	I would like to turn SAS123 right by x degrees. Is that OK?
5 Effects	I would like to let SAS123 fly 1.7 NM closer to Sector X than allowed to solve a conflict optimally, no conflicting traffic detected in Sector X. Is that OK?

the situation at Level 1 or 2 to understand why the automation wants to turn the aircraft (an optimized conflict solution) and the consequences of doing so (coming too close to the adjacent sector). It is not clear that this is about finding an *optimal* solution, which may raise questions from the ATCO about alternatives. This low-level information has to be transformed to a higher LACC by the ATCO to make it possible to compare them to the higher-level goals in order to make a decision.

All of this takes time, and may take the ATCO out of the loop of his/her work at higher levels. The second (Table 2) description clearly states that it relates to a conflict between a rule (required distance to the sector border, safety) and a goal (optimization), i.e., this description relates directly to the goals and trade-offs at Levels 4 and 5, where the ATCO is working, and makes it clear that no issues with respect to traffic other than that in the conflict to be solved have been identified. It is only necessary for the ATCO to decide whether it is feasible to coordinate with the adjacent sector to obtain permission for an exception from the border separation rule to reach the optimization goal. This is a Level 4 decision. Both descriptions are at the same LOA—the automation detects the problem and suggests a solution for the ATCO to approve or not approve. The difference is at what LACC it is communicated.

Note that when presenting the information to the human, the action to initiate the RAW (Fig. 4, Point 4) is taken on a lower level, 3, than the level of the information perceived by the ATCO, 5 (Fig. 4, Point 5). This transformation of the information is very important and one of the core ideas of the RAW, i.e. to align the communication with the human to the level at which the human is currently working. The human agent can then evaluate the information in a controlled manner starting from his/her current level. A decision is made on a slightly lower level, 4, while the action of providing the response is again located on Level 5 (Fig. 4, Points 6–7). This gives the ATCO the possibility to pace the process, and the LACC alignment should reduce the need for extensive information gathering at lower LACC.

The response is passed to the automation, which transforms it back to a low-level solution that is compliant with the revised goal (Fig. 5, Points 8–10). If positive, that is if the ATCO has deemed it a good idea to coordinate with the adjacent sector and has provided an affirmative answer, the automation implements the solution by modifying the route

of SAS123 to include a right turn. If negative, the automation implements the second-best alternative, and modifies the trajectory of SAS123 to include a left turn. During the procedure, it has not been necessary for the ATCO to consider the details of the possible solutions, and he/she has been able to continue to work with goals and trade-offs between the goals. The automation has made several transformations between cognitive levels, and in this way made such transformations unnecessary for the human. The ATCO may glance at the radar screen during the process, and in this way perceive information at a low LACC, to obtain an overview of the situation. That is, however, different from using the radar screen to obtain detailed, low-level information of each control process in order to use this information to determine how to follow high-level rules and goals when making a decision. Further, of course, nothing prevents the ATCO from taking in low-level information while it is present. What is important here is that the goal is that this information should not be *needed* for the RAW process, and the human can decide whether to gather it instead of it being forced upon him/her.

This ability to transform and align the information to suit the ATCO is a crucial aspect of the automation's competence. If the same procedure is always followed when initiating the RAW, the communication will be highly predictable – the human knows what to expect. Together with the timing of the communication, the RAW moves away from the unwanted, stressful handover situations and acts as a cooperating system. The process of trying until failing and then issuing an alarm is no longer necessary.

4.2 Rhythm and timing by adaptation

Initiation of the RAW, and thus the communication with the ATCO, must be made in a sensible way with respect to rhythm and timing in the information flow to avoid workload peaks. Hence, the automation must be provided with knowledge about the situation of the human with respect to attention, workload et cetera. For explaining the concept, we assume that this is available. In reality, this could be derived using techniques such as eye tracking and real-time analysis of system interaction. For this, we introduce a third JCF score, which shows how the automation can act when it knows what the ATCO is doing (Fig. 6, lowest score).

When a decision has been made to establish a RAW (Fig. 6, Point 3), the automation knows when it must obtain a response from the human to be able to implement the solution (Fig. 6, *t*3). The human must have sufficient time (Fig. 6, *t*2-*t*3) to reflect upon the RAW information and decide whether to respond and if so, how. The minimum time required depends on the application domain, but it should be a predefined, fixed amount, to provide predictability. Even if a certain amount of time is needed for the human to deal with the RAW, initiating it too early may result in an overload of RAWs. Furthermore, looking too far into the future will increase uncertainty due to the complexity of the real world: the weather may change, or other unforeseen events occur.

Early in the process, it is difficult to predict whether a situation will develop into a problem at all. Therefore, a RAW horizon should be established that defines the earliest time at which the automation can initiate a RAW (Fig. 6, *t*1). Consequently, a RAW can be presented between *t*1 and *t*2, the exact timing of which depends on the situation. While *t*2 is a fixed time, it is suggested that *t*1, the RAW horizon should be adaptable. The automation uses the knowledge about the ATCO's situation (Fig. 6, Point a) to decide on when between *t*1 and *t*2 to initiate the RAW (Fig. 6, Points b, 4, and 5). Even if *t*1, the horizon, is set by the human operator, the possible values depend on the implementation and domain-specific conditions.

Figures 3, 4, 5, 6 suggest that there is plenty of time, but the scores show only a few processes. It is probable that there are many more processes and events that affect the situation. These include other aircraft, the need to coordinate with other stakeholders, strategic planning activities, and so on. All of this must be considered when the automation decides when it is appropriate to establish the RAW, i.e. Points a and b, leading to the action at Point 4 in Fig. 6.

Importantly, the RAW is intended for use only in non-critical situations. Hence, if no response is given for some reason, the only effect will be a decrease in efficiency, not a decrease in safety. However, if missed responses start to occur more frequently, it is a clear signal that something is not working optimally. Critical issues and situations must continue to be handled by alarms and back-up systems and procedures.

5 Discussion

The JCF scores (Fig. 3, 4, 5, 6) clearly show the parallel human and automation control processes of the analyzed scenario, and how the processes are related to each other. It makes clear the importance of the temporal aspects of the RAW and how the automation should work with it. The order of the joints in the RAW follows the same pattern of

perception, decision, and action. It is critical that the ATCO has enough time between perception and decision (Fig. 6, Points 5 and 6). To achieve this, the automation must possess knowledge about the workload of the ATCO. Though illustrated as a point in the score (Fig. 6, Point a), it should be an ongoing task for the automation to measure and evaluate the workload continuously, so that the information is always available when the need for an RAW arises. When this happens, the automation must be able to answer three questions before establishing the RAW: "*Shall I consult the ATCO*, and if so, *when* should I do it, and *how*?". To answer the first question (Fig. 6, Point b), the automation must evaluate the possible benefit of obtaining help from the ATCO against the risk of overloading the ATCO. This means that the automation must not only know what the workload is, but also have access to a calibrated workload limit to compare it with.

The consequences of *not* following the RAW principles can be made clear by deconstructing the RAW analysis piece by piece. If the first question, *if*, is not addressed, the automation will always present the information, regardless of the situation. What has then been created is, with respect to the temporal aspect, an alarm. Points 3 and 4 (Fig. 6) take place practically simultaneously, and the information is presented to the ATCO earlier than necessary. If the first question, *if,* is addressed but not the second one, *when*, there would still be an alarm-like communication, but fewer occurrences. Finally, if the first two questions are addressed but not the third, *how*, the information is presented at the same level as that at which the issue arose, moving Point 5 (Fig. 6) in the analysis from LACC 5 to 3. This corresponds directly to the differences between the two LACC-dependent phrasings presented in Table 2. Thus, even if the timing is correct, the ATCO must do more work to understand the information and its relevance to the situation and the LACC at which he or she is working. This may, in turn, take time, and make it harder for the automation to estimate the time needed, creating a vicious circle of self-reinforcing bad timing. Remember though, that just asking the questions if, when, how, will not achieve the result we want without a thought-through way of answering them. The important part, and the novelty of the RAW, is *how* this should be done by using the LACC alignment and the timing based on the different control processes as modelled in the JCF framework.

Though using the same notion of "workspace" as in the models presented by Pacaux-Lemoine and Flemisch (Pacaux-Lemoine and Flemisch 2019), the RAW differs in some important aspects. In the Pacaux-Lemoine and Flemisch models, the workspace is separate from the controlled external process, and relates only to the communication between the automation and the human. Furthermore, the automation is still considered to be a support system. Instructions are passed through an interface, affect the process, and propagate back to the task goals in the process, in a cycle with no

defined duration. The RAW also differ from the described earlier workspaces by placing the interactions within and between the control processes on a timeline, thus adding the temporal aspect.

5.1 Working together

We have used a specific scenario to show and analyze the consequences of the RAW, but the principles are generic: similar scenarios should follow the same pattern for the human-automation cooperation. Similar situations are, for example, trade-offs between local and global optimizations and the receipt of inconsistent incoming data, which may confuse the automation. The parallel-process JCF scores in the analysis not only make the durations clear; they also elucidate the cooperation between the automation and the ATCO and how they work with the same object, the air traffic.

By working this way, side-by-side in a team, the human and the automation are continuously involved in controlling the same external processes, albeit on different levels. Hence, they both have an understanding of the situation and work alongside each other. If properly designed by the RAW principles, there should be no need for the human to understand in detail how the automation works. If not so, it imposes demands on the human not only to master the own control processes, but also to master the automation and how the automation handle *its* control processes, and then we are back to the automation ironies. To avoid that, it is important that the human can trust that the automation will ask for consultation when needed, and that the automation will adapt the communication to suit the situation of the human. In other words, the alignment of LACC and timing of information illustrated by the JCF scores in our example, are central. These are, of course, the same requirements as those of humans working together – we do not have to know in detail how other people are working, but we need to trust them in having the competency required, and that they are able to consult us in an appropriate way when needed. The design of the RAW can make this easier by providing predictable behavior: the human should not be surprised by issues communicated at an inconvenient LACC or with bad timing that interrupts the human's work or leaves too little time for response.

5.2 Future research

We consider RAW a design approach that should be suitable for any domain in which human operators work with highly autonomous systems. Vessel Traffic Service (VTS) and train control are closely related domains that share many properties with ATM with respect to both control processes and the control room environment. Future work should look at extending the RAW principles into other domains, and address how they can be used in more complex settings with teams larger than one human and one automation. Complex experiments might require quite some effort, but they are necessary to prove the ideas and produce knowledge needed to further develop the ideas of the RAW.

The RAW interaction design approach principles must be supported by human–computer interfaces (HCI) that facilitate the RAW. Their appearances, the look-and-feel, is a question for future research, and must be adapted to the particularities of the domain, the visual context, and the processes to which they are to be applied.

It is probable that artificial intelligence (AI) will be part of future systems, and it may make systems with learning abilities possible. However, if the automation can learn and improve its behavior, this will result in changes to the behavior. It will be important to investigate the ways in which this should be allowed and how it affects predictability and trust in human-automation cooperation.

Finally, can the information about when, how often, and why RAWs are established be used? It is possible that it can be used as a system-wide performance indicator. This would require not only monitoring the system status but also analyzing the data offline to gain deeper understanding of the human-automation cooperation. The possibility of using these data in an aggregated form by a supervisor role to indicate the overall status of the system could also be investigated.

6 Conclusions

We describe a design approach of a Reduced Autonomy Workspace (RAW) and analyze a case the Air Traffic Management (ATM) domain. From this, we draw four main conclusions:

Firstly, a notation for the control processes and interactions over time is needed to describe the RAW. The key characteristics are the adaptation of timing and the alignment of information with respect to Levels of Autonomy Cognitive Control (LACC) in the interaction between the processes of the automation and the human. The interactions are described in the Joint Control Framework (JCF) as cognitive joints.

Secondly, the occurrence of the cognitive joints in the RAW can be split into four phases:

1. Identification by the automation of the need for a RAW.
2. Evaluation by the automation of if, when, and how to present the RAW to the human.
3. Perception of the RAW by the human, and response to it.

4. Implementation by the automation of a solution based on the response from the human.

Thirdly, temporal adaptation must be carried out in real time, but the definition of target levels for the transformation and LACC alignment of information should be carried out offline. The rationale is that the level of cognitive control at which the human should work is expected to be fairly constant. If this procedure is rigorously followed, predictability can be maintained.

Last, but not least, the RAW modelled in this paper includes three processes: The work of the Air Traffic Controller with the traffic, the work of the automation with the traffic, and the monitoring by the automation of the ATCO's situation. This means that it is not necessary for the ATCO to monitor the automation. (Such a need, if present, can be shown in a fourth score, a process in which the ATCO is the subject, and the automation is the object to be monitored.)

To summarize: the RAW design approach can solve some issues often encountered in human-automation cooperation by allowing the automation to determine if, when, and how it should consult the human when autonomy has been reduced using the novel approach provided by the RAW. The emphasis on LACC rather than only Levels of Automation (LOA) as a starting point is a key to the RAW approach. By following the RAW approach, the risk for overloading the human by initiating communication at an inconvenient time is reduced, while sufficient time for the human to respond is maintained. The risk for surprises is reduced by providing information at a consistent level of cognitive control that is aligned to match the level at which the human mainly works. Finally, as the automation only consults the human if this is possible, it is not necessary for the human to continuously monitor the automation to try to foresee what the automation is up to.

Availability of data and materials Not applicable.

Code availability Not applicable.

Declarations

Conflict of interest There are no identified conflicts of interest.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Albus J, Antsaklis PJ, Meystel A, Passino K, Samad T (1998) Autonomy in engineering systems: what is it and why it is important? Proceedings of the 1998 IEEE ISIC/CIRA/ISAS Joint Conference: 520–521

Bainbridge L (1983) Ironies of automation. Automatica. https://doi.org/10.1016/0005-1098(83)90046-8

Baxter G, Rooksby J, Wang Y, Khajeh-Hosseini A (2012) The ironies of automation … still going strong at 30? ECCE '12: Proceedings of the 30th European Conference on Cognitive Ergonomics: 65–71

Bradshaw JM, Hoffman RR, Johnson M, Woods DD (2013) The seven deadly myths of "autonomous systems." IEEE Intell Syst 28(3):54–61. https://doi.org/10.1109/MIS.2013.70

Christoffersen K, Woods DD (2002) How to make automated systems team players. Adv Hum Perform Cogn Eng Res 2:1–12. https://doi.org/10.1016/S1479-3601(02)02003-9

Cooper GE, White MD, Lauber JK (1980) Resource management on the flight deck. Proceedings of a NASA/Industry Workshop Held at San Fransisco, California June 26–28,1979: 1–16

Endsley MR (2017) From here to autonomy: lessons learned from human-automation research. Hum Factors 59(1):5–27. https://doi.org/10.1177/0018720816681350

Erik Johansson BJ, Lundberg J (2017) Resilience and the temporal dimension–the chimera of timely response. Theor Issues Ergon Sci 18(2):110–127. https://doi.org/10.1080/1463922X.2016.1154231

Federal Aviation Administration (2019) NextGen implementation plan 2018–19. https://www.faa.gov/nextgen/media/NextGen_Implementation_Plan-2018-19.pdf

Flemisch F, Abbink DA, Itoh M, Pacaux-Lemoine MP, Weßel G (2019) Joining the blunt and the pointy end of the spear: towards a common framework of joint action, human–machine cooperation, cooperative guidance and control, shared, traded and supervisory control. Cogn Technol Work 21(4):555–568. https://doi.org/10.1007/s10111-019-00576-1

Gutzwiller RS, Espinosa SH, Kenny C, Lange DS (2018) A design pattern for working agreements in human-autonomy teaming. Adv Intell Syst Comput 591:12–24. https://doi.org/10.1007/978-3-319-60591-3_2

Hoc JM, Debernard S (2002) Respective demands of task and function allocation on human-machine co-operation design: a psychological approach. Connect Sci 14(4):283–295. https://doi.org/10.1080/0954009021000068745

Hollnagel E, Woods DD (2005) Joint cognitive systems—foundations of cognitive systems engineering. CRC Press, Boca Raton, FL

Itoh M, Pacaux-Lemoine MP (2018) Trust view from the human-machine cooperation framework. Proceedings—2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018, 3213–3218. https://doi.org/10.1109/SMC.2018.00544

Jamieson GA, Skraaning G (2018) Levels of automation in human factors models for automation design: why we might consider throwing the baby out with the bathwater. J Cognit Eng Decision Making 12(1):42–49. https://doi.org/10.1177/1555343417732856

Kaber DB (2018) Issues in human-automation interaction modeling: presumptive aspects of frameworks of types and levels of automation. J Cognit Eng Decision Making 12(1):7–24. https://doi.org/10.1177/1555343417737203

Kanki BG, Helmreich RL, Anca J (2010) Crew resource management (second). Academic Press. https://doi.org/10.1201/b13238-23

Klein G, Woods DD, Bradshaw JM, Hoffman RR, Feltovich PJ (2004) Ten challenges for making automation a "team player" in joint human-agent activity. IEEE Intelligent Systems, December, 91–95

Lee JD, Seppelt BD (2012) Human factors and ergonomics in automation design. Handbook of Human Factors and Ergonomics: Fourth Edition, August 2018: 1615–1642. https://doi.org/10.1002/9781118131350.ch59

Lemoine MP, Debernard S, Crevits I, Millot P (1996) Cooperation between humans and machines: first results of an experiment with a multi-level cooperative organisation in air traffic control. Comput Support Coop Work 5(2–3):299–321. https://doi.org/10.1007/BF00133661

Lundberg J, Johansson BJE (2020) A framework for describing interaction between human operators and autonomous, automated, and manual control systems. Cogn Technol Work. https://doi.org/10.1007/s10111-020-00637-w

Lundberg J, Arvola M, Westin C, Holmlid S, Nordvall M, Josefsson B (2018) Cognitive work analysis in the conceptual design of first-of-a-kind systems–designing urban air traffic management. Behav Inf Technol 37(9):904–925. https://doi.org/10.1080/0144929X.2018.1505951

Maurino DE, Murray PS (2010) Crew resource management. In: Wise JA, Hopkin VD, Garland DJ (Eds.) Handbook of Aviation Human Factors (Second, Vol. 53, Issue 9, pp. 224–243). CRC Press Taylor & Francis Group

Pacaux-Lemoine MP, Debernard S (2002) Common work space for human-machine cooperation in air traffic control. Control Eng Pract 10(5):571–576. https://doi.org/10.1016/S0967-0661(01)00060-0

Pacaux-Lemoine MP, Flemisch F (2019) Layers of shared and cooperative control, assistance, and automation. Cogn Technol Work 21(4):579–591. https://doi.org/10.1007/s10111-018-0537-4

Riera B, Debernard S (2001) Human-machine cooperation and supervisory control of automated processes. IEEE Int Conf Emerg Technol Factory Automat ETFA 1:635–644. https://doi.org/10.1109/etfa.2001.996424

Roth EM, Bennett KB, Woods DD (1987) Human interaction with an "intelligent" machine. Int J Man Mach Stud 27(5–6):479–525. https://doi.org/10.1016/S0020-7373(87)80012-3

Roth EM, Sushereba C, Militello LG, Diiulio J, Ernst K (2019) Function allocation considerations in the era of human autonomy teaming. J Cogn Eng Decis Mak. https://doi.org/10.1177/1555343419878038

Sheridan, T. B., & Verplank, W. L. (1978). Human and Computer Control of Undersea Teleoperators.

Shively RJ, Lachter J, Koteskey R, Brandt SL (2018) Crew resource management for automated teammates (CRM-A). In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 10906 LNAI. Springer International Publishing. https://doi.org/10.1007/978-3-319-91122-9_19

SJU (2020) European ATM master plan executive view. https://doi.org/10.2829/650097

Strauch B (2018) Ironies of automation: still unresolved after all these years. IEEE Trans Human Mach Syst 48(5):419–433. https://doi.org/10.1109/THMS.2017.2732506

Trapsilawati F, Wickens CD, Qu X, Chen CH (2016) Benefits of imperfect conflict resolution advisory aids for future air traffic control. Hum Factors 58(7):1007–1019. https://doi.org/10.1177/0018720816655941

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

Vanderhaegen F (2021) Heuristic-based method for conflict discovery of shared control between humans and autonomous systems—a driving automation case study. Robot Auton Syst 146:103867. https://doi.org/10.1016/j.robot.2021.103867

Wickens C (2018) Automation stages & levels, 20 years after. J Cogn Eng Decis Mak 12(1):35–41. https://doi.org/10.1177/1555343417727438