Research Article

Jonas Pöhler*, Nadine Flegel, Tilo Mentler and Kristof Van Laerhoven*

Keeping the human in the loop: are autonomous decisions inevitable?

https://doi.org/10.1515/icom-2024-0068
Received November 29, 2024; accepted March 5, 2025;
published online April 1, 2025

Abstract: Control rooms play a crucial role in monitoring and managing safety-critical systems, such as power grids, emergency response, and transportation networks. As these systems become increasingly complex and generate more data, the role of human operators is evolving amid growing reliance on automation and autonomous decision-making. This paper explores the balance between leveraging automation for efficiency and preserving human intuition and ethical judgment, particularly in high-stakes scenarios. Through an analysis of control room trends, operator attitudes, and models of human-computer collaboration, this paper highlights the benefits and challenges of automation, including risks of deskilling, automation bias, and accountability. The paper advocates for a hybrid approach of collaborative autonomy, where humans and systems work in partnership to ensure transparency, trust, and adaptability.

Keywords: control rooms; automation; human-computer collaboration; decision-making; safety-critical environments; autonomy

1 Introduction

Control rooms are the nerve centers that underpin critical infrastructure and public safety, ranging from small-scale cockpits for 1–2 operators to large-scale settings with dozens of them. Their systems are bolstered with complex IT infrastructures manned by skilled operators, monitoring, analyzing and reacting to the inbound data. Nevertheless, with the increasing pressure on these environments, the control room demand to work efficiently and in a timely manner also increases.

Operators in control rooms are now required to sift through millions of data points, handle thousands of alerts, make split-second decisions and manage more information than ever before – all leading to cognitive strain and the risk of information overload. Simultaneously, existing technology has evolved high-potent automated tools with potential to supplement or even supplant many decision-making functions. These new technologies such as machine learning algorithms that predict system failures and pervasive computing solutions that enable real time monitoring over large areas, etc., are changing the way operational control rooms work.

This leaves one to wonder: are these technological advancements going to lead to more autonomous decision-making in the control room, and inevitably less need for human operators? Some developments like Airbus push for single pilot operations in the cockpit lead to the assumption: ”Humans need not apply”.[1]

However, has this progression of innovation also led us to the point where we have to start asking whether or not we are moving towards a future whereby decisions within control rooms will be made entirely by automated computer systems, at the potential loss of human input? Fully autonomous decision-making offers tempting benefits but it also poses significant ethical and practical challenges, especially in safety-critical domains that have traditionally required human intuition and situational awareness. Recent empirical studies have shown that under reliable automated support, joint task performance increases with the degree of automation.[5] In such cases, additional human intervention may not only be redundant but could even degrade the quality of outcomes by introducing delays or errors. This evidence suggests that as automation becomes more robust, the efficiency gains and risk mitigation it offers could render traditional human oversight less beneficial for routine tasks. At one end, the complexity of the control room tasks and instant responsive actions suggest a need for automation. The use of autonomous systems may be able to mitigate human error, produce greater operational

***Corresponding authors: Jonas Pöhler and Kristof Van Laerhoven**, University of Siegen, Siegen, Germany,
E-mail: jonas.poehler@uni-siegen.de (J. Pöhler),
kvl@eti.uni-siegen.de (K. Van Laerhoven).
https://orcid.org/0000-0002-9942-8298 (J. Pöhler)
Nadine Flegel and Tilo Mentler, Trier University of Applied Sciences, Trier, Germany

efficiencies and a quicker response during crises.[6,7] However, full autonomy comes with new dangers: mode confusion, degradation of operator skills, and separation from the experience of critical situational awareness. The future of the control room therefore must balance between human decision making and automated insights.[4,8–10]

The key question this paper addresses is whether fully autonomous decisions in control rooms are, or should be, inevitable or if systems can be designed so that human operators remain actively involved in the decision-making loop? Through review of control room technological trends, operator independence attitudes and models of human-computer collaboration, this paper argues that such a balance may be the best solution available. This style of operation could enable the control rooms to utilize the speed and reliability offered by autonomous systems while still protecting what is unique about human decision-makers, especially in high-isolation situations that require agility and tact. In the first few chapter this paper deals with the changing role of control rooms and requirements which follows automation. This paper will then examine the state of autonomy in control room work, factors impacting operators' perceptions of tech and psychology, and the barriers to automation. Drawing from experiences in different domains, it will suggest ways to integrate automation while preserving human oversight. Finally this paper will end with a discussion of the potential impact on future design of control room systems.

2 Methodology

This paper applied a structured literature review to systematically map out the changing role of automation in control rooms and its implications for human supervision, with the aim of identifying, analyzing and synthesizing relevant scholarly work, maximizing the minimization of bias and reproducibility. The review commenced with a systematic search using Google Scholar, which was selected for its broad coverage of interdisciplinary research. Search terms included "control room automation," "human-in-the-loop decision-making," "safety-critical systems," "autonomous control rooms" and "operator attitudes toward automation". Boolean operators like AND and OR narrowed the search results, for example, by linking "control room" with "automation bias" or "legal accountability". The first set of queries returned more than 1,200 publications, which were sequentially filtered for relevance to the research aims.

The search results were then screened using prespecified inclusion and exclusion criteria. The analysis was centered on peer-reviewed articles, conference proceedings and book chapters published from the years 2000 onward through October 2024, which covered technical, human-factor or ethical aspects of automation in safety-critical contexts (e.g., aviation, energy grids, emergency response). Conversely, data acquired from non-peer-reviewed sources (blogs and white papers); publications irrelevant to control room operation and automation debates; non-English texts; and duplicate studies were excluded from review. This was to guarantee that only thematically relevant and methodologically sound works informed the analysis.

The selection happened in three stages. First, title and abstract screening rejected obviously irrelevant sources like papers that deal with fields that are unrelated to control room operations. This lead to 120 documents whose full text was screened for compliance with inclusion criteria. Next, through snowball sampling approaches, the reference lists of seminal works (e.g. Endsley [9] and Sheridan [11]) were explored to identify further relevant studies. This multistage process resulted in a total of 85 publications, nevertheless covering a wide range of domains (e.g., aviation, industrial control systems) and perspectives (e.g., technical, psychological).

Three major strategies formed the analytical framework. We performed thematic synthesis to find common themes (for example deskilling, trust in automation, ethical challenges) and to cluster them into meaningful domains. Finally, a review of the identified literature mapped against, and revealed areas where not much has been covered, specifically, the potential long-term cognitive implications of reduced manual control.

Despite its methodical rigor, this review was not without its limitations. Relying on Google Scholar as the principal database may omit some specialized repositories, and limiting the search to English-language publications could potentially miss crucial studies and innovations in certain regions.

3 The evolving role of control rooms and it's increasing complexity

Control rooms have seen a dramatic evolution over the last few decades. As the systems, they manage, have grown larger and more complex, so too has the cognitive burden on human operators.[12–14]

Historically, control room operators have been a bridge between complex physical systems and high-level decision-makers who relied on them. They had to monitor hundreds of metrics, analyze trends and data and then take action to keep the system running smoothly and safely. Yet, electronic control systems, ubiquitous computing and a multitude of data sources have profoundly broadened the horizon – and particularly expanded the scale of what a control room operation looks like.

As one source points out, ”The subsequent twin effects of increasing geographic cover, and economies of scale, has raised the stakes for getting control rooms right the first time”.[15] Control rooms now monitor an interconnected networks stretching across country or even continent, facing a dramatically higher volume and velocity of information flowing in to them – while this is a small nugget of information among many, and combined with pressure to respond quickly when something new arises, this creates quite the cognitive burden on human operators.

Control and command rooms are key in energy management, emergency response, and transportation as places of real-time decision-making and monitoring/management of complex systems. Control rooms serve as the connective tissue of critical infrastructure, requiring operators to constantly evaluate information, react to alarms, and make consequential decisions that affect safety or operational efficiency. The job of a control room today is changing fast – driven by technology, but also emerging demands from a rising population and increased network complexity.[16,17]

Control rooms, as it has traditionally been configured, are built around a multi-screen workstation, where operators work with mouse and keyboard to control and interact with processes through established user interfaces.[12,18,19] This approach focuses on the usability and workflow required by operators with a design that puts humans at the center of decision-making, where all key decisions are made by operators. This model has kept operators engaged for decades, being able to use their expertise in novel situations and provide an element of flexibility needed in uncertain conditions. It is also viewed as an advantage in dealing with complex, ambiguous and unanticipated scenarios where technological solutions may not suffice.

As the number of trouble-potential situations rises, so does the reliance on a human-in-the-loop approach for control room operators; but the limitations of this approach are becoming evident. As the number of variables to be considered by operators increases – especially in environments where immediate and multisystem coordination is needed, such as in military or public safety settings – control rooms that depend on human oversight may not scale properly. That is where autonomous machines taking over could outperform human operators – faster, more consistent, and with greater computational power. Over the years, control rooms have evolved from basic monitoring stations to active, information-rich spaces where operators perform more tasks, analyze live data and process a larger volume of alerts. There are many reasons for this growing complexity:

Control rooms are now interfaced with multiple data sources, ranging from environmental sensors to connected subsystems. This immense amount of data needs constant filtering, prioritizing and interpretation.[20–23]

Control room operations have become broader in geographic range and scale, with standalone facilities responsible for distributed infrastructure and networks.[17,24–27]

With the urbanisation and expansion of critical infrastructure, control rooms are being confronted with a growing number of incidents. With more events, the requirement for faster decision-making can try operators’ mental acuity and concentration. All this means that control room operators are experiencing greater cognitive loads than they have ever faced before. In an environment where the consequences of inaction or prolonged inactivity could be catastrophic, human error is never far from the surface and always worse under such pressure. Unsurprisingly, overstressed human operators have been shown to be less accurate in processing information, with consequent degradation of situational awareness and decision making quality.

With an ever-increasing number of responsibilities and cognitive loads placed on control room operators, manual control becomes increasingly difficult to sustain for both efficiency and accuracy. Research suggests that as task complexity increases, operators can be overwhelmed and may respond too late or make mistakes, resulting in decreased situational awareness.[34]

High-stakes environments like these highlight the way that many issues necessitate automated or semi-automated technological solutions to help operate some functions themselves, or at least a reduced mental load for the operator.

The burgeoning complexity has resulted in greater interest in automation and autonomous systems as a way to lessen the load. Autonomous systems might be able to help operators by filtering information, flagging important alerts, or even autonomously making certain simple decisions, freeing human operator efforts for more complex or less-well-defined work.[4,9,35] With automation, that may allow for real-time response in situations, where a human simply cannot do as much once the situation escalates past a certain level of stress. Now, given this transition, it also raises the question of the balance of control and the extent to which human intuition and expertise are important in safety-critical environments.

4 Decision-making in safety-critical environments

The process of making decisions in control rooms is a complicated task that requires real-time data evaluation, situational awareness and critical thought, often under time constraints. Especially in safety-critical scenarios (e.g., energy grids, emergency response traffic management), these decisions are rightfully linked with public safety and operational integrity. This chapter describes decision-making – what it is, how it works in control rooms, and why it may change with new automation.

Decision-making in the control room is seldom that simple. However, operators have many competing interests to juggle like decisions under time pressure, uncertainty and ambiguity, or coordinating with multiple actors:

- Decisions evolve within a high-stakes, time-critical context; demanding innate urgency and complexity. When multiple, interconnected systems are involved – there are many potential data sources; and the potential for your actions to cascade through other systems – the complexity manifold multiplies.
- Uncertainty and Ambiguity: Control room environment is complex and has multiple variables, and incomplete or uncertain information.
- Multiple Actors: Decision making in control rooms is not an individual process; it typically involves one or more operators, and sometimes between organizations (e.g. coordination between fire departments and energy providers). Transitioning into this collaborative space involves both communicating across roles and departments as well as aligning on priorities.
- Data-Driven Grounding: Operators take long care in basing decisions on data from monitoring systems, models, and simulations. The decision foundation contains real-time data as well as predictive models sufficiently accurate that protocols or rules can mediate action when the conditions of specific scenarios have been met.

Such characteristics highlight that control room decisions are evolving and influenced by multiple interdependent variables, which make them unique (compared to routine, low-stakes decision-making scenarios) as these belong to high-intensity situations.[36,37]

In safety-critical areas, decisions can be classified into different types depending on what type of decision it is and the degree of responsibility. Some possible classifications are:

- Strategic and Tactical Decisions: Strategic decisions are those that are long-range in nature like policy decisions, whereas tactical decisions are more short-term or operational. In a control room situation, for instance, tactical decisions are made on the spot by an operator to react and adapt to evolving situations that occur in real time, while strategic decisions take place at higher levels.[38,39]
- Routine Decisions vs Non-Routine Decisions: Routine decisions are those which either all of us make regularly without straying too far from standard protocols and procedures, whereas non-routine operations are more creative and innovative in nature and require consideration with unique context in mind.[37,40]
- Roles of Responsibility and Accountability: Decision making in control rooms comes with accountability. While operators make the decision and therefore own their decisions, responsibility for that action may be held at supervisory or organizational levels. For example, an operator may own a particular incident but ownership of the effect on the organization may not be at their level. Such a spread of responsibility requires clarity on roles and decision-making in hierarchies.[41]

By the integration of automation, this situation is complicated by accountability issues where some decisions are taken independently. When systems work independently or even automatically, it's essential to establish who is responsible when an error occurs, or worst will happen. The speed of technological growth outpacing legal frameworks, along with the question about who is legally accountable for decisions made autonomously – that being the technology or human user(s), also raises ethical and operational dilemmas.[42–44]

Decision-making is profoundly influenced by several organizational factors, like who has an expertise and power to make what access control decision. In cases where quick responses are essential, clearly defined roles and escalation processes will make it a lot easier to decide who drives the bus.

Guidance on procedures (i.e. standard operating procedures or SOPs) can provide a scaffolding for decision-making by describing trajectories of steps and best practices taken in response to common scenarios. Many of these guidelines are especially important in safety-critical situations where escalating events can be disruptive and even dangerous.[45,46]

By automating data processing, scenario analysis and recommendation generation, the decision-making process can also be improved. But as we approach a future of automating decisions, it will impair effective automation in situations that require more subtle judgement which can adapt to the situation at hand.[47–49]

Automation can further alleviate cognitive load on the operators as well. Automated systems allow operators to concentrate on higher-level, knowledge-based behavior [50] by taking over mundane tasks and presenting only the most pertinent information – increasing both efficiency and accuracy in a high-stakes environment. However, for automation to be viable the decision-making must be transparent. To be able to trust these systems, the operators should comprehend how these automated recommendations are produced. Openness is necessary so that operators trust the outputs and understand how to read what came out of the system.[51–53]

All in all, the need for greater control, an increased workload and a rising volume of data in Control Room environments really create situations which challenge even people to manage every part of decision-making.

Automation plays an essential role by getting rid of the unreliability associated with human decision-making, but poses its own unique challenges and dangers when introducing AI into safety-critical parts of decision-making. One of the big challenges to address is trust and accountability. During critical situations, automated systems must be trusted by the operators. It poses the risk of automation bias, because there is a chance that operators will start to rely on automated recommendations and not question what outputs the system delivers (miss an error). However, absence of trust may make people reluctant to use automated support.

The other challenge relates to the decisions themselves. Not every decision made in control rooms can be distilled to objective data. Most of the decisions are subjective and machine does not have human intuition required to replicate it as experience comes with time. These subjective, non-reproducible nature of some decisions suggests that automation can be helpful only in cases where tasks are objective and should be performed by humans if they involve complex decision-making or judgement.

Legal and ethical considerations further complicate the use of automation in decision-making. As mentioned, legal frameworks often lag behind the rapid development of autonomous systems, creating ambiguities in accountability. When automated decisions lead to adverse outcomes, determining liability can be complex, especially in scenarios where operators only supervise automated processes rather than making direct choices.

5 Human affinity for technology and operator attitudes

While the technical feasibility of autonomous systems will always be a key component in their eventual success, a large factor surrounding automation use in control rooms is actually the attitudes and preferences of operators themselves. The adoption of automated tools in control room environments is largely influenced by operators: their enthusiasm for technology and autonomy as well as the trust they have built with these tools. This chapter focuses on the human behavioral and psychological characteristics of operators interacting with technology; it investigates how operator attitudes influence the acceptance and use of automation in safety-critical environments.

5.1 Technology interaction preference

Affinity towards technology interaction (ATI) is defined as disposition to interact with technology.[54] Operators present in control rooms typically have high ATI as their job involves constant interaction with complex IT systems and data visualization tools.[20] But both research and anecdotal evidence show affinity for various technologies is not the same across domains, nor even influenced by particular ages, incomes or explicit experience with that tech, but still part of some operational need. As an example, operators on emergency services and public utilities may be more technophilic than their counterparts in maritime control, which can potentially be explained through the structural congruence of tasks within these fields alongside how environmental conditions shape new technology interaction.[20]

High ATI operators tend to be more open to technological solutions and may be less resistant in making the switch to sophisticated decision-support systems and automation tools. These operators are more open to autonomous capabilities that automate clear tasks and enable them to focus on higher-level strategic decision-making. Operators with lower ATI, on the other hand, may hold a more skeptical or resistant stance towards automation – especially when it changes existing workflows significantly or requires substantial behavioral change. Insights into such diversity will help in the design of automation that meets the needs of varying human user types within control rooms.[55,56]

5.2 Need for independence and authority

Equally, the longing for independence and command is a paramount reason behind operator behavior in favorable conditions regarding automation. Decision-making: Control room operators work in high-stakes environments where maintaining a degree of controllability over decision-making is desirable; they want to be the one deciding actions when the result influences public safety and operational integrity. That autonomy matters to operators, for it allows them incorporate their experience in lesser-defined scenarios while making situationally aware and nuanced decisions.[9,57]

But the growing complexity of control room tasks and the demand for quick responses are driving many operators to the realization that semi-autonomous or autonomous systems can help. Operators are often more comfortable delegating routine or well-defined tasks to automation because they want to hold authority over the higher-level, high-impact decisions that can be very complex. This selective stance toward autonomy also demonstrates a tendency toward collaborative automation, which perceives systems as partners to aid or assist rather than replace human judgment.[58,59]

Whether operators prefer to have control also affects the levels of trust they put in automation. Operators are more trusting and willing to use systems when these system designs enhance, rather than supersede, their decisions. On the opposite side, fully autonomous systems that take control away from operators can be anxiety-provoking and mistrusted; users perceive their knowledge and judgment as devalued. This points to the necessity of designing automation that works with operators in a way that supports their perceived control instead of undermining it.[10]

Trust is key to successful automation in control rooms operators need to trust that automated systems will perform as expected, respond with accurate information and assist them in achieving their objectives. In safety-critical environments, trust is even more important as they need to have confidence in the fact that systems will not fail nor produce some unintended gray swans under high impact scenarios.[60–62]

And this is where the transparency has a major role. A good understanding of the automation process and how it arrives at decisions builds operator trust, especially when faced with novel situations. Systems that provide explainable recommendations and transparency in rationale for automated actions will garner acceptance by operators. For example, systems that act as a black box – producing an answer but without the explanation for why its said answer is correct – breeds distrust in an operator who might not want to provide that guidance at a moment of critical need.[52,63]

The transparency and interpretability of automation systems would also reduce the likelihood of operators engaging in automation bias (i.e., an over-trust on what these systems output). This understanding helps operators to critically assess the tool's recommendations and provide interference when appropriate, which in turn improves safety and contributes to effective operation.

5.3 General attitudes towards automation in specific domains

Depending on the nature of the work, environmental considerations as well as operational constraints, automation is often regarded in very different ways across control room domains. Additionally, in some domains where rapid decision making has to be done under high pressure situation as in emergency response control rooms, the operators might find automation which help filtering information and prioritization of alerts more acceptable. The reason for this acceptance is to remove mental burden and speed up answers in time-sensitive situations.[20]

In contrast, operators in public utility control rooms – including the folks controlling our energy grids – may have a more ambivalent view about automation. They will embrace autonomous tools for routine monitoring and diagnostics of systems, but expect them to be prevented from making massive changes without human supervision due to widespread effects. Likewise, in the highly variable environment of maritime control rooms, operators tend to not use automation due to uncertainty over capability and robustness in unstructured environments.[20,64]

Such domain-specific attitudes further demonstrate that automation design requires a construction method. Automation systems must be designed for the types of functionality that each control room domain truly needs, matching automation with human oversight to conditions and operator expectations.

5.4 Collaborated automation: assistance vs control?

As a rule, operators prefer a middle-ground approach to automation – systems that assist but do not take full control. This collaborative automation model emphasizes human-automated system partnership – combining strengths to achieve effectiveness. In this model, automation takes care of mundane, data-heavy tasks while operators maintain control of complex, high-stakes decisions that demand judgment and agility.[4,11,65]

Collaborative automation also enhances operator engagement by allowing them to focus on tasks that align with their expertise and reduce the cognitive burden of low-level monitoring. By presenting automation as a tool that augments human capabilities rather than replaces them, control rooms can foster an environment where operators feel valued and empowered. This approach also encourages operators to trust and rely on automation, as they perceive it as an ally rather than a competitor.

6 Autonomy and human oversight

Automation can benefit the cognitive load, lead to quicker responses and process repetitive tasks but should never fully replace the human at centre stage in making decisions, particularly for safety-critical ones. In this chapter we will touch on the significance of balancing human autonomy with oversight, dangers of becoming overly reliant on automation, and possibilities to create a collaborative system that respects both technology and human skill.

Especially in safety-critical environments such as control rooms, human decision-making adds vital flexibility and adaptability into the mix, alongside ethical aspects that may go beyond what an autonomous system can provide. These trade-offs between humans and autonomous systems may lead to synergy; for example, autonomous systems are quick and efficient with large amounts of data – but human operators bring contextual understanding, judgment, and the ability to respond to novel situations, abilities that will be vital for unpredictable high-stakes scenarios.

Finally, human oversight is a must: Automated systems are often not equipped to handle unexpected scenarios or variables that might be more nuanced and therefore may not fall into predetermined parameters. Abstract – For instance, an autonomous system in the control room for a power network would be capable of identifying equipment faults as well as trigger steps to mitigate them based on known guidelines. However, it would likely not be able to make those complex decisions when resource allocation must be made during a regional outage in circumstances of human interactions needed with real-time situational awareness required to quantify trade-offs and prioritize actions.

Additionally, the presence of human oversight enables accountability in terms of ethics and law. Where the consequences of poor judgment could have life or death ramifications, or widespread societal impact (think technologies affecting many people), there should be a human in the loop to ensure that action is ethical and that we can point fingers when things go wrong. While autonomous systems can based on data make decisions without human intervention, they cannot apply the moral and ethical reasoning that often is required in a critical situations.[66]

Automation can improve efficiency, but excessive dependence on autonomous systems comes with a number of dangers. A risk here is automation bias where operators are likely to be overly reliant on automated recommendation and won't interrogate system outputs. This bias may foster some level of acceptance in where the operators will not critically engage with the output and accept whatever comes from the system. Under a safety-critical domain, if an incorrect recommendation cannot be challenged based on automation bias, the results can be devastating.

A second risk, called deskilling, is gradually losing important flap operator skills over time. With growing automation in performing routine tasks, it may happen that operators would not have an idea of how to perform manual procedures and they might be compromised on their ability to intervene during a failure. But in emergencies, if an autonomous system faces a new problem, operators may have to take over control using their manual skills as they are weakened through underuse. In the long term, if human skills are not retained, it may be challenging and costly to shift back towards a more human-led approach instead of an unsustainable stat that is solely dependent on automation – a situation commonly referred to as deskilling.[67]

Moreover, autonomous systems induce confusion in modes if operators do not have a clear view of the current status, mode and level of automation of the system. Such confusion can potentially cause mistakes if operators misassume the extent of control they have or are unaware of a system operating autonomously. Further, if mode confusion does occur, it may make the division of responsibility between human-run and autonomous systems less clear than in other scenarios – especially problematic in critical moments when a quick response is needed.[68]

Due to the difficulties and dangers of achieving complete autonomy a partially automated solution is therefore an attractive compromise, often referred to as "collaborative autonomy". Automation plays the role of an assistant to human operators, performing low-level data-intensive tasks, but leaving decisions that would be considered critical in human hands (the augmenting path). With collaborative autonomy, control rooms can reap the efficiencies of automation while keeping human operators – who provide crucial advantage to high-stakes environments – front and centre. Metnler et al. used the metaphor of a shepherding dog within a flock to describe an ideal human-AI collaborative system: the autonomous system identifies opportunities and threats, provides warnings and suggestions, but the human maintains the final say.[4,69]

Human-centered autonomous systems are collaborative, pursues the goal of supporting and enhancing human capabilities instead of replacing them. As an example, a collaborative system can observe various parameters of the system and identify potential problems, but it would defer to the human operator for making a decision. It also allows for shared control between the system and the human operator, while allowing human intervention at any point if complex situations require subtle judgement.

From a practical standpoint, adaptive automation is one manner in which the principle of collaborative autonomy can be realized, where the level of automation varies dynamically as a function of situational requirements. The system may take over more when in a non-safety critical, repetitive setting and use less supervision for rudimentary tasks. In high-stress or uncertain situations, though, the system would ask the operator to intervene, effectively handing off control back to the human operator for context-sensitive decision-making. An adaptive model that caters to the automation levels according to the type of task and situation, ensuring human involvement where needed.[70,71]

Finally, training and skill retention are vital to ensure that operators remain proficient even as automation increases. To mitigate the risks of deskilling, regular training programs should be implemented to keep operators proficient in both manual and automated procedures. Simulation-based training can help operators maintain critical skills by providing practice scenarios that require manual intervention. Training programs should also address the specific challenges of working with automated systems, such as managing automation bias and responding to mode confusion.

Balancing autonomy and human oversight also requires careful consideration of ethical and legal implications. In safety-critical environments, autonomous systems must adhere to ethical principles that prioritize human well-being and safety. This adherence is particularly important in decisions involving potential risks to human life, where ethical reasoning is essential for making choices that align with societal values.[66]

Regulatory frameworks and legal oversight are also crucial to ensure that autonomous systems are designed and deployed responsibly. For this regulations like EU 2024/1689 and national laws must be carefully considered.[72,73]

Legally, accountability in automated decision-making must be clearly defined. When control room systems operate autonomously, it can be difficult to assign responsibility in cases of error or failure. Legal frameworks often lag behind technological advancements, which may leave gaps in accountability. To address these issues, organizations may need to establish policies that define responsibility for autonomous actions, clarify when operators are expected to intervene, and specify protocols for monitoring autonomous systems to ensure compliance with ethical and regulatory standards.

7 Practical implications

Although the particular incorporation of automation into control rooms might pose a challenge distinctive to the individual situation, comparable issues are present in any field where coordination between human operators and autonomous systems is essential.

This chapter examines two case studies – aircraft cockpits and industrial control rooms – that illustrate the practical considerations of managing pilots and automation in small-scale and large-scale aeronautical socio-technical systems.

The differences we see here highlight the diversity of what is needed/expected – and challenges for automation to handle/control room – which in turn provide valuable lessons for determining best practices on where autonomy ends and human staff should take over.

7.1 Cockpits of aircrafts – a miniature control room

By definition, a “cockpit”, can be seen as a control room – and therefore an aircraft cockpit might possibly be the most extreme case of control room with highly specialized tasks in very limited space. Pilots in this environment operate adjacent to their automated systems, with little physical separation from the flight controls that command complex and safety-critical flight operations. Cockpit automation, including autopilot systems and automated flight management systems (FMS), has improved significantly over the last several decades providing efficient and secure handling of routine activities. On the other hand, this huge development has posed a unique set of challenges around human oversight, situational awareness and trust in these automated systems.

Automation is more prevalent in a modern cockpit for routine tasks that have less significance while this brings strategic decision-making into focus by the pilots. Current autopilot systems can control altitude, speed and heading largely without input from the pilot, while the FMS can automatically determine fuel requirements, routes and ETA. By alleviating some of the cognitive stress of standard flights, these systems let pilots concentrate on monitoring how well the aircraft is performing and addressing irregularities. Yet, as we have become more automated in our procedures this has led to unintended consequences (e.g. ”Automation Complacency” where pilots are so reliant on the automation they lose situational awareness).[74–76]

As the 2009 crash of Air France Flight 447 illustrated, there can be dangers from excessive dependence on cockpit automation; after erroneous airspeed readings caused the autopilot to turn off in this case, the pilots lost control of their aircraft and crashed. For example, in this case the abrupt transition became a phrasing for poor pilot training and lack of situational awareness. Many lives were lost that day, as the pilots were not ready to take control under those conditions. The lesson: in uncertain environments like those faced by cockpit crews, where situations can rapidly go from bad to worse, automation must never be allowed to overshadow human intervention.[77,78]

Another example for overshadowing the human operator with automated operations is the Crash of Lion-Air Flight 610 and Ethiopian-Airlines Flight 302 in these cases the automated MCAS system encountered an unexpected state and even though the pilots tried to override the system they could not gain control quick enough.[79]

In response to these problems, the aviation industry has implemented some best practices – including improved pilot training and simulation on manual control proficiency and automation awareness. Current training emphasizes the need for automation awareness, practice in responding to abrupt changes in balance control modes, and recognition of states where pilots may be required to take manual control. Cockpit system design also continues trends of increased transparency and intuitive displays to enhance pilots' situational awareness by providing information on the level of automation currently in use, its status, and potential threats.[7,80]

Current trends in aircraft control aim to reduce the workload of pilots to the point where it may be feasible to operate an aircraft with a single pilot. Airbus' push for single-pilot operations underscores the extent to which automation has already assumed a dominant role in aircraft control systems.[81–83]

7.2 Industrial control rooms – large-scale automation

Unlike the small, extremely focused cockpit environment, industrial control rooms are large-scale, often complex environments where operators monitor extensive interrelated systems across whole entities like power plants, oil refineries and manufacturing facilities. It's common that such control rooms have many monitors, a lot of data input and lots of controls, allowing operators to manage processes hundreds or thousands of miles away with terabytes (if not more) of real-time data. Industrial control rooms are fundamental in handling these large-scale operations, usually taking care of routine monitoring, fault detection and process adjustments through a utomation.[84]

In an industrial control room, automation is used extensively to facilitate the handling of everyday work and ensure stable operation. An example would be power generation plants where automation systems monitor the electricity requirements, balance the load and identify faults. Such automated processes free up operators for high-level supervision, instead of being tied to basic system parameters. While this automation provides a lot of benefits, it also poses some challenges in the case which are unexpected. If automated systems encounter scenarios they cannot handle, operators must be able to take over easily – that's not easy to do if they've been away from manual control for long periods of time.[6,85]

An iconic example of the perils of mass automation in industrial control rooms is the explosion at the Texas City Refinery in 2005. Subsequent investigations into the incident found that it was exacerbated by the failure of automated systems to warn operators about unsafe conditions in the facility, leading to a catastrophic failure with multiple fatalities and enormous environmental damage. This incident also demonstrated the dangers of overdependence on automated systems that are not augmented by sufficient human perception: The operators did not have a good enough picture of the situation to manage a growing threat appropriately. Following the incident there were widespread demands for better training and better alarms, as well as stricter safety protocols.[86,87] In the aftermath of events similar to Texas City, a focus within the industry has been on automation and human control room interaction design in industrial environments. Such initiatives include an overhaul of alarm models that ensures the avoidance of overload while filtering out non-critical alerts so that operators can get a glimpse of real threats. Training programs have adjusted to encourage the need for manual skills and knowledge on how to read the alerts generated by automated systems and what action to take when they appear. Furthermore, there is exploration into adaptive automation strategies where the control transfers dynamically between human operators and systems based on the operational context to enhance human-system collaboration while exploiting strengths of automation.

8 Are autonomous decisions inevitable?

With a positive development in automation technology, and mounting pressures on control rooms, many wonder: Will autonomous decision-making become unavoidable in safety-critical environments? Across various industries, control rooms are increasingly implementing advanced levels of automated systems that can integrate numerous complex data processing and fault detection as well as some decision-making capabilities. However, although the trend seems inevitable that autonomous systems will take over more and more standard processes, there are fundamental legal, ethical and operational arguments for believing that human intervention will still be necessary in many cases.[88]

8.1 The argument for full autonomy

Over the long term, human operators in control rooms are more likely to transition from active decision making to watching and overseeing autonomous systems. Many of the tasks we carry out regularly, especially those that contain repetitive watching and analysis or predictable decision points, may be automated. Autonomous systems can analyze big data as it emerges, perform pattern recognition and deploy pre-defined neural networks to make immediate decisions in relatively simple situations. This transition is especially valuable in areas with high cognitive load and multi-tasking which can lead to decreased human performance, as well as higher error rates.[4,89]

For many of the more mundane activities and tasks in various sectors such as manufacturing as well as energy management, automation has already handled these functions with human operators now taking on a supervisory role. Automation unburdens operators from repetitive tasks and allows him/her to focus on higher-level duties like – analysis of complex situations, handling unforeseen events, and making strategic decisions. Essentially, human operators become automation managers, managing the processes to ensure they comply with safety and objectives.

This shift to being more of a monitor is most useful in contexts where you can draw on routine work, and the consequences of error can be reduced through preplanned modalities. With adaptive automation getting stronger, systems will also get better in knowing when the need for human intervention arises, thereby lessening the dependence on regular human involvement in routine tasks. The monitoring role does not mean we can abandon human interaction entirely, it means that we capture the nature of human engagement by oversight and intervention rather than control.

While legal, ethical, and operational challenges may constrain the extent of full autonomy in control room decision-making, a growing body of research provides compelling arguments for increasing reliance on automation. First, empirical studies have demonstrated that higher degrees of automation significantly enhance routine performance and reduce operator workload.[5,90] Under reliable support, joint task performance increases with the degree of automation, as autonomous systems efficiently handle data filtering and well-defined tasks, minimizing human error in safety-sensitive environments.

Complementing these efficiency gains is the evolving nature of human–automation interaction. As automated systems consistently deliver reliable outcomes, operators tend to develop increased trust and dependency on them. This phenomenon, often referred to as ”automation bias”, manifests as reduced manual oversight – even when human supervision remains a necessary safeguard.[91,92] Notably, when automation reaches a high level of reliability and performance, further human intervention may actually lower the overall quality of outcomes. In such cases, operators' attempts to override or second-guess the automation can introduce delays, errors, or disruptions that ultimately degrade system performance.

Moreover, the risk mitigation capabilities of autonomous systems further bolster the case for increased automation. In high-stakes contexts – where human error can have catastrophic consequences – autonomous decision-making ensures consistent adherence to optimal protocols. Although explainable AI approaches aim to maintain transparency and allow for human intervention when needed, the overall evidence suggests that minimizing unnecessary human interference in routine processes substantially reduces the risk of error [5] (Figure 1).

The traditional Fitts List [93,94] delineates human strengths in sensory functions, perceptual abilities, flexibility, judgment, and reasoning. However, recent

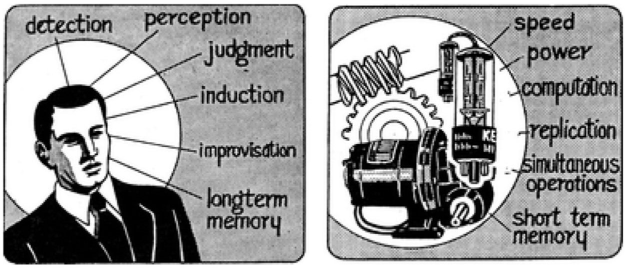

Figure 1: Fitts list consists of tasks where the human is best at (on the left) and where the machine is best at (on the right).[93]

advances in artificial intelligence necessitate a critical reappraisal of this framework. Modern technologies not only rival but, in many cases, exceed human performance in these domains. Lets look at the different claims:

Sensory Functions: Historically, humans have been celebrated for their acute sensory thresholds. Yet, modern sensors - such as LiDAR and hyperspectral imaging – detect phenomena far beyond human capabilities. These machine "eyes" not only see more, but they react faster – processing images in milliseconds, which is impossible for human vision's 200 ms response limit. In image recognition benchmarks, AI vision has surpassed humans: by 2015, deep neural networks achieved lower error rates in object recognition than people, marking the first time machines beat humans at classifying diverse images.[95] In fields like medical diagnostics, computer systems analyze MRI and CT scans to identify anomalies that may elude human radiologists.[96–98]

Perceptual Abilities: Humans excel at generalizing sensory input under varying conditions, a strength once thought unique. Today, deep learning networks have transformed pattern recognition. Advanced facial recognition systems operate with over 99 % accuracy across diverse scenarios, and computer vision applications in quality control and autonomous navigation demonstrate a level of consistency and scalability that challenges human perceptual reliability. On the ImageNet challenge (a broad test of visual object recognition), AI surpassed human performance: a 2015 deep network by He et al. achieved only 4.5 % error, slightly better than the 5 % error of expert humans.[95] In fields like medical imaging, this pattern-recognition prowess has tangible impact. A deep learning system at Stanford (CheXNeXt) was able to screen chest X-rays for 14 types of disease in seconds, performing as well as radiologists on most conditions – and even outperforming human experts on one pathology.[99]

Flexibility: While human adaptability in novel situations has long been admired, machines have made significant strides in this area. Reinforcement learning algorithms – exemplified by systems like AlphaZero – develop strategies through self-play, often surpassing human ingenuity in complex games.[100] Humans typically specialize (the best chess grandmaster isn't an elite Go player), but AlphaZero showed an algorithm could flexibly apply itself to multiple complex domains and achieve superhuman skill in each.

Judgment and Selective Recall: Human judgment has traditionally relied on the selective recall of experiential knowledge. However machines have shown times and times again that they are more and more able to learn from a large troth of knowledge. Recent advancements in LLM technology shown that selective recall is a discipline where machines have eclipsed the human. Making judgement calls based on the available information has also been shown to be above human performance. One study showed an AI outperforming the traditional Breast Cancer Risk Assessment Tool in predicting which women would develop breast cancer within 5 years.[101]

Reasoning: The human capacity for inductive reasoning has been a cornerstone of scientific inquiry. Yet, machines have begun to master both inductive and abductive reasoning. Systems such as AlphaFold accelerate scientific discovery by predicting protein structures, while large language models generate creative hypotheses and solutions across a range of applications – from experiment design to complex code generation. In one demonstration, GPT-4 scored around the top 10 % of test-takers on a simulated bar exam, which includes complex legal reasoning and essay writing, exceeding the performance of most law graduates.[102]

This transition is also reflected in the changing role of human operators. As automation takes over more routine functions, human operators are increasingly redefined as supervisors and strategic decision-makers. This evolution preserves critical human expertise for non-routine and complex scenarios, ensuring that human judgment remains central when it matters most.[103] Finally, technological progress and economic incentives drive the irreversible adoption of higher automation. Once autonomous systems exceed human capabilities in specific domains, cost savings and reduced liability risks further accelerate their integration.[5] Additionally, studies in human–robot interaction indicate growing social acceptance of autonomous agents, suggesting that both operators and the public are increasingly comfortable with systems that assume greater control over routine tasks.[104] Collectively, these factors converge to argue that increased automation in control rooms is not only technologically feasible but also strategically advantageous. The combined benefits of enhanced efficiency, effective risk mitigation, and a redefined supervisory human role make the shift toward more autonomous decision-making a natural and necessary evolution in managing complex, safety-critical systems.

8.2 Legal and ethical limitations to complete autonomy

While the advantages of automation are evident, there will be deep-rooted legal and ethical safeguards against machines taking over all aspects of control room decision-making. Safey-critical environments, e.g., healthcare or aviation, often focus on ethically and socially consequential decisions. By these terms, the stakes are astronomical – a bad call could mean someone dies, an ecosystem remains irrevocably fractured, or economies collapse. These types of situations call for moral reasoning and ethical judgments that autonomous systems, which operate blindly according to algorithms or pre-set directions, cannot provide.[3,105,106]

Legally, it is hard to find fault when an autonomous system goes awry.

Constraints on the ability of machines to operate independently are also established by legal frameworks, especially when decisions influence human lives. Liability is a big deal when it comes to this. Determining liability becomes complicated when people are harmed due to a choice made by an autonomous system. As legal accountability currently falls mostly on human operators or organizations, the decision chain from machine to human must remain intact in order for humans to take responsibility for outcomes. In high-stakes settings, the legal and ethical issues with fully autonomous systems cannot be overcome without clear accountability.[3,4,106–109]

In addition, in complex real-world situations, unexpected edge cases or novel scenarios that the programming has not been previously exposed to may create problems for autonomous systems.

In addition, ethical perspectives highlight the need for human oversight in areas where compassion, empathy, or value-based decisions are needed – things machines simply cannot do. Although an autonomous system might perform well when rules and the environment are highly structured, in complex moral dilemmas or where values conflict the autonomous system is unlikely to do a good job as it lacks context. So, for example, in a control room for emergency response, the decision to allocate scarce resources in relation to an ongoing crisis cannot be purely evidence based because that will inevitably involve balancing competing priorities and trade-offs which embody social values and ethical principles. Those types of decisions are human, and should never be fully delegated to a machine.[110,111]

While autonomous systems are taking over most of the routine tasks, human operators will always be imposing their ways in high-stakes decision-making scenarios, where flexibility, critical thinking, and adaptive choices should be exercised. The nature of autonomous systems, which operate with fixed constraints and set algorithms, means that they struggle to cope in situations where the right action or sequence of actions is not sufficiently pre-established, and they may fail to identify the factors driving a scenario.

By contrast, human operators have the cognitive flexibility to evaluate unusual situations, interpret incomplete data, and rapidly adapt in ways that machines can't. When a high-pressure situation arises, or system failure occurs, sometimes it requires human voices to override the automated impulse reaction and instead construct a solution through thinking outside of the box, or action across multiple systems and organizations. Our ability to adapt is especially important in any kind of control room scenario where events can be unpredictable and responses cannot always be mapped out ahead of time.Human operators will be involved primarily in exception handling: that is, where the automation exceeds its functional limits or happens upon an unanticipated situation. And, they may even serve as a crucial successful hand-in-hand validator for validating organization-wide critical machine-made choices concerning ethical, legal, and operational aspects. Human operators are still there for a safety net, making sure that automation does not operate in a vacuum but under human judgment and control as they maintain the final authority in critical situations.[88,112,113]

However, with control rooms also being increasingly furnished with autonomous systems, there are significant implications for design and training. Firstly, control rooms need to be a hybrid between the human operator and autonomous systems, yet still have humans engaging in the loop and capable of intervening when necessary. These systems call for interfaces that convey the state, intent, and constraints of the system in a clear manner so that operators know what is going on and when they need to intervene as automation reaches its limits. Providing visual or sound indication if the user is going to interfere or need an operation with more detail. This will not only make your operator responsible but also keep him on a lookout when there are automated actions being executed.

Thirdly, training programs need to adjust in order that operators will be trained for their new tasks. Operators will require training, not for manual interventions but for monitoring skills and critical assessment of automated outputs with an easy intervention method when there is a need to handle exceptions. Operators can benefit from simulation-based training that gives them practice by experiencing many scenarios in which they will have to deal with the automatic-to-manual control handoff; In addition, it helps reduce deskilling among operators because they remain familiar with the underlying systems while taking on a supervisory role.[6,65]

Human factor smart systems will also include autonomous decision-making, and therefore regulatory frameworks and formalized ethical guidelines will have to adapt to the pace of technological change. In addition, explicit and concrete policies on accountability, transparency and such triggers for intervention are required to avoid that automation systematically goes beyond what we have decided as a society or through our law. Organizations need to develop clear frameworks of when operators need to intervene, as well as the roles of humans and machines in shared control situations.[9,59]

Ultimately, yes autonomous systems will help augment many aspects of control room operations in the future but humans are still key. Because accountability, ethical judgment, and situationally adaptive decision-making are important in high-stakes environments, operators will always be the final backstop – always holding authority and an ability to override automated decisions with human expertise and judgment.

9 Conclusions

While much of the discussion will focus on the opportunities, incorporating these technologies will also raise challenges as well. Automation has already been shown to increase efficiency and decrease cognitive burden for humans in environments where data needs to be processed rapidly, and decisions made every few seconds. Automation is now crucial for managing complicated frameworks inside industrial control rooms to aircraft cockpits, and keeping them from going haywire. Given the increasingly complex and demanding requirements of a modern control room, the potential areas where automation could help – faster responses, less errors, ability to handle repetitive tasks – are significant. Across many industries, human agents have already started taking on the role of a supervisor who oversees and monitors an automated process with minimal intervention. With this change in focus from verifiers to overseers, people can concentrate on higher level tasks related to oversight and strategy while autonomous systems deal with daily operations. So will this trend continue and the human niche become smaller and smaller until the machine controls everything? Maybe. The technical possibility for a level of automation is there. Operators are open to more automation and more technology in their work environment. In some cases it seems that only the missing legal framework stands in the way of full autonomous decisions. However there are also compelling legal, ethical and operational reasons why full autonomy is unlikely to happen soon. Even if they reach higher efficiency, autonomous systems lack the moral reasoning and flexibility required to confront challenging ethical problems and unknown situations. Finally, existing legal codes limit the amount of jurisdiction that machines can have – especially in liability and accountability-sensitive aspects. While control rooms of the future may have all human or full autonomy, it could be messy to assign accountability in case of failure or error with little transparency built-in mechanisms.

For the near future, these legal hurdles seem to form the control as a hybrid environment where automation will take over control of routine aspects, but humans are not likely to be removed from the loop completely, especially in safety-critical situations. This balanced partnership harnesses the strengths of both machines and humans to improve efficiency and resilience in complex systems.

Looking ahead, it will be crucial to monitor how regulatory frameworks evolve and how swiftly technology advances. This next phase of automation will determine the balance between human oversight and machine-driven processes, ensuring accountability, ethical considerations, and efficient operation. To summarise, future control rooms are not a choice between man and machine, for now, but rather a partnership bringing both strengths to the table. In the long run, if the intelligence and capabilities of the machine continues to develop at rapid speed, there may be a time when we say: Humans need not apply.

Acknowledgments: The authors acknowledge and thank all participants of the 11. Workshop Mensch-Maschine-Interaktion in sicherheitskritischen Systemen at Mensch und Computer 2024 for discussions which influenced this paper.
Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Use of Large Language Models, AI and Machine Learning Tools: Grammarly for spelling and grammar correction.
Conflict of interest: The authors state no conflict of interest.
Research funding: This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 425868829 and is part of Priority Program SPP2199 Scalable Interaction Paradigms for Pervasive Computing Environments.
Data availability: Not applicable.

References

1. CGPGrey, Humans Need Not Apply. https://www.youtube.com/watch?v=7Pq-S557XQU.
2. Verdegay, J. L.; Lamata, M. T.; Pelta, D.; Cruz, C. Artificial Intelligence and Decision Problems: The Need for an Ethical Context. *Metaverse* **2021**, *2*, 13.
3. Douer, N.; Meyer, J. The Responsibility Quantification Model of Human Interaction with Automation. *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 1044—1060.

4. Nothwang, W. D.; McCourt, M. J.; Robinson, R. M.; Burden, S. A.; Curtis, J. W. *The Human Should Be Part of the Control Loop?*; Resilience Week (RWS): Austin, USA, 2016; pp 214–220.
5. Onnasch, L.; Wickens, C. D.; Li, H.; Manzey, D. Human Performance Consequences of Stages and Levels of Automation: An Integrated Meta-Analysis. *Hum. Factors: J. Hum. Factors and Ergon. Soc.* **2014**, *56*, 476–488.
6. Lockhart, J. M.; Strub, M. H.; Hawley, J. K.; Tapia, L. A. Automation and Supervisory Control: A Perspective on Human Performance, Training, and Performance Aiding. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Thousand Oaks, Vol. 37, 1993; pp 1211–1215.
7. Sengupta, S.; Donekal, A. K.; Mathur, A. R. Automation in Modern Airplanes – A Safety and Human Factors Based Study. *INCOSE Int. Symp.* **2016**, *26*, 386–394.
8. Sprengart, S. M.; Neis, S. M.; Schiefele, J. *Role of the Human Operator in Future Commercial Reduced Crew Operations*; IEEE: New York, 2018; pp 1–10.
9. Endsley, M. R. From Here to Autonomy: Lessons Learned from Human–Automation Research, 2017. https://journals.sagepub.com/doi/10.1177/0018720816681350.
10. Ueda, S.; Nakashima, R.; Kumada, T. Influence of Levels of Automation on the Sense of Agency during Continuous Action. *Sci. Rep.* **2021**, *11*, 2436.
11. Sheridan, T. B. Individual Differences in Attributes of Trust in Automation: Measurement and Application to System Design. *Front. Psychol.* **2019**, *10*, 1117.
12. Cochran, E. L. Designing Control Rooms for the Year 2000: New Technologies, New Techniques? In *Proceedings of the Human Factors Society Annual Meeting*; SAGE Publications: Thousand Oaks, Vol. 36, 1992; pp 458–459.
13. Partini, J. Welcome to the Future Control Room Working Environment, 2023. https://www.slideshare.net/JetonPartini/welcome-to-the-future-control-room-working-environment.
14. The Changing Nature of Control Centre Ergonomics. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Thousand Oaks, Vol. 44, 2000; p 506.
15. Wood, J. A Standard Approach to Control Room Design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. *44*, 2000; pp 6–417. 6–420.
16. Meshkati, N. Control Rooms' Design in Industrial Facilities. *Hum. Factors and Ergon. Manuf. Serv. Ind.* **2003**, *13*, 269–277.
17. Stokes, H.; Gummersall, S.; Shukla, J.; Johnson, R. Advanced Control Rooms for Boiling Water Reactors: The Nuclenet Control Complex. *IEEE Trans. Nucl. Sci.* **1973**, *20*, 786–800.
18. David, H.; Handley, H. A. H. Observations of Real-Time Control Room Simulation. 2020, 245–258. https://onlinelibrary.wiley.com/doi/10.1002/9781119698821.ch13.
19. Lau, N.; Powers, D. System-Task Analytical Framework for Monitor Assessment. In *2016 Resilience Week (RWS)*, 2016; pp 184–189.
20. Flegel, N.; Wessel, D.; Pöhler, J.; Van Laerhoven, K.; Mentler, T. Autonomy and Safety: A Quantitative Study with Control Room Operators on Affinity for Technology Interaction and Wish for Pervasive Computing Solutions. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* 2023; pp 1–10.
21. Al-Dabbagh, A. W.; Hu, W.; Lai, S.; Chen, T.; Shah, S. L. Toward the Advancement of Decision Support Tools for Industrial Facilities: Addressing Operation Metrics, Visualization Plots, and Alarm Floods. *IEEE Trans. Autom. Sci. Eng.* **2018**, *15*, 1883–1896.
22. Jr, C. G. S.; Hintz, K. J. Sensor Management in a Sensor-Rich Environment. In *Signal Processing, Sensor Fusion, and Target Recognition IX*; SPIE Press: Bellingham, Washington, USA, 2000; pp 48–57.
23. Kim, D. Y.; Kim, J. How Does a Change in the Control Room Design Affect Diagnostic Strategies in Nuclear Power Plants? *J. Nucl. Sci. Technol.* **2014**, *51*, 1288–1310.
24. Russell, J.; Masiello, R.; Bose, A. Power System Control Center Concepts. In *IEEE Conference Proceedings Power Industry Computer Applications Conference*, 1979; pp 170–176.
25. Staying in Control IEEE Power & Energy Magazine. https://magazine.ieee-pes.org/january-february-2012/staying-in-control/.
26. Van Oort, E.; Rosso, R.; Cabello-Montero, J. Evolution of Real-Time Operations Monitoring: From Concept to Global Implementation. In *SPE Annual Technical Conference and Exhibition*; SPE: Dallas, Texas, 2005; p 97059–MS.
27. Murugesan, R. Evolution of Industrial Automation. *Int. J.Comput. Appl. Technol.* **2006**, *25*, 169.
28. Cook, M. Mediated Decision Making in Multi-Crew Systems. In *International Conference on People in Control (Human Interfaces in Control Rooms, Cockpits and Command Centres)*; IEEE: Bath, UK, 1999; pp 235–241.
29. Adams-White, J. E.; Wheatcroft, J. M.; Jump, M. Measuring Decision Accuracy and Confidence of Mock Air Defence Operators. *J. Appl. Res. Mem. Cognit.* **2018**, *7*, 60–69.
30. Hudoklin, A.; Rozman, V. Human Errors versus Stress. *Reliab. Eng. Syst. Saf.* **1992**, *37*, 231–236.
31. Cummings, M. L.; Bruni, S.; Mitchell, P. J. Human Supervisory Control Challenges in Network-Centric Operations. 2010, *6*, 34–78.
32. Price, T.; Tenan, M.; Head, J.; Maslin, W.; LaFiandra, M. Acute Stress Causes over Confidence in Situation Awareness. In *2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*; IEEE: San Diego, CA, USA, 2016; pp 1–6.
33. Harris, S. D.; Narkevicius, J. M. Emergent Failure Modes and what to Do about Them. *INCOSE Int. Symp.* **2016**, *26*, 1044–1058.
34. Breslow, L. A.; Gartenberg, D.; McCurry, J. M.; Trafton, J. G. Dynamic Operator Overload: A Model for Predicting Workload during Supervisory Control. *IEEE Trans.Hum.-Mach. Syst.* **2014**, *44*, 30–40.
35. Taylor, G. S.; Reinerman-Jones, L. E.; Szalma, J. L.; Mouloua, M.; Hancock, P. A. What to Automate: Addressing the Multidimensionality of Cognitive Resources through System Design. *J. Cognit. Eng.Decis. Mak.* **2013**, *7*, 311–329.
36. Malakis, S.; Kontogiannis, T.; Kirwan, B. Managing Emergencies and Abnormal Situations in Air Traffic Control (Part I): Taskwork Strategies. *Appl. Ergon.* **2010**, *41*, 620–627.
37. Al-Tarawneh, H. A. The Main Factors beyond Decision Making. *J. Manag. Res.* **2011**, *4*; https://doi.org/10.5296/jmr.v4i1.1184.
38. Whittenburg, J. A. Automated Tactical C2 Systems: Do They Support the Decision Maker? In *Proceedings of the Human Factors Society Annual Meeting*; SAGE Publications: London, Vol. 34, 1990; pp 1153–1157.

39. Rhine, R. J. Command-and-Control and Management Decision Making. *Hum. Factors: J. Hum. Factors and Ergon. Soc.* **1964**, *6*, 93—100.
40. Cohen, I. Improving Time-Critical Decision Making in Life-Threatening Situations: Observations and Insights. *Decis. Anal.* **2008**, *5*, 100—110.
41. Veasey, D. A.; McCormick, L. C.; Hilyer, B. M.; Oldfield, K. W.; Hansen, S.; Krayer, T. H. *Confined Space Entry and Emergency Response*, 1st ed.; Wiley: New Jersey, USA, 2005.
42. Asaro, P. M. The Liability Problem for Autonomous Artificial Agents. In *2016 AAAI Spring Symposia*; Stanford University: Palo Alto, California, USA, 2016.
43. Porter, Z.; Habli, I.; Monkhouse, H.; Bragg, J. The Moral Responsibility Gap and the Increasing Autonomy of Systems. In *Developments in Language Theory*; Hoshi, M.; Seki, S., Eds.; Springer International Publishing, Vol. *11088*: Cham, 2018; pp. 487—493.
44. Elish, M. C. Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction. *Engaging Sci., Technol., Soc.* **2019**, *5*, 40—60.
45. Saleh, J. H.; Marais, K. B.; Favaró, F. M. System Safety Principles: A Multidisciplinary Engineering Perspective. *J. Loss Prev. Process Ind.* **2014**, *29*, 283—294.
46. Degani, A.; Wiener, E. Procedures in Complex Systems: the Airline Cockpit. *IEEE Trans. Syst. Man, and Cybern. — Part A: Syst. Hum.* **1997**, *27*, 302—312.
47. Bayrak, A. E.; McComb, C.; Cagan, J.; Kotovsky, K. A Strategic Decision-Making Architecture toward Hybrid Teams for Dynamic Competitive Problems. *Decis. Support Syst.* **2021**, *144*, 113490.
48. Alasmri, N.; Basahel, S. Linking Artificial Intelligence Use to Improved Decision-Making, Individual and Organizational Outcomes. *Int. Business Res.* **2022**, *15*, 1.
49. Waldman, A.; Martin, K. Governing Algorithmic Decisions: The Role of Decision Importance and Governance on Perceived Legitimacy of Algorithmic Decisions. *Big Data & Soc.* **2022**, *9*; https://doi.org/10.1177/20539517221100449.
50. Fleming, E.; Pritchett, A. SRK as a Framework for the Development of Training for Effective Interaction with Multi-Level Automation. *Cogn. Technol. Work* **2016**, *18*, 511—528
51. Akash, K.; McMahon, G.; Reid, T.; Jain, N. Human Trust-Based Feedback Control: Dynamically Varying Automation Transparency to Optimize Human-Machine Interactions. *IEEE Control Syst.* **2020**, *40*, 98—116.
52. Atoyan, H.; Duquet, J.-R.; Robert, J.-M. Trust in New Decision Aid Systems. In *Proceedings of the 18th International Conference on Association Francophone d'Interaction Homme-Machine — IHM '06*; Association for Computing Machinery: Montreal, Canada, 2006; pp 115—122.
53. Nicodeme, C. Build Confidence and Acceptance of AI-Based Decision Support Systems — Explainable and Liable AI. In *2020 13th International Conference on Human System Interaction (HSI)*; IEEE: Tokyo, Japan, 2020; pp 20—23.
54. Franke, T.; Attig, C.; Wessel, D. A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *Int. J. Hum.—Comput. Int.* **2019**, *35*, 456—467.
55. Silvast, A.; Virtanen, M. J.; Abram, S. Habits over Routines: Remarks on Control Room Practices and Control Room Studies. *Comput. Support. Coop. Work (CSCW)* **2024**, *33*, 39—58.
56. Flegel, N.; Pöhler, J.; Van Laerhoven, K.; Mentler, T. I Want My Control Room to Be": On the Aesthetics of Interaction in a Safety-Critical Working Environment. In *Mensch und Computer 2022*; Association for Computing Machinery: Darmstadt Germany, 2022; pp 488—492.
57. Rochlin, G. I. Safe Operation as a Social Construct. *Ergonomics* **1999**, *42*, 1549—1560.
58. Johnson, C. D.; Miller, M. E.; Rusnock, C. F.; Jacques, D. R. Applying Control Abstraction to the Design of Human—Agent Teams, 2020, *8*, 10**.**
59. Flemisch, F.; Heesen, M.; Hesse, T.; Kelsch, J.; Schieben, A.; Beller, J. Towards a Dynamic Balance between Humans and Automation: Authority, Ability, Responsibility and Control in Shared and Cooperative Control Situations. *Cognit., Technol. Work* **2012**, *14*, 3—18.
60. Knight, J. Safety Critical Systems: Challenges and Directions. In *Proceedings of the 24th International Conference on Software Engineering*; ICSE 2002: Orlando, FL, USA, 2002; pp. 547—550.
61. Neumann, P. G. Risks of Untrustworthiness. In *2006 22nd Annual Computer Security Applications Conference (ACSAC'06)*, IEEE, 2006; pp. 321—328.
62. Dunn, W. Designing Safety-Critical Computer Systems, 2003. https://ieeexplore.ieee.org/document/1244533/.
63. Wiegmann, D. A.; Rich, A.; Zhang, H. Automated Diagnostic Aids: The Effects of Aid Reliability on Users' Trust and Reliance. *Theor. Issues in Ergon. Sci.* **2001**, *2*, 352—367.
64. Wahlström, M.; Hakulinen, J.; Karvonen, H.; Lindborg, I. Human Factors Challenges in Unmanned Ship Operations — Insights from Other Domains. *Procedia Manuf.* **2015**, *3*, 1038—1045.
65. Battiste, V.; Lachter, J.; Brandt, S.; Alvarez, A.; Strybel, T. Z.; Vu, K.-P. L. Human Interface and the Management of Information. In *Information in Applications and Services*; Yamamoto, S.; Mori, H., Eds.; Springer International Publishing, Vol. *10905*: Cham, 2018; pp. 479—493.
66. NXP Ethical Framework for Artificial Intelligence. 2020; https://www.nxp.com/docs/en/white-paper/AI-ETHICAL-FRAMEWORK-WP.pdf.
67. Downey, M. Partial Automation and the Technology-Enabled Deskilling of Routine Jobs. *Lab. Econ.* **2021**, *69*, 101973.
68. Lee, S.; Hwang, I.; Leiden, K. Intent Inference-Based Flight-Deck Human-Automation Mode-Confusion Detection. *J. Aero. Inf. Syst.* **2015**, *12*, 503—518.
69. Mentler, T.; Flegel, N.; Pöhler, J.; Van Laerhoven, K. Man's (And Sheep's) Best Friend": Towards a Shepherding-Based Metaphor for Human-Computer Cooperation in Process Control. In *Proceedings of the 33rd European Conference on Cognitive Ergonomics*; Association for Computing Machinery: Kaiserslautern Germany, 2022; pp 1—4.
70. Jou, Y.; Yenn, T.; Yang, L. Investigation of Automation Deployment in the Main Control Room of Nuclear Power Plants by Using Adaptive Automation. *Hum. Factors and Ergon. Manuf. Service Ind.* **2011**, *21*, 350—360.
71. Kaber, D. B.; Endsley, M. R. The Effects of Level of Automation and Adaptive Automation on Human Performance, Situation Awareness and Workload in a Dynamic Control Task. *Theor. Issues in Ergon. Sci.* **2004**, *5*, 113—153.
72. Borges, G. In *Law and Technology in a Global Digital Society*; Borges, G.; Sorge, C., Eds.; Springer International Publishing: Cham, 2022; pp. 3—26.

73. Regulation — EU — 2024/1689 — EN — EUR-Lex. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.
74. Dorneich, M. C.; Rogers, W.; Whitlow, S. D.; DeMers, R. Analysis of the Risks and Benefits of Flight Deck Adaptive Systems. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Thousand Oaks, Vol. 56, 2012; pp 75—79.
75. Gouraud, J.; Delorme, A.; Berberian, B. Autopilot Mind Wandering, and the Out of the Loop Performance Problem. *Front. Neurosci.* **2017**, *11*, 541.
76. Dehais, F.; Peysakhovich, V.; Scannella, S.; Fongue, J.; Gateau, T. Automation Surprise" in Aviation: Real-Time Solutions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*; Association for Computing Machinery: Seoul Republic of Korea, 2015; pp 2525—2534.
77. Dalcher, D. Why the Pilot Cannot Be Blamed: a Cautionary Note about Excessive Reliance on Technology. *Int. J. Risk Assess. Manag.* **2007**, *7*, 350.
78. Marks, P. Are You Ready to Get on a Pilotless Plane? *New Sci.* **2014**, *223*, 30—33.
79. Curran, N. T.; Kennings, T. W.; Shin, K. G. Analysis and Prevention of MCAS-Induced Crashes. 2023. https://arxiv.org/abs/2301.08779.
80. Fominykh, D.; Levin, D. Engineering-psychological and Ergonomic Design of the Aircraft Crew Workplace. *E3S Web of Conf.* **2023**, *383*, 05003.
81. Myers, P. L.; Starr, A. W. Single Pilot Operations IN Commercial Cockpits: Background, Challenges, and Options. *J. Intell. Rob. Syst.* **2021**, *102*, 19.
82. Bell, K. Pilot Study for Cabin Crew's Willingness to Operate on Single Pilot Operations. *Coll. Aviat. Rev. Int.* **2024**, *42*; https://doi.org/10.22488/okstate.24.100226.
83. Harris, D. Single-pilot Airline Operations: Designing the Aircraft May Be the Easy Part. *The Aeronaut. J.* **2023**, *127*, 1171—1191.
84. Ahlen, A.; Akerberg, J.; Eriksson, M.; Isaksson, A. J.; Iwaki, T.; Johansson, K. H.; Knorn, S.; Lindh, T.; Sandberg, H. Toward Wireless Control in Industrial Process Automation: A Case Study at a Paper Mill. *IEEE Control Syst.* **2019**, *39*, 36—57.
85. Hancke, T. Ironies of Automation 4.0, 2020, *53*, 17463—17468.
86. Moulton, B.; Forrest, Y. Accidents Will Happen: Safety-Critical Knowledge and Automated Control Systems. *New Technol. Work and Employ.* **2005**, *20*, 102—114.
87. Hoffman, R. R.; Hawley, J. K.; Bradshaw, J. M. Myths of Automation, Part 2: Some Very Human Consequences. *IEEE Intell. Syst.* **2014**, *29*, 82—85.
88. Zerilli, J.; Knott, A.; Maclaurin, J.; Gavaghan, C. Algorithmic Decision-Making and the Control Problem. *Minds and Mach.* **2019**, *29*, 555—578.
89. Kim, S. W.; Kong, J. H.; Lee, S. W.; Lee, S. Recent Advances of Artificial Intelligence in Manufacturing Industrial Sectors: A Review. *Int. J. Precis. Eng. Manuf.* **2022**, *23*, 111—129.
90. Hoesterey, S.; Onnasch, L. Operators Over-rely Even More when Automated Decision Support Is the Exception and Not the Norm. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Thousand Oaks, Vol. 66, 2022; pp 1070—1074.
91. Onnasch, L.; Hoesterey, S.; Fahrner, V. To (Dis)agree with Automation: Effects of Automation Levels on Trust Attitude and Trust Behavior in High-Risk Situations. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Thousand Oaks, Vol. 67, 2023; pp 1393—1399.
92. Pithayarungsarit, P.; Rieger, T.; Onnasch, L.; Roesler, E. The Pop-Out Effect of Rarer Occurring Stimuli Shapes the Effectiveness of AI Explainability. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Thousand Oaks, Vol. 68, 2024; pp 352—358.
93. De Winter, J. C. F.; Dodou, D. Why the Fitts List Has Persisted throughout the History of Function Allocation. *Cognit., Technol. Work* **2014**, *16*, 1—11.
94. Fitts, P. M. *Human Engineering for an Effective Air-Navigation and Traffic-Control System*; National Research Council: Canada, 1951.
95. Time for AI to Cross the Human Performance Range in ImageNet Image Classification. 2020; https://aiimpacts.org/time-for-ai-to-cross-the-human-performance-range-in-imagenet-image-classification/.
96. Trivedi, H.; Wawira Gichoya, J. The LLM Will See You Now: Performance of ChatGPT on the Brazilian Radiology and Diagnostic Imaging and Mammography Board Examinations. *Radiol.: Artif. Intell.* **2024**, *6*, e230568.
97. Wang, Z.; Luo, X.; Jiang, X.; Li, D.; Qiu, L. LLM-RadJudge: Achieving Radiologist-Level Evaluation for X-Ray Report Generation, 2024. https://arxiv.org/abs/2404.00998.
98. Keshavarz, P.; Bagherieh, S.; Nabipoorashrafi, S. A.; Chalian, H.; Rahsepar, A. A.; Kim, G. H. J.; Hassani, C.; Raman, S. S.; Bedayat, A. ChatGPT in Radiology: A Systematic Review of Performance, Pitfalls, and Future Perspectives. *Diagn. Interv. Imag.* **2024**, *105*, 251—265.
99. Rajpurkar, P.; Irvin, J; Ball, R. L.; Zhu, K.; Yang, B.; Mehta, H. Deep Learning for Chest Radiograph Diagnosis: A Retrospective Comparison of the CheXNeXt Algorithm to Practicing Radiologists. *PLOS Med.* **2018**, *15*, e1002686.
100. Silver, D.; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifre, L.; Kumaran, D.; Graepel, T.; Lillicrap, T.; Simonyan, K.; Hassabis, D. A General Reinforcement Learning Algorithm that Masters Chess, Shogi, and Go through Self-Play. *Science* **2018**, *362*, 1140—1144.
101. Arasu, V. A.; Habel, L. A.; Achacoso, N. S.; Buist, D. S. M.; Cord, J. B.; Esserman, L. J. Comparison of Mammography AI Algorithms with a Clinical Risk Model for 5-year Breast Cancer Risk Prediction: An Observational Study. *Radiology* **2023**, *307*, e222733.
102. Martínez, E. Re-evaluating GPT-4's Bar Exam Performance. *Artif Intell. Law* **2024**, https://doi.org/10.1007/s10506-024-09396-9.
103. Onnasch, L.; Hösterey, S. Stages of Decision Automation: Impact on Operators' Role, Awareness and Monitoring. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Thousand Oaks, Vol. 63, 2019; pp 282—286.
104. Roesler, E.; Winhold, L.; Onnasch, L. Compete, Cooperate, or Both? Exploring How Interaction Settings Shape Human-Robot Interaction. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*; Association for Computing Machinery: Boulder CO USA, 2024; pp 901—905.
105. Bostrom, N.; Yudkowsky, E. *The Cambridge Handbook of Artificial Intelligence*; Frankish, K., Ramsey, W. M., Eds., 1st ed.; Cambridge University Press: Cambridge, 2014; pp 316—334.

106. Morris, A. T.; Maddalon, J. M.; Miner, P. S. On the Moral Hazard of Autonomy. In *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)*, San Antonio, TX, USA; 2020; pp 1—9.
107. Porter, Z.; Ryan, P.; Morgan, P.; Al-Qaddoumi, J.; Twomey, B.; McDermid, J.; Habli, I. Unravelling Responsibility for AI. 2024. https://arxiv.org/abs/2308.02608,arXiv:2308.02608.
108. Ryan, P.; Porter, Z.; Al-Qaddoumi, J.; McDermid, J.; Habli, I. What's My Role? Modelling Responsibility for AI-Based Safety-Critical Systems. 2023. https://arxiv.org/abs/2401.09459,arXiv:2401.09459.
109. Awad, E.; Levine, S.; Kleiman-Weiner, M.; Dsouza, S.; Tenenbaum, J. B.; Shariff, A.; Bonnefon, J.-F.; Rahwan, I. Blaming Humans in Autonomous Vehicle Accidents: Shared Responsibility across Levels of Automation. 2018. https://arxiv.org/abs/1803.07170,arXiv:1803.07170.
110. *Responsible AI for Disaster Risk Management: Working Group Summary — World | ReliefWeb*. 2021; https://reliefweb.int/report/world/responsible-ai-disaster-risk-management-working-group-summary.
111. Khalil, K. M.; Abdel-Aziz, M.; Nazmy, T. T.; Salem, A.-B. M. The Role of Artificial Intelligence Technologies in Crisis Response, 2008. https://arxiv.org/abs/0806.1280,arXiv:0806.1280.
112. Douer, N.; Redlich, M.; Meyer, J. Operator Responsibility for Outcomes: A Demonstration of the ResQu Model. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. *64*, 2020; pp. 278—282.
113. Bainbridge, L. Ironies of Automation. *Automatica* **1983**, *19*, 775—779.