Chartered Institute
of Ergonomics
& Human Factors

Human Factors in Highly Automated Systems

WHITE PAPER

#ciehf

> *“The sociotechnical system, including the engineers, flight controllers, and programmers on the ground, as well as pieces of machinery, was impressive, precise, even wondrous, achieving a successful landing on all six attempts. But it was not perfect. Programs alarmed, guidance over-shot, boulders appeared, people misspoke and buttons failed. In each case, human abilities intervened in unplanned ways, made decisions and landed the spacecraft on the Moon”.*

David Mindell, ‘Digital Apollo’

> *“Successful efforts going forwards will be those that wrap new machine intelligence capabilities around human competencies in order to get the best out of each.”*

Dr Alonso Vera, Chief of the Human Systems
Integration Division at NASA Ames Research Center

Contents

Foreword

Looking back, it seems surprising that despite a career of some 40 years working on human factors aspects of the research, development, verification and use of complex systems primarily in high-hazard industries, I have never worked directly on an automation project. Clearly, automation and digitisation have been all around me. It could hardly have been otherwise in a career that started by patching programs into an analogue computer and writing programs in Fortran, progressed rapidly through the launch of 'portable computers' (my first weighed over 7kg and was carried over the shoulder in a bag more than 75mm high), to the phenomenal power and capabilities of today's devices carried casually in our pockets that are increasingly central to our daily lives, even down to being able to order food in a restaurant.

I was, however, extremely fortunate to have been a student of the late Neville Moray, my psychology professor. Neville was a leading world authority on the characteristics and limitations of the processes by which the human brain pays attention to and makes sense of information and events in the world around us. Much of his research concerned human factors in industrial process control systems. The knowledge and ways of thinking about the role of people in complex systems I gained from Neville have remained central to my professional work throughout my career.

It is against that background that, despite not working directly in automation, I have had innumerable opportunities to study and learn from incidents when things have gone wrong in highly automated systems. Not only incidents with serious health, safety or environmental outcomes but also many times when systems have failed to deliver what they promised, with consequent fallback on relying on people to put things right. I have also been aware of the amount of high-quality research published in the human factors and applied psychology literature that could, and should, inform the development of highly automated systems.

It has long seemed clear to me that those responsible for investment in automation repeatedly both overestimate the abilities and reliability of automation and underestimate the extent to which their systems will continue to rely on people as well as the wider impact the introduction of automation can have on people and organisations. In particular, they underestimate the psychological challenges involved in expecting people to be able to monitor and attend to systems they rarely interact with,

understand what the automation is doing, and remain alert, competent and in a fit state to intervene if they need to. At the time when I held responsibility for human factors engineering standards in a major global corporation, the company's project engineering process assumed that the more highly automated a proposed new system would be, the less effort would need to be paid to human factors in design and development. That assumption was wrong at the time, and it remains wrong now.

The project to develop this white paper began with a working group comprising around 20 highly experienced human factors practitioners from across a range of industries. Through discussion, the enormity of the challenge we faced and the wide differences between industries and applications rapidly became clear. That led to the need to constrain the aspirations of what the paper could achieve. It also became clear however, that despite the differences, there were a core set of issues and challenges around human factors in automation that everyone shared. These led to the principles that are the core of the paper.

Being in the last year of my professional career, I very much hope that this white paper, issued with the authority of the Chartered Institute of Ergonomics & Human Factors and the vast experience of its members behind it, will make a positive and lasting contribution to raising awareness of the importance of human factors to highly automated systems. Awareness not only of the need to properly consider the human element but also of the kind of questions and challenges that typically need to be addressed. And awareness of the scope of work that needs to be properly funded and led by properly competent people when developing and implementing those systems.

Professor Ron McLeod
Leader, Human Factors in Automation
White Paper project

ACKNOWLEDGEMENTS

This paper was written by CIEHF Fellow Ron McLeod with support from a large number of CIEHF members and others.

Special thanks are due to Nora Balfe from Irish Rail, who made important contributions and suggestions as well as acting as a reviewer, and Alison Starr from the Manufacturing Technology Centre, who carried out a rigorous review of the final draft.

Thanks are due to the following individuals for reviewing a near-final draft:

- Mark Sujan from Human Factors Everywhere
- Rasmus Adler from Fraunhofer IESE
- Kevin Wong from Shell
- Paul Traub from Paul Traub Associates
- Paul Bowie from NHS Education for Scotland

Thanks are also due to the following individuals for their contributions to the Working Group tasked with preparing draft content for the paper:

- Joe Smyth
- Pascal Severin
- Emily Shaw
- Robin Burgess-Limerick
- Neil Correia
- Filip Florek
- Andrew Sutherland
- Iain Stuart
- Ned Hickling
- Andy Wright
- Julie Bell


Changzhou
浒墅关
Xushu Guan
苏州西
Suzhou West
江桥镇
Jiangqiao Zhen
Departure Yard

1. Introduction

In 1983, Stanislav Petrov was the duty officer at the command centre for a Russian nuclear early-warning system when the system reported that the United States had launched nuclear missiles at Russia. Fortunately, Petrov judged the reports to be false. His decision to disobey orders, against Soviet military protocol, is credited with having prevented a nuclear war. With the current international situation, it is unsettling to consider whether an individual in an equivalent position in the 2020s, whatever their responsibility, and even if they had the authority and the ability, would come to the same decision given the speed, reliability and extent of integration and opaqueness of today's automated systems.

In these challenging times, this white paper has been produced to address a widespread concern among members of the Chartered Institute of Ergonomics & Human Factors (CIEHF) over the frequent lack of adequate consideration of the roles and responsibilities of people when highly automated systems and products are being designed, developed and implemented. While many organisations have prepared guidance on introducing automation into products, manufacturing and business processes, seemingly few provide detailed and specific consideration of the human factors implications and risks associated with their introduction or how to address them.

These concerns must be set in the context of the prospect that the immediate future is likely to see a further, even more significant change in the relationship between people and technology. A move to the situation where technology increasingly has the authority – indeed, even the moral and/or legal responsibility – to make decisions and act autonomously without any direct involvement from people at the moment decisions are made and actions are taken. Increasingly, that includes decisions with major potential for conflict with civil liberties, personal freedoms and human rights.

Future historians looking back at the first half of the 21st century may recognise this as being the point when the balance between humans and technology fundamentally shifted. When a threshold was crossed from a world where technology supported humans, to one where technology effectively took control. Crossing that threshold will have a profound impact. The implications extend well beyond technology and economic benefits, to impacting on the legal and financial systems, as well as on the moral and ethical framework in which societies operate.

This white paper is intended for non-specialists who may have little or no professional background in human factors and ergonomics but who are influential in the way decisions are made about the development and use of technology. The knowledge and guidance it contains is based on both fundamental scientific and applied research, as well as from deep study and learning from adverse events[1].

Much of this knowledge has been generated in the aerospace, defence and nuclear power industries, where the reliance on technology to operate systems and manage risk has historically been especially prominent. More recently, a significant investment in high-quality research is being conducted in support of the development of automated and autonomous vehicles.

The paper provides an overview of some of the key human factors issues and concerns with highly automated systems. It is intentionally high level and broad in scope and is not linked to any specific application area. It draws on numerous lessons from both scientific research as well as applied experience that readers can use as points of practical learning and reference.

---

[1] See CIEHF White Paper on Learning from Adverse Events (2019).

1.1 PURPOSE

The purpose of this paper is to raise awareness in four areas:

- Key human factors challenges when introducing highly automated systems.
- The depth of knowledge available in the scientific and applied literature about how to optimise the role of people in those systems.
- Important learnings from adverse events where there has been a breakdown in the relationship between people and highly automated systems.
- Key principles that can help guide the development and implementation of highly automated systems.

The paper is based around nine principles. Their aim is to provide an easy-to-follow guide to human factors issues which need to be addressed when developing and implementing highly automated systems. The principles are:

1. Understand the potential influence of other elements of the system on the automated components, as well as how the introduction of automation can affect those components. Automation must be seen in the context of the overall socio-technical system it exists in.

2. Recognise that automation nearly always changes, rather than removes, the role of people in a system. Those changes are often unintended and unanticipated. They can make the tasks people need to perform more difficult and can disrupt established relationships, lines of communication and the ability to exert authority.

3. Be clear about which of the four core functions (acquiring information, extracting meaning from it, making decisions and taking action) automation will have the ability to perform for each system task, and under what conditions it will be given the authority to control those functions without human oversight.

4. Be realistic in acknowledging that people, at some level, are going to have to monitor, supervise, and hold responsibility for, the performance of the automation. Design, introduce and support the automation such that those people can maintain awareness of the state of both the automation and the world it operates in.

5. Ensure effective, transparent and unambiguous communication between the automation and the human elements of the system, such that the human is able to remain in the loop and situationally aware at all times.

6. For each task or function an automated system has the ability to perform be as explicit as possible where the balance between authority, responsibility and control lies. Be clear about what the expectations about responsibility imply for the different stakeholders in the system.

7. Ensure the people relied on to support the automation understand what the system is doing and why. There should be no automation surprises.

8. Avoid making unrealistic assumptions about the ability of people to monitor and effectively intervene in any system where there is little for them to do over sustained periods.

9. Recognise that automated systems can increase the levels of task difficulty and workload imposed on the human elements in the system as well as the level of human reliability needed in the inspection, calibration, maintenance and testing of system components.

LEARNING FROM THE SCIENCE BASE: THE IRONIES OF AUTOMATION

Despite being published 40 years ago, Lisa Bainbridge's concerns over the 'Ironies of Automation' remain relevant today (Bainbridge, 1982). Introducing automation rarely removes the human; rather it changes their role. Intentions behind the introduction of automation are often naive to this nuance. Bainbridge identified a number of ironies that often occur when automation is introduced:

- Designers use technology to perform processes and tasks that are easy to automate. They are well defined and predictable. This leaves the human with the tasks that cannot be readily automated. The tasks left to the human can be arbitrary and with little thought given to providing support for them.

- Errors introduced by designers in developing the automation can themselves be major source of problems.

- People who were previously highly skilled and experienced performing tasks manually, suddenly become novices in their new role.

- Without focused effort in the design and support of the system, people rapidly lose the awareness, skills and competence needed to be able to intervene and support the automation when they are expected to.

If not properly managed, the ambition to remove the risk of 'human error' through automation simply moves the risk to elsewhere in the system.

1.2 CONSTRAINTS

The scope of material covered in this white paper is constrained in four main ways:

- It is primarily concerned with highly automated systems intended to perform or support continuous, real-time control tasks in an industrial context and where the consequences of a system not performing as it is intended are likely to be significant. These might be in domains such as process control, transport, defence, manufacturing, mining, healthcare or other real-time operations. Much of the content and principles are, however, equally applicable to many consumer products (such as automated vehicles and the automation of domestic appliances and systems).
- The paper is focused on human factors issues and principles related to the early stages of thinking about the development and implementation of highly automated systems. The aim is to encourage and support organisations embarking on the development or procurement of automated systems or products to ask the right questions and initiate the necessary work to ensure the role of people is taken into consideration in decision making around the capability, design and use of the new systems. For example, there are very significant human factors issues associated with the management of change and transition to automation, as well as training and support of those who use or work with automated systems. There are also significant issues associated with the use, or interaction with, highly automated systems by groups of people with specialist needs, including not least the elderly. Both of these are complex but are beyond the scope of this white paper.
- The scope covers systems roughly up to and including what the Global Mining Guidelines Group refers to as Level 4, or 'Highly Autonomous Systems' – see figure 1.
- The paper is limited to human factors aspects of the role of the individuals most directly involved with the automation in real-time, and their immediate relationship with the system. The wider social, legal, organisational – and, indeed, emotional – issues that automation can create are outside the scope of the paper.

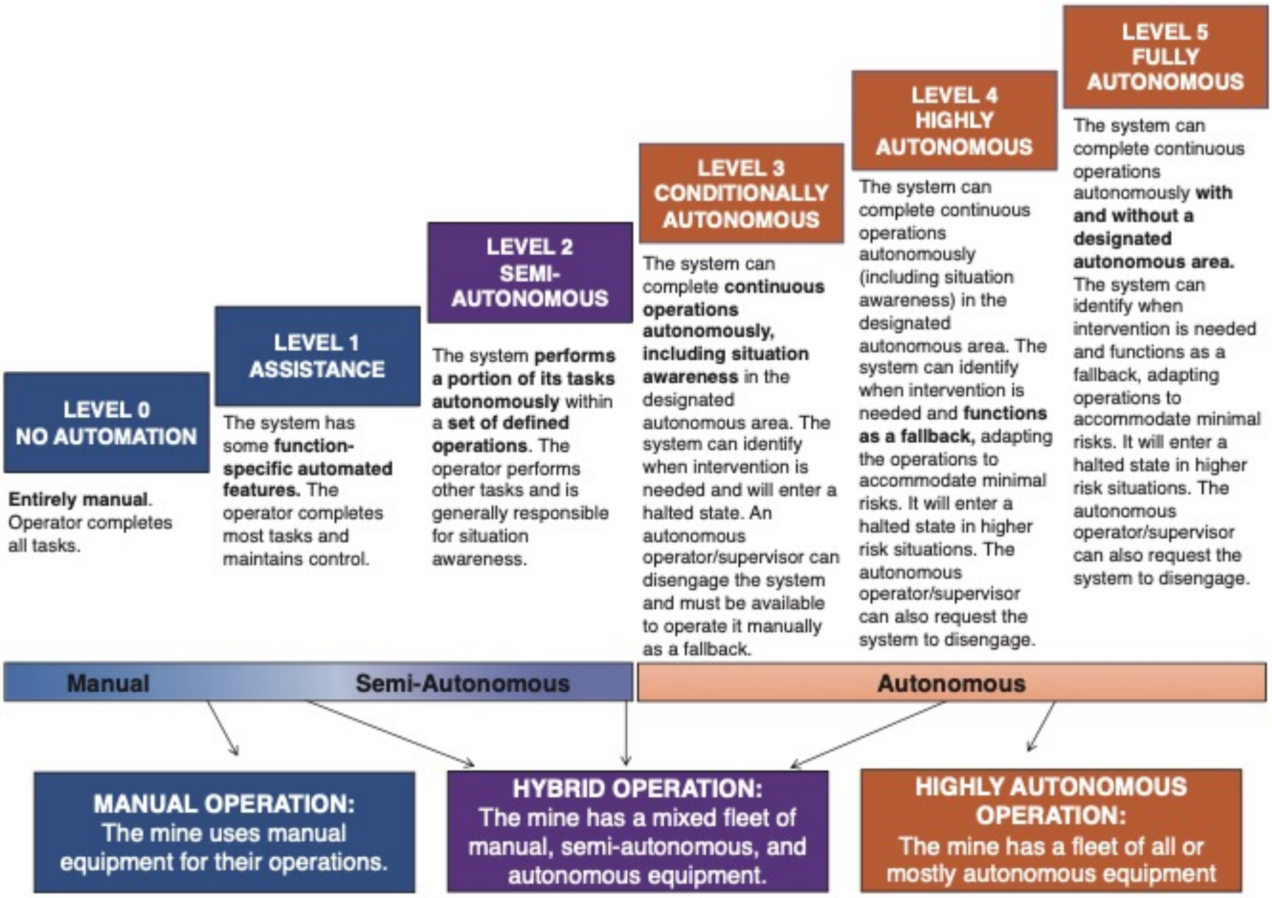
LEARNING FROM EXPERIENCE: FATAL INCIDENT WITH AN AUTOMATED VEHICLE

In March 2018 an Uber self-driving Volvo XC90 test vehicle hit and killed a pedestrian pushing a bicycle across the road. The vehicle failed to correctly classify the pedestrian with a bicycle and project their path as a potential collision risk.

The design and licensing of the vehicle assumed that the "safety driver" would be able to take-control in the event the auto-pilot did not perform as expected.

Unfortunately, the driver was distracted using their mobile phone to watch a TV show. Consequently, they were not in a state to recognise the failure of the auto-pilot and take control in the time available.

> *"..if you build vehicles where drivers are rarely required to respond, then they will rarely respond when required."*
>
> **Peter Hancock**
> National Transportation Safety Board, (2018).



Entirely manual. Operator completes all tasks.

The system has some **function-specific automated features.** The operator completes most tasks and maintains control.

The system **performs a portion of its tasks autonomously** within a **set of defined operations**. The operator performs other tasks and is generally responsible for situation awareness.

The system can complete **continuous operations autonomously, including situation awareness** in the designated autonomous area. The system can identify when intervention is needed and will enter a halted state. An autonomous operator/supervisor can disengage the system and must be available to operate it manually as a fallback.

The system can complete continuous operations autonomously (including situation awareness) in the designated autonomous area. The system can identify when intervention is needed and **functions as a fallback,** adapting the operations to accommodate minimal risks. It will enter a halted state in higher risk situations. The autonomous operator/supervisor can also request the system to disengage.

The system can complete continuous operations autonomously **with and without a designated autonomous area.** The system can identify when intervention is needed and functions as a fallback, adapting operations to accommodate minimal risks. It will enter a halted state in higher risk situations. The autonomous operator/supervisor can also request the system to disengage.

MANUAL OPERATION: The mine uses manual equipment for their operations.

HYBRID OPERATION: The mine has a mixed fleet of manual, semi-autonomous, and autonomous equipment.

HIGHLY AUTONOMOUS OPERATION: The mine has a fleet of all or mostly autonomous equipment

FIGURE 1: Mining Automation Maturity Model. (Taken from ‘Guideline for the implementation of autonomous systems in mining. Global Mining Guidelines Group, 2018)

1.3 STRUCTURE

Following this introduction, the paper is in four main sections:

- **SECTION 2** sets the scene by pulling together some important concepts and clarifying some of the terminology used.

- **SECTION 3** considers differences between the needs of different industries and discusses six human factors challenges that have repeatedly been found to impact on the reliability, performance or safety of highly automated systems.

- **SECTION 4** aims to help readers identify whether the issues about the role of people in highly automated systems are likely to apply to the kind of systems they are interested in. The section includes a screening tool to help structure thinking about whether there is a need to put effort into the human factors aspects of the design and development of the system.

- **PRINCIPLES AND RECOMMENDATIONS** are made throughout the document in the context in which they arise. Section 5 summarises the nine principles.

LEARNING FROM EXPERIENCE: OVER-RELIANCE ON A CYCLE COMPUTER: A PERSONAL STORY

As a keen cyclist, I recently invested in a cycle computer. Over time, I came to rely on it to go ever deeper into areas I didn't know, trusting the computer to tell me where and when to turn. Until the day I came to realise the risks I had fallen into by trusting it without properly integrating it into my planning.

One winter's evening, with dusk approaching, I found myself deep in the countryside a long way from anywhere familiar. With the temperature dropping, out of food and water and with no warm clothing, I realised the battery on the computer was about to run out.

The experience made me realise the extent to which I had put my safety into the hands of the technology. Being impressed and delighted with the new computer, I hadn't realised the implications of having no spare power or back-up navigation aids with me.

A cycle computer is simple. Complex systems have the potential to degrade or fail oin ways that can be much more difficult, sometimes impossible, to predict. Though the lessons from this experience are as true for the adoption of automation anywhere as they are for cycling.

The role and limitations of the technology, as well as the role of the people relying on it, need to be properly understood and prepared for. Not only when the technology works as expected, but when it is not available, or does not perform as expected. Ensuring people are available and capable of taking control if they need to is something no system that relies on automation can do without.

Ron McLeod
CIEHF Fellow

2. Setting the scene

1 **PRINCIPLE:** Understand the potential influence of other elements of the system on the automated components, as well as how the introduction of automation can affect those components. Automation must be seen in the context of the overall socio-technical system it exists in.

Automation is everywhere: in our homes, in our hands and in our workplaces, in transport and energy systems, in our hospitals and in the way our governments work. While the principles and experience set out in this white paper apply to most instances of automation, some will be more important than others depending on the particular application and context: while automatic doors in public places can be a risk should they fail in the event of fire, it is easy to identify simple and practical solutions.

"…Automation does not simply supplant human activity but rather changes it, often in ways unintended and unanticipated by the designers of automation, and as a result poses new coordination demands on the human operator." (Parasuraman et al, 2000)

This section sets the scene for the remainder of the paper by introducing some core concepts that are central to consideration of the role of people in highly automated systems:

- Being clear about the difference between 'machines' and 'automation'.
- Understanding different types and levels of automation from a human factors perspective.
- The importance of being clear about the relationship between the technological and human elements of systems in terms of which has the 'ability', 'authority' and 'responsibility' for system performance, as well as where the actual 'control' of performance lies at any time.

2.1 AUTOMATION OR MACHINE?

The term 'automation' can mean very different things to different people, depending on their objectives and the context of its use[2]. It can be a source of much confusion. The advent of the horse-drawn plough undoubtedly allowed a massive reduction in the physical exertion required in farming, as well as the productivity achieved, although few people would seriously argue that the horse-drawn plough was a 'machine' in the conventional sense. The invention of the steam engine, however, genuinely led to the widespread use of 'machines'. The key was the replacement of humans or animals with mechanical, electrical, thermo-dynamic or other sources of power.

Similarly, in the age of automation, while automated cruise-control in cars and aircraft remove the need for a human to manually control the vehicles speed, cruise-control is a long way from what is now intended when we refer to 'automated', and especially, 'autonomous' or even 'intelligent' systems.

For the purpose of this white paper, the following distinction is made:

- **MACHINE:** If the designers have very little uncertainty about the details and variability

---

[2] There is often confusion between uses of the terms 'automation' and 'autonomy'. In line with the UK Civil Aviation Authority (CAP 1377) the term 'autonomy' is taken to refer to systems that have "...self-determination and independence of decision-making..." Not all 'highly automated systems' need have autonomy.

of either the domain the product or system is expected to operate in, or exactly what the automation needs to do, or if the system does not have the capability to deal with unplanned variability in the domain, then the system is better thought of as a ‘machine’, rather than ‘automation’.

- **AUTOMATION:** By contrast, if there is significant uncertainty or unplanned variability about either the domain the product or system is expected to perform in, or the way functions are to be performed, but the system is capable of dealing with those uncertainties with little or no reliance on a human, then the system is considered as having ‘automated’ those functions. It not only has the ability, but it is given the authority to behave autonomously in performing one or more of the core functions without relying on human input.

The essence of this distinction is that to be considered as automation rather than a machine, the product or system must have the ability to detect and understand changes in the environment or circumstances it is operating in, and to adapt its behaviour accordingly. That is, it must have some degree of autonomy. Automation does not necessarily operate under a wider range of conditions than machines; but it has a sophisticated ability to detect changes in its environment, and to vary its actions in response to those conditions, that machines do not possess.

2 **PRINCIPLE:** Recognise that automation nearly always changes, rather than removes, the role of people in a system. Those changes are often unintended and unanticipated. They can make the tasks people need to perform more difficult and can disrupt established relationships, lines of communication and the ability to exert authority.

LEARNING FROM THE SCIENCE BASE: WHAT, AND HOW MUCH, TO AUTOMATE?

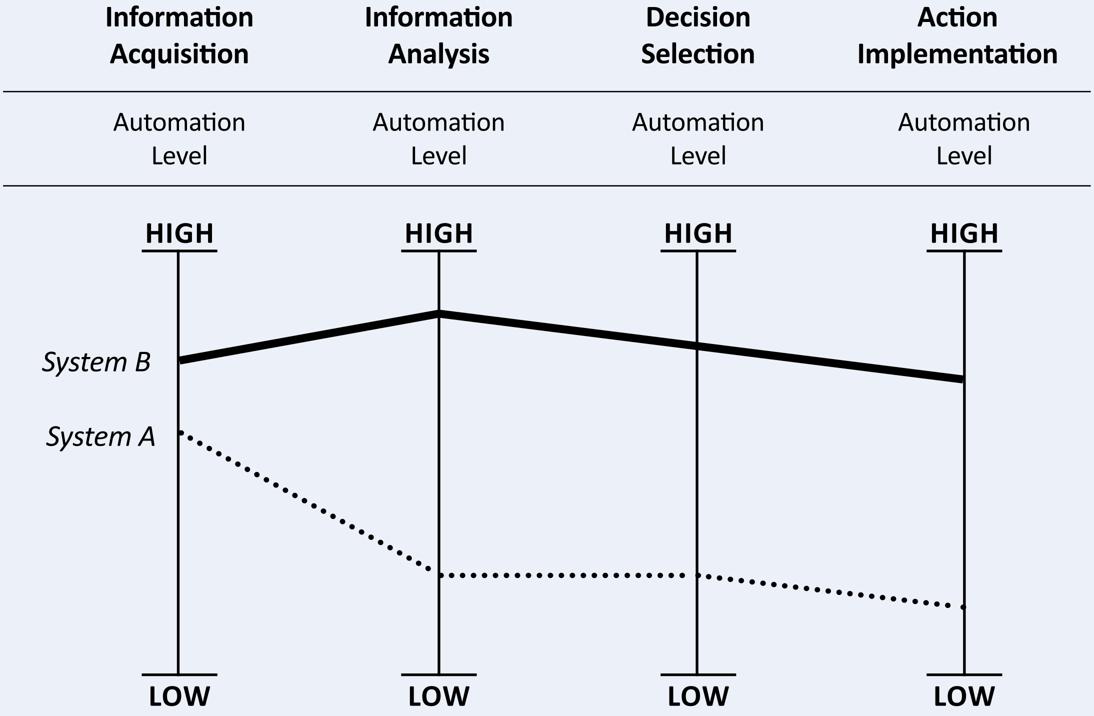
In a classic paper, Parasuraman et al (2000) addressed the question of when technology makes it possible to automate many aspects of a system, which functions should be automated and to what extent. To answer the question, they proposed that automation can be applied to four classes of function that map generally to a simple four-stage model of how the human brain processes information:

1. **Information acquisition**
2. **Information analysis**
3. **Decision and action selection**
4. **Action implementation.**

Parasuraman and his colleagues proposed that each of the four functions can be automated to differing degrees, as illustrated in the figure below.

Different combinations of types and levels of automation could be evaluated using primary criteria of human performance (mental workload and situation awareness as well as the potential for complacency and skill degradation) as well as secondary criteria (the reliability of the automation and the potential costs that might arise if the performance of the overall system was either incorrect or in appropriate).

Parasuraman et al did not claim their model provided comprehensive design guidance. Rather it was seen as providing a useful starting point for considering what types and levels of automation to implement in any system. Since its publication, the model has been widely used both by researchers as well as those involved in developing highly automated systems in a variety of industrial applications.



2.2 TYPES AND LEVELS OF AUTOMATION

Much of the technical and scientific literature has tried to distinguish between different ‘types’ and ‘levels’ of automation. It is important to be clear about the difference between these two terms.

From a human factors perspective, the term ‘types of automation’ is most usefully taken as referring to whether technology has the capability to control the performance of one or more of four core functions necessary to perform operational tasks[3];

- **Acquiring information:** Attending to sources of data about the state or nature of the world the system is expected to operate in that is relevant to achieving system goals and converting the data into information that is available for use in the system.
- **Extracting meaning:** Extracting meaning from the information attended to in a way that is directly relevant to performance of operational tasks in the short or long-term[4].
- **Making decisions:** Based on the meaning extracted from real-time information, making decisions about modifying or changing how the operational task is performed to continue to satisfy the system’s goals.
- **Taking action:** Effecting a change either on the system or, via the system or other agents, on the external world.

From a human factors perspective, these four core functions should be the starting point in defining what automation will do.

3 **PRINCIPLE:** Be clear about which of the four core functions (acquiring information, extracting meaning from it, making decisions and taking action) automation will have the ability to perform for each system task, and under what conditions it will be given the authority to control those functions without human oversight.

By contrast to types of automation, the term ‘levels of automation’ refers to the extent to which automation has the authority to control the performance of one or more of these four generic functions. There is an extensive literature looking at ways of describing different levels of automation[5]. Different industries have adopted slight variants of the definitions of levels of automation to suit their needs. However, for the purpose of this white paper, five simple levels can usefully be distiguished[6]:

1. None: entirely human, no automated support.
2. Low level automation.
3. Medium level automation.
4. High level automation.
5. Fully automated, performed with no human support.

Some research, especially in aviation, has tried to integrate the concepts of types and levels of automation to give a single indication of how highly automated a particular system is.

---

[3] The levels are derived from Parasuraman et al, 2000.

[4] The distinction between short and long-term relevance of information is important. Skilled – and especially ‘expert’ – people are good at recognising patterns over time and recognising the possible longer-term implications of information in a way that can drive future task performance, such as knowing to check in future if a possible ‘weak’ signal is developing towards a problem, or initiating tests or checks early if there is possible concern.

[5] The 10-point definition developed by Sheridan and Verplank (1978) has been used as the basis for many subsequent attempts to define generic of levels of automation.

[6] These are based on the Civil Aviation Authority’s CAP 1377.

The purpose has been to try to make comparative judgements across systems having different combinations of types and levels of automation. It seems to make intuitive sense that the more of the core functions the product or system is authorised to perform, and the higher the level to which each of those functions is automated, the higher the level of overall automation. However, other than for the purpose of giving a general impression of the nature of an automated system, little practical value is gained by trying to integrate combinations of types and levels of automation and to make relative judgements about the human factors implications of different combinations[7].

In the table below, the number refers to the 5 levels of automation. There is little practical value in attempting to make relative judgements about the human factors implications of the three hypothetical systems listed. Such judgements are only possible based on a detailed understanding of the tasks to be performed, and how the authority and control of each function is shared in different circumstances.

	System A	System B	System C
Acquire information	4	4	O
Extract meaning	4	3	4
Make decision	3	1	4
Act	0	4	5

2.3 ABILITY, AUTHORITY, CONTROL AND RESPONSIBILITY

Frank Flemisch and his colleagues (Flemisch et al, 2012) published an influential paper in 2012 that explored the importance of four 'cornerstone concepts' in the design of human-machine systems: ability, authority, control and responsibility.

- **Ability:** having the means and resources to execute control.
- **Authority:** What the actor (people or technology) is or is not allowed to do. Both the authority to exert control, as well as the authority to change the control authority.
- **Control:** Acting on the situation so it develops in a preferred way.
- **Responsibility:** being accountable for the consequences of control. Responsibility is assigned before control is exerted and evaluated afterwards.

Figure 2 summarises the relationship between these four cornerstone concepts[8]. These four concepts provide a powerful approach to thinking about and analysing the characteristics of highly automated systems that draws attention to the role and responsibilities of the human in the system.

---

[7] Though it is noteworthy that a considerable body of research effort, most notably in the aviation sector, has gone into trying to draw equivalences across different combinations of type and levels of automation.

[8] Flemisch et al (2012) also introduced a powerful graphical tool, known as the "A2CR" diagram, which can be used to visualise the relationship between the cornerstone concepts in a single diagram.

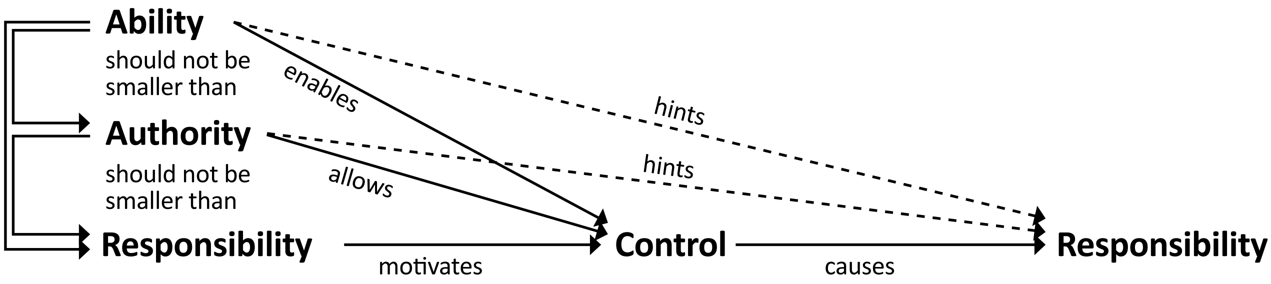


FIGURE 2: Relationship between Ability, Authority, Control and Responsibility (from Flemisch et al, 2012)

2.4 THE GULFS OF RESILIENCE AND ACCOUNTABILITY

Figure 3 illustrates differences between a wide variety of 'automated' types systems: some that are in operational use, others still in development. The systems shown on figure 3 are distinguished in terms of two dimensions:

A. The extent of the system performance envelope in which the automation is designed to accept the authority to exert control.

B. The ability of the system to detect and respond to unexpected events within the system boundary.

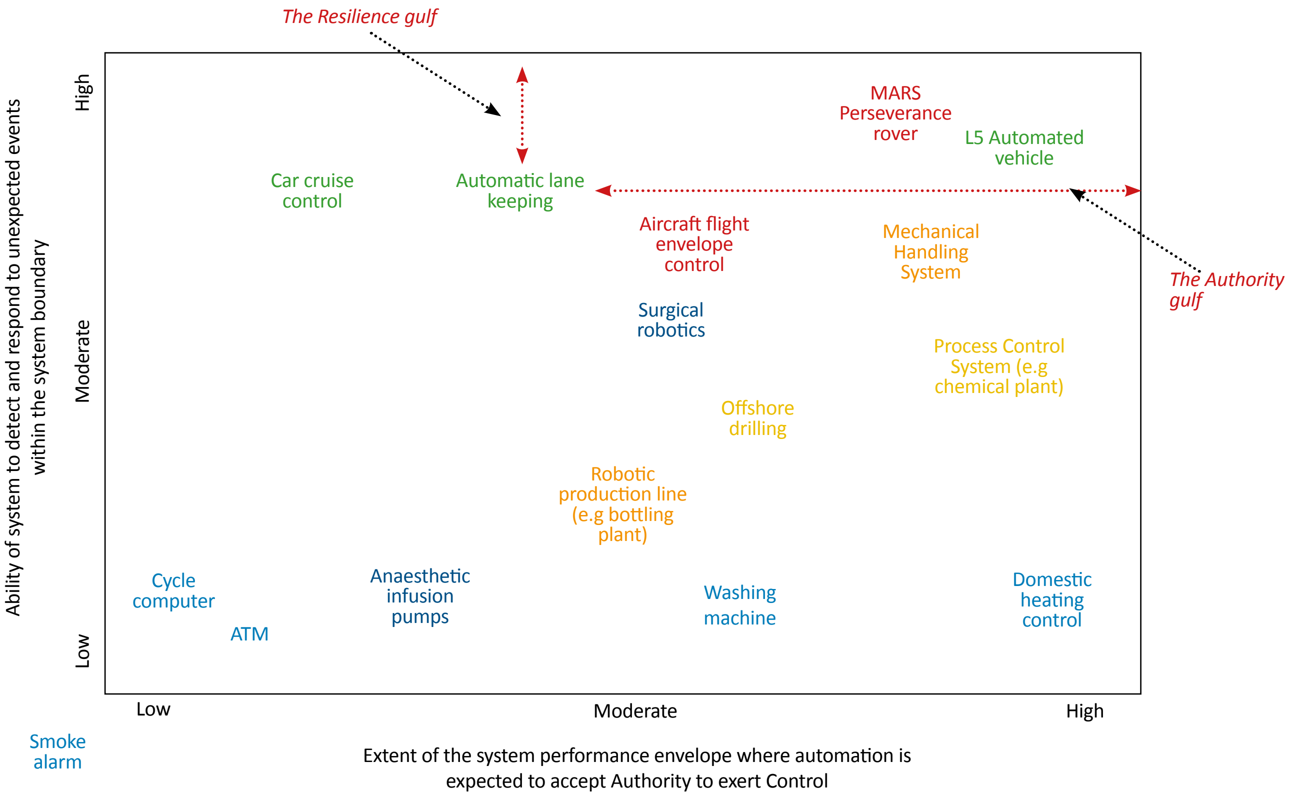


FIGURE 3: Two-dimensional framework for distinguishing between automated systems in terms of their reliance on human performance

Figure 3 directly conveys an impression of two areas where the automated system is going to place a high reliance on human performance. These are referred to on the diagram as two 'gulfs'[9]:

1. The 'resilience gulf', indicating the extent to which the system is not capable of responding to unexpected events within the system boundary without human support.

2. The 'authority gulf', indicating the extent of the overall system boundary where the automated system lacks the authority to exert control.

As an example, the extent of automated lane keeping in current generation automated vehicles is shown on figure 3 as being approximately 40% of the entire performance envelope needed of cars: so, it is located slightly to left of the centre of the horizontal axis. While the ability of current generation self-driving vehicles to detect and respond to lane limits is impressive, there remain a range of conditions and events associated with lane keeping (such as fog and poorly supported or maintained road infrastructure) that are beyond the abilities of current automated vehicles to detect and respond to without falling back on driver support.

The data used to position the automated lane keeping system on the two dimensions on figure 3 are both hypothetical. But they illustrate the point that with automated lane keeping in vehicles there remain gulfs of resilience to unexpected events, and of authority to exert control. The system relies on the human driver to bridge both gulfs. Those gulfs are part of the basis of the recent legal discussions in the UK and the introduction of the term 'user in charge'[10] to refer to the role of the human in such vehicles.

LEARNING FROM THE SCIENCE BASE: WHAT KINDS OF TASKS TO AUTOMATE?

The success of automation depends on the type of tasks involved. The key is understanding what, as well as what not, to automate. Cummings (2018) used the well-established Skill, Knowledge and Rule-based taxonomy of human error to explore the kind of functions where automation is best suited to support or replace human performance.

Automation is generally most suitable for tasks performed at a skilled level; vehicle control, precision milling, interpreting complex imagery, or even some forms of surgery for example. The required performance is well defined and performed in a highly constrained environment. Skill-based performance is largely sub-conscious, though it includes regular conscious monitoring to check that there are no unexpected problems.

Rule-based performance occurs when monitoring detects a problem that is recognised, and where there are learned and practised responses. Similarly with automation, if the problem has been anticipated in the design, and algorithms have been developed to deal with it, then automation should be capable of handling the problem.

But if the problem is not recognised, and there is no previously learned response to it, the human reverts to knowledge-based performance. That means drawing on deeper knowledge, experience, and recognition of similarities with other situations to work out how to deal with the problem. This can be difficult and complex, and an ideal situation for cooperation between the human and the automation.

In summary, human cognition, intuition and judgment are powerful resources in complex, poorly defined situations with a high degree of uncertainty. In these situations, cooperation and shared allocation of tasks between people and automation is often the most appropriate solution.

---

[9] In his 2013 book 'The Design of Everyday Things", Donald Norman introduced the concepts of two 'gulfs' people face when trying to interact with things. The "Gulf of Execution" is about figuring out how the item works or how to interact with it; the "Gulf of Evaluation" is about figuring out what happened in response to some user action. The "Resilience Gulf and "Authority Gulf" shown in figure 3 draw on these ideas, though they are different in nature.

[10] Law Commission of England and Wales and the Law Commission of Scotland (2022).

LEARNING FROM THE SCIENCE BASE: THE HUMAN AS A SYSTEM MONITOR

Introducing automation nearly always introduces the need for people to monitor the system; indeed, monitoring is frequently the main role of the human in the system. Unfortunately, as much experimental research shows, due to the nature of the human attention system, people are not good at sustained monitoring over the long term. A paper by Mica Endsley (Endsley, 2017) examines some of the challenges in the human monitoring of automated systems.

Perhaps surprisingly, passive monitoring of an automated system can be more difficult, with a higher mental workload, than if the automated task was performed manually. This is due to the human being ‘out of the loop’ for much of the time.

Maintaining situation awareness can become much more difficult through being out of the loop and as a result of poor communication between the system and the human: lack of transparency about how the system works, what it is doing, and what it is trying to do. In designing and introducing highly automated systems, effort must be given to providing effective support to the human monitoring role, and especially maintaining high levels of situation awareness.

3. Issues and challenges

3.1 INDUSTRY-SPECIFIC ISSUES

Most industries hold similar expectations of how they will benefit from introducing ever more automation: improved operational efficiency, quality consistency and pace, increased safety and avoidance of HSE risk, as well as lower costs associated with anticipated reduced workforce levels. At the same time, different industries have different priorities, and can have their own drivers for their pursuit of automation, or for implementing different types or levels of automation. For example:

- The nuclear industry is looking to automation primarily to increase safety.
- Mass transport industries seek to increase capacity and flexibility in their operations.
- Healthcare aims both to improve the quality of diagnosis and treatment as well as expanding its ability to offer services by automating administrative tasks and logistical processes.
- Oil and gas exploration is looking to access reserves in increasingly difficult and remote locations that would not be practical or economic – or, indeed, possible - if they had to rely purely on human performance.
- The defence sector is bound by rules of engagement and the Geneva convention on the use of autonomous weapons.

The situation and context in which different systems or products are used can be very different across industries and types of application. Differences range from what automated systems can do and what they are allowed to do without human intervention, to the role and relationship between people and technology, and, significantly, where responsibility for the overall performance and consequences of their use – as well as misuse - lies.

These differences arise from a combination of how and where automated systems are used, the legacy and shared industry experience and the types of technology applied, as well as the legal and regulatory framework in which an industry operates.

Different industries have adopted different models to describe what they mean by different degrees of automation: both type and levels. The transport industries have been most prominent in developing frameworks to describe different levels of automation. Though these models themselves only apply to a limited set of activities. For example, frameworks developed by the road and rail industries apply to driving operations, but not to control rooms; those in aviation apply to air traffic management but not to flight deck operations. Furthermore, these models only refer to systems that are in operational use; they do not cover the role that automation can play in the design, construction, maintenance or decommissioning of a system.

LEARNING FROM EXPERIENCE: APOLLO 14

“The advantages of manned space flight were again clearly demonstrated on this mission by the crew’s ability to diagnose and work around hardware problems and malfunctions which otherwise might have resulted in mission termination.”

Digital Apollo, p. 248

> *"…examples of healthcare AI applications include the use of patient-facing chatbots, mental health applications, ambulance service triage, sepsis diagnosis and prognosis, patient scheduling, planning of resources, quality improvements, and even the development of COVID-19 vaccines."*
>
> CIEHF Human Factors in Healthcare AI White Paper, 2021.

As the Mars Explorer has demonstrated so impressively, automation has made it possible to perform operations that would otherwise simply be inconceivable. As impressive as these systems are, automation in space exploration is still very far from being genuinely autonomous of human involvement despite the time delays associated with the vast distances involved. This raises challenges for shared human and technological control that are probably not experienced by any other earth-bound application. Earth-based mission controllers, analysts and scientists are relied on to perform detailed mission planning, as well as interpretation of data acquired by the remote vehicle. They must continually re-plan, update mission objectives and send short-term plans for activities to be carried out by the planet-based Explorer. Ground-based engineers are also required to oversee and monitor the performance of the vehicles numerous systems and to advise on optimising the systems performance and working life.

LEARNING FROM EXPERIENCE: DIGITAL FLIGHT STRIPS

> *"The planned replica of paper flight strips in electronic form…did not fly with our air traffic controllers. Despite slick algorithms for moving the electronic strips and sorting them in time or by level, the controllers just closed them and developed new ways of controlling the traffic."*
>
> Tony Licu, Head of Safety Unit at EUROCONTROL, describing experience introducing automated fight strips into air traffic management system in the 1990s. HINDSIGHT 33, Winter 2021-2022

Different industries and services also have their particular context and regulations. Defence applications, for example, must comply with the Geneva convention which constrains the operations that can be undertaken without human oversight.

There has been massive investment in the development of genuinely autonomous cars (as opposed to simply electric vehicles) in the past decade or so, together with growing national investment in the required supporting infrastructure. These investments have been driven through a combination of expectations about potentially enormous future returns on investment by the companies involved, combined with government expectations for improving road safety and reducing growing congestion problems. Compared with most other industries, however, the human factors challenges for autonomous vehicles are in some ways significantly more difficult. For example, a major part of the market for autonomous vehicles are private owners who are not professional drivers, but who are still expected to be able to buy and use autonomous vehicles safely off-the-forecourt, with minimal training or organisational support.

Table 1 summarises some of the ambitions, expected benefits and challenges of automation in different sectors.

TABLE 1: Examples of aspirations, expected benefits and challenges of automation in different sectors

Sector	Aspiration	Expected/achieved benefits	Example challenges
Oil, gas and renewables	Automated drilling	Enhanced safety and reliability. Increased efficiency. More consistent performance.	Uncertainty of the environment. Lack of digital knowledge and skills/ re-skilling. AI-enabled production. Automated grid-balancing.
Nuclear	Automation of production and protection functions	Improved cost-efficiency. Reduced risk and exposure to nuclear materials. Reduction in human error. Increased reliability and safety.	De-skilling. Over-trust/ over-reliance. Concerns over cyber security. Demonstrating automation reliability.
Commercial aviation	Automate as much of the tactical control of aircraft as possible.	Optimal balance of airspace capacity. Fuel efficiency and safety. Reduced costs. Reduced turn-round times/more revenue earning time in the air.	Keeping the pilot in the loop, skilled and situationally aware. Resilience/degraded operations. Clarity of roles and responsibilities. Human-automation interaction. Mode confusions. Conflicting goals of stakeholders. Legal liability (who is responsible for separation – currently the ground authorities).
Rail	Fully autonomous train driving, particular metro systems	Increased capacity. Improved safety. Fuel efficiency. Reduced operating costs and de-manning.	De-skilling. Complexity of mainline networks.
Road	Fully autonomous road vehicles for both commercial and private use	Improved safety. Increased capacity with reduced congestion. Improved access for persons of reduced mobility. Increased driver comfort and productivity.	Keeping the driver in the loop, skilled and situationally aware. Trust and acceptance. Motion sickness in non-driving related tasks. Mixed traffic composition (automated/ non-automated). Calibration as vehicles age.

TABLE 1: Examples of aspirations, expected benefits and challenges of automation in different sectors (continued)

Sector	Aspiration	Expected/achieved benefits	Example challenges
Defence	Fully autonomous defence systems	Reduced personnel costs. Reduced exposure/ improved survivability for own staff. Increased lethality for red forces.	Compliance with Geneva convention. Potential for accidental catastrophic consequences. Legal liability.
Healthcare	Automation of medical, administrative and logistical processes, and safety checks	Improved detection and diagnosis of disease. Higher patient throughput. Reduction of human errors. Improved reliability and consistency in administration of medication. Improved/ enhanced surgical procedures. Freeing up clinician time.	Interoperability of systems. Complexity. Complacency. Deskilling. Trust and acceptance. Training requirements.

3.2 KEY HUMAN FACTORS CHALLENGES

4 **PRINCIPLE:** Be realistic in acknowledging that people, at some level, are going to have to monitor, supervise and hold responsibility for the performance of the automation. Design, introduce and support the automation such that those people can maintain awareness of the state of both the automation and the world it operates in.

Despite the significant differences between industries, experience from numerous types of automated systems repeatedly identifies the same six human factors challenges that impact on the reliability, performance or safety of highly automated systems. These are:

1. Keeping the human ‘in the loop’ and situationally aware.

2. Enabling people to retain the skills they need to be effective performing their roles in the system.

3. Finding the right assignment of authority and responsibility balanced between the human and automation components of the system.

4. Avoiding people developing an uncritical sense of trust in the system, leading to complacency.

5. Inadvertently increasing the difficulty and/or mental workload involved in fulfilling their role in the system.

6. Avoiding automation ‘surprises’.

Note that these challenges are concerned with the role of the individuals directly involved with the automation in real time, and their immediate relationship with the system. They do not address the wider social, legal, organisational – or indeed emotional – issues that automation can create but that are outside the scope of this paper. The six challenges are however, remarkably consistent across industries and applications.

Challenge 1: Keeping the human ‘in the loop’ and situationally aware

5 **PRINCIPLE:** Ensure effective, transparent and unambiguous communication between the automation and the human elements of the system, such that the human is able to remain in the loop and situationally aware at all times.

One of the major issues that arises in the relationship between people and automation is when automation is so reliable that it is in control most, but not all, of the time, leaving the human with little to do other than monitor the system looking for signs that they may need to intervene in what the automation is doing. It is extremely difficult for people to concentrate and focus attention for more than short periods when all they are expected to do is monitor. Lack of attention and awareness of what is happening, and what the system is doing, or intending to do, quickly leads to the human becoming ‘out of the loop’ and losing situation awareness.

LEARNING FROM THE SCIENCE BASE: SITUATION AWARENESS (SA)

The concept of ‘situation awareness’ (SA) is fundamental to effective human performance in virtually any domain, from healthcare to nuclear power and from politics to sport. Inevitably, the concept is also central to assuring the role of people in highly automated systems (See Endsley, 2003).

Most commonly, SA is considered as comprising three increasingly complex levels of knowledge about the world around us:

- Level 1 SA is about being able to perceive information that tells us about the state of the world we are in.
- Level 2 SA is about understanding what that information means in terms that are relevant and directly useful to the tasks we are facing.
- Level 3 SA is about being able to use that understanding to project and allow us to predict and prepare for what is likely to happen in the future.

The applied psychology and human factors literature contains a great deal of knowledge about the nature, limits and properties of situation awareness. There is also a significant amount of knowledge about how to design systems that support people in developing and maintaining high levels of situation awareness.

The CIEHF white paper, ‘Human Factors in Healthcare AI’, contains examples of approaches to supporting situation awareness in a healthcare context.

“...it is impossible for even a highly motivated human being to maintain effective visual attention towards a source of information on which very little happens, for more than about half an hour. This means that it is humanly impossible to carry out the basic function of monitoring for unlikely abnormalities...” **Bainbridge (1972)**

Keeping the human in the loop and situationally aware demands that the human users remain actively involved and attentive to what is going on, such that they retain awareness of both the state of the world the automation is operating in and the state of the automation itself. This relies on features in the overall system design that provides the humans with some form of active engagement sufficient to allow the human to quickly and seamlessly support the system when needed. It also required the automation components to be transparent in the actions and processes they undertake in a way that the humans can engage with.

The most effective way of avoiding people becoming out of the loop and losing situation awareness is to ensure the people involved need to behave proactively in looking for and using information about the state of the world and the system. The alternative, being reactive, and simply responding to system-generated alerts, leaves the human disengaged from the system, with the consequent drift to a loss of situational awareness.

LEARNING FROM THE SCIENCE BASE: PROACTIVE OPERATOR MONITORING

How do people monitor proactively? How do they control how they allocate their limited attention when there are many sources of information that need to be checked?

In 1964, John Senders reported what became a classic experiment to help understand how operators in a process control environment allocate their visual attention across different information displays. The work was driven by concern over information overload in nuclear control rooms. There was a need to understand how humans deal with situations where they are expected to pay attention to a number of information sources changing at different rates.

Over time, people build an internal 'mental model' of the statistical properties of the world they are expected to monitor. That mental model is used, subconsciously, to decide when and how often to look at different information sources.

Challenge 2: Enabling people to retain the skills they need to be effective supporting automation

Introducing automation that is highly reliable, though not perfect, can make it very difficult for people to retain the knowledge and skills needed to be able to both recognise the need to intervene before it is too late, and to be willing and capable of intervening when needed in a timely manner.

> *"...a formerly experienced operator who has been monitoring an automated process may now be an inexperienced one"*
>
> **Bainbridge (1972)**

In recent decades, aviation has dealt with concerns over pilots becoming de-skilled with the introduction of increasingly capable flight deck automation, through massive investment in sophisticated, highly realistic and type-approved aircraft simulators. These allow pilots to develop and practice manual – and especially emergency – skills in a safe context. Simulation-based training, supported by international agreements and standards for simulator and pilot licencing, is central to modern commercial aviation operations. Despite this effort, incidents such as the loss of the Air France Airbus AF447 over the North Atlantic in 2009 (see separate box) show that, even if pilots have the necessary skills, the circumstances, including the speed of events in which they need to recognise the need to apply those skills, can sometimes overwhelm even highly experienced pilots such that they may not draw on those skills when needed.

Other, less highly regulated and safety critical industries, have taken different approaches to the issue of de-skilling. Rail, for example, has encouraged manual driving at off-peak times to allow drivers to maintain their skills. In the case of healthcare, the potential for automation to lead to a loss of essential skills is yet to be fully addressed. This seems, at least in part, to be due to a combination of automated systems being much newer and tending to be introduced based on local initiatives and the availability of local resources, including, not least, enthusiastic medical sponsors. Recognition among medical professions, professional bodies and regulators of the likely impact of ever-increasing reliance on automated systems in healthcare on the skill of medical professionals seems at risk of lagging behind the enthusiasm to deploy the systems.

Challenge 3: Finding the right assignment of authority and responsibility balanced between the human and automation components of the system

6 **PRINCIPLE:** For each task or function an automated system has the ability to perform, be as explicit as possible where the balance between authority, responsibility and control lies. Be clear about what the expectations about responsibility imply for the different stakeholders in the system.

Automation increasingly offers the ability to perform tasks and functions more efficiently, reliably, and accurately than can be achieved by people. Consequently, automation has often been introduced, or organisations have invested in automation based on expectations about the benefits and returns that the investment will deliver in performing specific tasks or operations.

There is, however, a very significant difference between automation that has the ability and is given the authority to perform a task or function under prescribed conditions, and where the ultimate responsibility for the performance of the system lies. System developers, and companies introducing automation to their processes often like to assume that the users of the systems will retain ultimate responsibility for the performance of the systems. However, if the design and implementation of a system creates situations where the human is pushed out of the loop, is unable to intervene, lacks situation awareness, or has lost the skills and knowledge to be able to ensure system performance, there are real concerns. Further, the willingness of the humans in the system – whether individual consumers or employees who may or may not be members of unions or other trade bodies – to accept responsibility (and, indeed, legal liability) when their ability to influence and exert overall control over system performance is degraded can be a major issue.

Probably the most high-profile example of this currently lies in the area of automated vehicles. There is a major debate underway between regulators, lawyers, insurers and vehicle developers about where responsibility in the event of automated vehicles being involved in accidents ultimately lies. Similar debates are, or will need to be, held in every domain, from aviation to medicine.

"…the analysis of problems with highly automated aircraft has shown that where responsibility is ambiguous or poorly indicated in the control station, several problems arise... less than perfect understanding of each other's abilities and characteristics could lead to misunderstanding between operator and machine, and hence competition for control."

Moray & Inagaki (1999)

Challenge 4: Avoiding people developing an uncritical sense of trust in the system, leading to complacency

The concept of complacency, what it means, and how, if at all, it differs from notions such as mindlessness and normalisation of deviance, is confusing. A reading of the technical and scientific literature suggests at least three ways of thinking about complacency: organisational, automation-induced, and situational[11]. While both organisational and situational complacency can be important in influencing human performance in highly automated systems, the principal issue for this white paper is around automation-induced complacency.

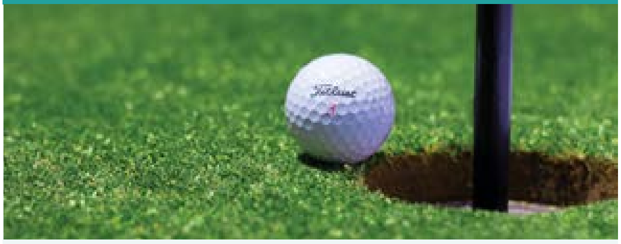
Automation-induced complacency (also referred to as 'automation bias') is the over-trust in technology that frequently develops when people use systems that are usually highly reliable and consistent. This is the case even when it is well known that they are not perfect and will, at some point, rely on human intervention.

LEARNING FROM EXPERIENCE: COMPLACENT BEHAVIOUR FOLLOWING SAT NAV

In 2009, the BBC reported an incident where a car was left teetering on a cliff edge after the driver followed sat nav directions down a footpath. The driver continued to follow the instructions when they told him the narrow, steep path he was driving on was a road. He only stopped when his car hit a fence above a railway bridge. The driver – who told the police he relied on his sat nav for his job – was charged with driving without due care and attention.

http://news.bbc.co.uk/2/hi/uk_news/england/bradford/7962212.stm

LEARNING FROM EXPERIENCE: COMPLACENT BEHAVIOUR ON THE GOLF COURSE



I was about to play my second shot on the 16th hole at my local golf club. My partner and I looked at our golf watches which both said there were 143 yards remaining. "It seems a lot further than that," my partner said. I agreed. We both hit 8 irons and came up 50 yards short.

As we approached the green, my partner asked me what hole my golf watch said we were playing. "14th," I said. As we were playing winter greens, our watches had not automatically moved the holes on as we played. Despite both of us, from what we could see with our own eyes, doubting the distance shown, we trusted the watches. And we both played the wrong shot.

"..even though performance on the task was substantially degraded….almost half the pilots used the automation when it failed…"

Parasuraman and Riley (1997)

Unfortunately, there is no simple solution to avoiding automation-induced complacency. As Moray and Inagaki (1999, see box) have shown experimentally, it seems to co-vary with factors including the inherent reliability of the system, as well as people's self-confidence in their own abilities. What is clear is that simply relying on people to avoid placing undue trust in the automation they use and work with is not an effective approach.

---

[11] See McLeod, 2020, for a discussion and a model of different types of complacency.

Challenge 5: Inadvertently increasing the difficulty and/or mental workload involved in fulfilling their role in the system

When automation is being introduced, it is easy to assume that the work of the people who remain part of the system will be made easier. Unfortunately, experience shows that that is often not the case. Indeed, when the role of the human changes to being one of monitoring a system for signs that it is not performing well or supporting it in performing tasks that rely on human judgement or decision making, the workload and difficulty of the human's tasks can increase. That is especially the case when things go wrong, and the human is expected to diagnose what has happened and intervene effectively, often under time pressure. Increased task difficulty is often associated with the challenge of remaining situationally aware, or of working out why the system is not working properly when, most of the time, it operates successfully with little or no human intervention.

In the case of automation of manual tasks, it is usually a fair assumption that the physical effort remaining for the humans involved will be reduced. However, even in manual handling systems, the humans can, for example, still be expected to perform physical activities that are awkward and expose them to musculoskeletal injury in manoeuvring heavy and awkward items into a position where the automation is able to take over.

Challenge 6: Avoiding automation 'surprises'

"Automation surprises begin with miscommunication and misassessments between the automation and users which lead to a gap between the user's understanding of what the automated systems are set up to do, what they are doing, and what they are going to do"

Woods & Sarter (2000)

A fundamental tenet of human performance is that, over time and as people gain more experience of a system, they build an internal mental representation of the properties of the system. The individual draws on this 'mental model' to help them understand how the system works and to help them predict what is likely to happen in the immediate future[12]. A user's mental model will initially be based on knowledge and information about the system from training and user manuals, as well as their expectations based on experience with (at least what they consider to be) similar systems. Over time, the model will be updated and modified by the conclusions and explanations the individual makes about how the system works based on their own observations and experiences working with and observing the system. Experience and learning from many incidents involving automated systems have shown that, even with highly trained and experienced people, the internal model people build to help them make sense of how the system works and what it is doing can be very far from reality.

The term 'automation surprises' refers to situations where an automated system does something that the people working with or using the system did not expect, and do not, at least immediately, understand. Surprises usually arise from a gap between the individual's mental model of how the system works and the reality of how it actually works.

Being surprised by something the system does is not necessarily a problem. If there is time to understand what has happened and, if necessary, intervene, or if there is in-built redundancy that lets the user quickly regain their understanding there may be no adverse consequences. Problems arise, however, if a surprise in a critical system makes it clear to the people involved that the mental model they hold of how the system works or what it is doing is wrong: that they have lost situation awareness and are completely 'out of the loop'. Even more critical are situations where the automation is suddenly unable to function, and the human is surprised to find themselves having to take over manual control of the operation unexpectedly. This is precisely what happened in the crash of the Air France Airbus over the North Atlantic in 2009.

---

[12] Psychologists and human factors professionals refer to this understanding as the human's 'mental model' of the system.

7 **PRINCIPLE:** Ensure the people relied on to support the automation understand what the system is doing and why. There should be no automation surprises.

Designing automated systems without the potential for serious automation surprises can be a major challenge. Meeting that challenge requires, among other things, a very clear understanding of the relative roles of the human and automated elements in performing critical tasks and functions, especially where there is a need for collaborative control. In particular, it requires situations where the system could need to hand control over to the human, or could need to rely on the human to make a critical decision with little advance warning, to be recognised and be taken proper account of during system design. Avoiding surprises in those critical situations also demands an effective strategy for real-time communication of what the system is doing and why, combined with strategies for ensuring the human is alert and situationally aware.

LEARNING FROM THE SCIENCE BASE: TRUST AND COMPLACENCY

At the heart of the success of systems that rely on people to monitor and support automation is people's willingness to intervene if they think the automation needs support. Central to this is the issue of trust.

A body of human factors research has explored the concern that, as people get used to increasingly sophisticated and reliable automation, they become complacent. The more they trust the automation, the less likely they are to monitor it effectively and to intervene when they should.

Moray and Inagaki (1999) reviewed the experimental research into these issues. Results show that trust in automation is a complex subject affected by many factors: how it is introduced; how self-confident people are in their ability; how much opportunity they have to perform manually; as well as long-term expectations, among other things.

People appear to have a high level of trust when they first experience an automated system. That trust declines as they become aware of the system's limitations, before increasing again as they learn and develop strategies to overcome the system's weaknesses.

An important determinant of trust, especially where automation is used to support fault diagnosis or decision making, is whether the human agrees with the automation. Even if the system is correct, if people do not agree with its diagnosis or recommendation, they will tend to lose trust in it. There is also evidence that people are less tolerant of errors made by automation than of the same errors made by other people.

Trust is also affected by how easy it is to know what the system is likely to do, as well as how dependable it is: "...several crashes of modern airliners have been due to a failure to understand the behaviour of the automation".

4. Human factors in the development of highly automated systems

"Designers and developers of AI, individuals with responsibility for procuring AI applications, regulators, and bodies funding research and development need to move beyond the technology-centric view, and instead approach AI from a systems perspective, i.e., to consider from the outset the interaction of people with AI as part of the wider clinical and health system."

CIEHF Human Factors in Healthcare AI White Paper, 2021

All socio-technical systems of any complexity, and especially those performing functions that are in any way critical, require attention to human factors principles in their design, development and implementation. Standards, best practices and guidelines, as well, in some cases, as regulations, exist and are now widely applied to meet the needs of different industries in integrating human factors into their projects. These aim to ensure human-machine systems are well-designed, and the role of the humans, operators, maintainers and other stakeholders, are properly supported through training, support systems and in other ways.

There is, however, a seemingly widely held misconception that, as automation is usually intended to simplify or replace human tasks, there is little need to spend money on human factors activities during the development of highly automated systems. Further, the business case for investing in highly automated systems often relies on savings in the cost of labour (i.e., fewer or less skilled people) or reductions in health, safety, security and environment risk. These assumptions can also lead to a justification not to allocate resource to human factors aspects of design that would be recognised as being needed in more conventional systems. Decades of experience, supported by learning from innumerable incidents, demonstrate that the opposite is, in fact, often true. The necessity of designing and implementing highly automated systems that are effective in keeping human operators 'in the loop', situationally aware and capable of intervening at short notice when needed means that human factors input to the development of those systems can be every bit as important, indeed, often even more so, than with other types of systems. Highly automated systems can require attention to human factors issues that is different to the attention typically needed in the design of other types of human-machine systems[13].

LEARNING FROM EXPERIENCE: INTRODUCING DIGITISED FORMS

A rail operator tried to introduce 'automated forms' to replace paper-based forms used by signallers. But rather than acknowledge that moving the forms to a tablet fundamentally changed the task, they simply transferred the same format of paper form to a digital environment.

The opportunity to use automation to pre-populate fields or otherwise assist with completing the forms correctly was lost. The 'automated' forms took longer to complete and the operator no longer had the ability to manually annotate them as they had before.

---

[13] Based on consideration of experience in the mining industry, Burgess-Limerick (2020) suggests the scope of a human-systems integration programme suitable for use during the procurement of automated systems.

LEARNING FROM THE SCIENCE BASE: THE NEED FOR BALANCE

Norman describes some of the challenges of designing an automated system without a comprehensive understanding of the human need. He explains the need to balance the capability of the technology with that of people.

The human cannot be too isolated from the automation in case there is a need to override or take control. Though there is little to be gained if the automation does not take over at least some of the human's role.

Automation must be balanced and appropriate. Simply aiming for 'more automation' is not only futile, but potentially dangerous.

Norman, D (2010)

The six key human factors challenges described in the previous section are complex, and highly cognitive in nature (leaving aside the wider social and organisational impact of introducing automation). Solutions that address these challenges cannot simply be specified in engineering standards. They require careful analysis, as well as an approach to developing and testing solutions, that is different both in nature and in technical difficulty from many other human factors issues.

4.1 CHARACTERISTICS OF HIGHLY AUTOMATED SYSTEMS

There are several characteristics of automated systems that indicate when special attention will need to be given to the human factors implications during their design, development and implementation. For example, considering the system in its entirety, if all of the following characteristics exist, there will almost certainly be a strong imperative for investing in a suitably focused human factors effort throughout the design, development and proving of the system:

1. The system is considered critical. Failure to perform any of the operational tasks to the standard expected has the potential to result in significant adverse consequences, such as in terms of health, safety, or financial performance, environmental control or reputation.

2. Achieving the system goals requires sustained and continuous performance over time (as opposed to being a momentary intervention, for example during an emergency, such as emergency shut-down systems or emergency braking systems).

3. The system is designed to work in an environment that is constrained and whose limits can be specified to a high level of detail (referred to as the Operational Design Domain or ODD[14]).

4. Within the limits of the ODD, the system is expected to perform one or more of the four core functions with little or no human support:
    - Acquiring information
    - Extracting meaning
    - Making decisions
    - Taking action.

5. Without relying on the human element in the system, it is not possible to design the system either to:
    - perform completely all of the operational tasks needed within the ODD, or
    - detect and respond in a satisfactory way to all events or situations likely to lead to the system having to continue functioning outside the boundaries of the ODD.

6. The system does not have the complete ability to monitor its own state and/or recognise when it needs calibration or unplanned maintenance (i.e., outside the manufacturers' recommended planned maintenance schedule).

---

[14] The term Operational Design Domain is taken from the vehicle automation industry. See SAE J3016 (2018).

4.2 CHALLENGES FOR HUMAN FACTORS DESIGN ANALYSIS FOR HIGHLY AUTOMATED SYSTEMS

Projects developing highly automated systems must ensure they fully understand the role of people in their systems before they commit to solutions. In critical systems, failure to do so can potentially lead to catastrophic consequences with resulting damage not only to an organisation's reputation, but also to the public confidence.

Setting out a comprehensive approach for integrating human factors into projects that meets the needs of highly automated systems is beyond the scope of this white paper. However, human factors considerations nearly always demand that detailed analysis, often supported by research in the case of more novel or critical applications, is conducted early on in a project's lifecycle to ensure decisions about the design and implementation of the systems are properly informed in several key areas. Examples of issues that need to be addressed at the different stages of the development process are summarised on the following pages.

"Design the human into the process. Design systems around what humans need in order to respond to unanticipated events."

Dr Alonso Vera, Chief of the Human Systems Integration Division at NASA Ames Research Center

LEARNING FROM EXPERIENCE: INTRODUCING AUTOMATION TO DRILLING

Until around 1990, drillers in oil and gas exploration usually stood at the driller station which was open and exposed to the weather. They would have one hand on a long metal brake lever, the other on a clutch and the foot resting on the throttle. They looked like a spread eagle standing at the driller station.

But they received a lot of feedback on how well the drilling was progressing from the vibration of the brake handle, the sound of the rotary table and the sight of the cable reel on the draw-works.

With improvements in technology, the driller was moved to a comfortable chair inside an environmentally controlled driller's shack. They now control the brake, clutch and everything else through joysticks and watch the operation through screens mounted in front of the chair.

This workstation has undoubtedly improved performance and reliability for most drillers, though many experienced drillers recognise that the loss of the ability to directly see and feel the operation has diminished their sense of situation awareness and control over the operation.

From McLeod, R W (2015) Designing for human reliability.

8 **PRINCIPLE:** Avoid making unrealistic assumptions about the ability of people to monitor and effectively intervene in any system where there is little for them to do over sustained periods.

LEARNING FROM EXPERIENCE: INCORRECT ASSUMPTIONS ABOUT AUTOMATION LED TO POOR WORK ENVIRONMENT

The design of a process unit for a new refinery made the assumption that a nearby river could be used as a source of cooling water without any need for cleaning the water.

Based on this assumption, the design team concluded that operators would only need to access valves to manually control the flow of water very rarely. The valves were therefore assessed as being of low priority. Little effort was therefore made to ensure the valves were easy to access or operate.

Once the plant became operational it was quickly realised that the quality of water from the river did not meet the standard assumed in the design. As a consequence, a number of operators had to be regularly assigned to open large manual valves. This led to a much a higher demand on operator time than was expected.

The work involved was also physically exhausting with the potential for musculoskeletal injury due to the poor accessibility and working postures forced by the lack of adequate design to provide good access to the valves. Furthermore, the lack of consideration of human factors in the design created the potential for human error and led to rule-breaking by encouraging operators to stand on piping to access and operate the valves.

During feasibility studies and concept development

- Ensure functions to be automated are recognised as existing in the context of the overall socio-technical system in which they exist. Ensure the potential influence of the wider system on the automated components are understood, as well as the potential impact of the automation on other system components.

- Be realistic about where responsibility for the performance of the overall system will ultimately lie, and whether it might be acceptable, legally or morally, for the owner of that responsibility to vary depending on the circumstances of use.

- Understand how the introduction of the automation is likely to change the roles, tasks and responsibilities currently assigned to people, as well as how those changes might influence established relationships between different people and roles.

During initial design

- Be as clear as possible about the limits of the ODD and the extent to which the system will be capable of monitoring whether the current conditions are diverging from the conditions defined in the ODD.

- Based on the required system capability, be clear about where the technical and human components will operate and if this can change between them at any time. And, based on that capability assessment, be clear about where, and under what conditions, technical components can safely be given the authority to perform functions without human input.

- Recognise that automation is likely to increase the workload and levels of human reliability required. Ensure the needs of those who will be relied on to inspect, calibrate, maintain and test system components are properly taken into account in the design of the system, including requirements for training the human components.

During detailed design

- Identify and understand the total set of functions that must be performed, as well as how performance of those functions is to be allocated between the human and technical components of the system. This includes understanding what capability will be needed to monitor and detect developing problems, and either to recover when automated systems are detected as failing, or to gracefully degrade to human control when necessary.

- Ensure the requirements for communication and transparency between the automation and the human elements of the system are understood, such that the human can remain in the loop and situationally aware to the extent necessary.

- Understand what features, information, space and support are going to be needed to ensure the human elements can be effective in filling their role in the system. That includes ensuring those features are designed in a way that is consistent and compatible with the need for people to be able to perform safely and efficiently and without risking their health or wellbeing.

The types of human factors analyses involved, and the levels of human factors competence and resources needed to support the system development, will vary depending on the response to considering each of these issues[15].

There are times when it is necessary to recognise that the optimal solution might mean allocating some functions to people, even though the technology has that ability. At other times it means recognising that, even although technology will be given the authority and responsibility for performing some functions within the ODD, there is still a need to invest in human factors to deal with situations that are at the limit, or go beyond, the ODD. This means recognising not only the type of people likely to be involved, but the situation and context they are likely to be in when they are called on to act. It also means ensuring both that the system provides the information and controls they will need, and that those features are implemented in a way that is compatible with human factors design principles. Finally, it can mean recognising the need to invest in training, or other forms of user support, as part of the overall system development, to ensure that people have the capability to fulfill their role in system. Given the nature of highly automated systems and the ways technology is developing, some of those user training and support solutions will demand novel and original approaches, that themselves demand human factors attention in their development.

LEARNING FROM EXPERIENCE: TRAIN DOORS LEFT OPEN

In 2019, a train in southern England travelled 16 miles at speeds of up to around 80 mph with a door to one of its passenger coaches open (RAIB, 2019).

Operation of the door was controlled by the driver from the cab. The open door was not visible to platform staff. Because the door's automated interlock system failed, the driver was given a visual indication that all doors were closed and the train was able to move when commanded by the driver.

Due to poor maintenance and inspection, two screws had worked loose from a bracket attached to the door. A microswitch detected the piston rod had moved to the doors closed position even though the door itself remained open. This released the interlock, allowing the train to move away with the doors open.

---

[15] McLeod and Balfe (2022) have described an example of an analysis method developed during the course of preparing this white paper and based on these areas that has the potential to be used early in the development of highly automated systems to identify where focused human factors effort is likely to be needed.

9 **PRINCIPLE:** Recognise that automated systems can increase the levels of task difficulty and workload imposed on the human elements in the system as well as the level of human reliability needed in the inspection, calibration, maintenance and testing of system components.

LEARNING FROM EXPERIENCE: A PACKET ROCKET: HOW SEEMINGLY MINOR CHANGES CAN BREAK YOUR SYSTEM

In 2008, an engineer at the Edwin I Hatch nuclear power station inadvertently forced a controlled emergency shutdown of the plant after installing a software update on an office computer on the business network.

The update, designed to synchronise data between business system and control system environments, caused a reset on the control system and the subsequent activation of plant safety systems. Whilst all safety systems performed as designed, the incident led to a loss of electricity generation for the licence holder, Southern Nuclear.

4.3 A HUMAN FACTORS SCREENING TOOL FOR AUTOMATION PROJECTS

Many industries now recommend initiating the integration of human factors into their system development processes by means of some form of human factors screening (see for example Energy Institute, 2020) or an ‘Early Human Factors Analysis’ (MoD, 2021). These screenings recognise that, with increasing awareness of basic human factors principles across the many engineering and design communities and with human factors technical specifications being incorporated into numerous engineering standards, many human factors requirements are now met through work routinely conducted by other engineering and design disciplines. So, the question arises whether there is any need for a project to include focused effort by people with higher levels of human factors skills and competence in collaboration with the effort of other disciplines. A human factors screening seeks to answer that question by quickly assessing key aspects of the human factors implications of the introduction of a highly automated system.

Table 2 presents a ‘human factors in automation’ screening tool that could be used by any type of organisation setting out to develop a highly automated system. The tool is intended to be applied from the earliest stage of thinking about the new system.

The tool involves seventeen challenges, based on consideration of the following seven themes[16]:

1. The criticality of the overall system.

2. The impact on the roles of people in the system.

3. Where responsibility for system performance will lie, and how current responsibilities might change.

4. The balance of abilities between the human and automated components of the system.

5. The extent to which the automation is expected to be given authority to perform one or more system tasks.

6. The extent of control of system functions expected to be delivered by people.

7. How the transition from the existing situation to reliance on the automated system will be managed.

---

[16] Note that themes 3,4,5, and 6 are derived from the work of Flemish et al (2012).

LEARNING FROM EXPERIENCE: AIR FRANCE FLIGHT 447: LOSS OF SUPERVISORY CONTROL

On 1 June 2009, Air France Airbus A330-230, flight AF 447 crashed into the Atlantic while en-route from Rio-de Janeiro to Paris. All 228 passengers and crew on-board perished.

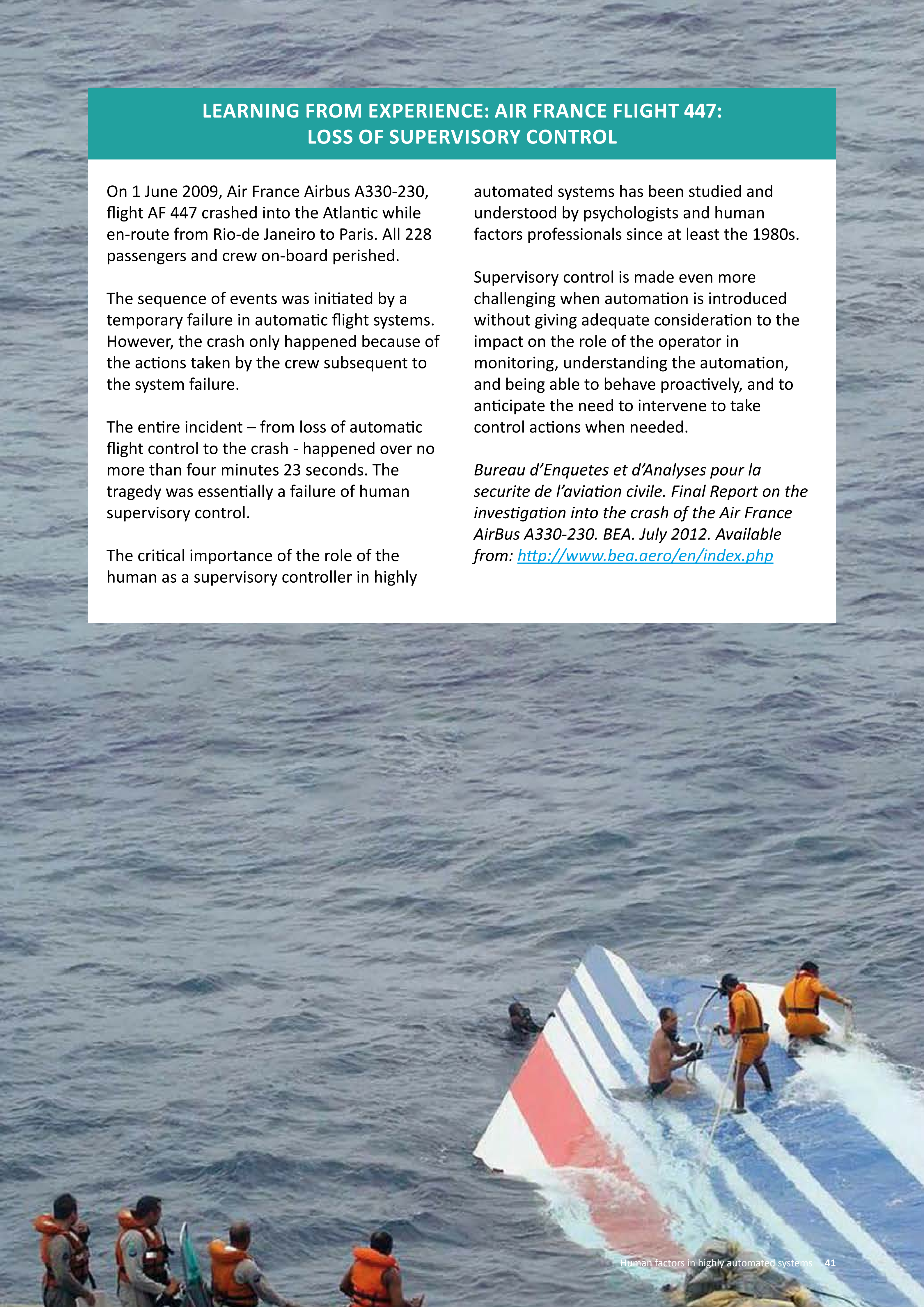
The sequence of events was initiated by a temporary failure in automatic flight systems. However, the crash only happened because of the actions taken by the crew subsequent to the system failure.

The entire incident – from loss of automatic flight control to the crash - happened over no more than four minutes 23 seconds. The tragedy was essentially a failure of human supervisory control.

The critical importance of the role of the human as a supervisory controller in highly automated systems has been studied and understood by psychologists and human factors professionals since at least the 1980s.

Supervisory control is made even more challenging when automation is introduced without giving adequate consideration to the impact on the role of the operator in monitoring, understanding the automation, and being able to behave proactively, and to anticipate the need to intervene to take control actions when needed.

Bureau d'Enquetes et d'Analyses pour la securite de l'aviation civile. Final Report on the investigation into the crash of the Air France AirBus A330-230. BEA. July 2012. Available from: http://www.bea.aero/en/index.php



LEARNING FROM EXPERIENCE: THE BOEING 737-MAX

The loss of two aircraft, Ethiopian Airlines Flight 302 and Lion Air Flight 610, within a short space of time highlighted problems with the design of flight deck automation on the new Boeing 737 MAX.

The project introduced an automated system which had the authority to push the aircraft nose down under certain conditions. As well as failing to fully analyse the potential malfunctioning of the system, Boeing advised airlines buying the aircraft that their pilots would not need any additional training when transferring from older 737 aircraft to the MAX.

The contributing factors have been widely reported and include a major engineering programme driven by financial and time pressures, underpinned by inadequate risk assessment and project controls, and overseen by decision makers who failed to grasp the ramifications of modifying complex digital systems without addressing end user needs.

Pilots flying the 737 MAX were unaware not only of the capabilities of the new automation, but also of its existence. They had received no training on what to do in the circumstances the pilots found themselves in.

The challenges shown in table 2 are intended to promote critical thinking around key aspects of how the proposed automation is likely to impact on people, the relationship between the human and automated components, and how they will need to work together. By encouraging thinking around these issues, the challenges are intended to promote awareness of areas where human factors effort is likely to be needed. The awareness and insight gained by applying the tool will help the organisation sponsoring development of the system understand the kind of human factors activities, and the level of human factors specialism that is likely to be needed to support the design and implementation of that system.

As a rule-of-thumb, the more questions that are answered between ‘not sure’ and ‘definitely applies’, the stronger the case would be for investigation in human factors effort during development and implementation of the system. Where that is the case, a suitably qualified human factors professional[17] should be consulted to interpret the results of the screening and to plan a programme of work.

---

[17] For example, a Registered Member of the CIEHF or equivalent from another relevant professional body

TABLE 2: Human Factors in Automation Screening Challenges

Themes	Challenges	Definitely does not apply		Not sure		Definitely applies	Implications for human factors (where a 'definitely applies', or 'not sure' response is given)
Criticality	The system is considered critical. If it failed to perform as intended, and to the expected standards, there could be serious consequences in terms of health, safety, environmental, financial or other loss.						A suitable strategy for integrating human factors into project management planning should be prepared. Relevant human factors engineering standards should be identified and incorporated into the project baseline. The role of people in preventing incidents should be fully incorporated into planning for barrier management analyses and safety cases[18].
Roles and tasks	Introduction of the system will significantly change, though not completely remove, some of the tasks currently performed by people, or will change how existing tasks are performed (for example by introducing an electronic implementation of what are currently manual or paper-based tasks).						A suitable analysis of operator roles and tasks should be completed as part of system analysis.

---

[18] Such as Layers of Protection Analysis, Safety Integrity Levels, Bowtie Analysis, etc. See 'Human Factors in Barrier Management', CIEHF (2016).

Themes	Challenges	Definitely does not apply		Not sure		Definitely applies	Implications for human factors (where a 'definitely applies', or 'not sure' response is given)
Roles and tasks	People will be expected to be active in monitoring and supporting the automation over extended periods.						The system will need to be designed to ensure people are able to be proactive in understanding what the system is doing. It will not be sufficient to simply rely on alerts or other user prompts. Consideration should be given to whether it is reasonable to expect users to be able to maintain the alertness needed to monitor the system effectively. Consideration should also be given to how the system will keep the user informed of its projections about the effect its intended actions might have.
	The automation will take over some or all of the tasks previously performed by people, but people will still be relied on in the event of degraded system performance.						Analysis will be needed to understand what skills, knowledge and other abilities users will need, and how they will gain or retain them over time.
	Performance of the automation will rely on high levels of human reliability in inspecting, testing, calibrating and maintaining system components.						Change management should assess the impact of the automation on existing maintenance teams and maintenance policies. Task and human reliability analysis should be performed on inspection, testing, calibration and maintenance of critical components. The design of workspaces and components involved in inspection, testing, calibration and maintenance of critical functions must comply with human factors design principles.

Themes	Challenges	Definitely does not apply		Not sure		Definitely applies	Implications for human factors (where a 'definitely applies', or 'not sure' response is given)
Responsibilities	Introduction of the automation will change the balance of responsibilities between the various stakeholders in the system (such as users, their employers, manufacturers, etc.).						Work should be performed to ensure responsibilities for performance of the system between stakeholders are clear, understood and supported by legislation, conditions of work or sale or other agreements. Change management should consider the potential impact of changes in organisational/team roles and responsibilities. Risk assessment should take into account the implications if the human users are not aware of their new or changed responsibilities or are not willing or able to accept them. Analysis will be needed to understand what skills, knowledge and other abilities users will need, and how they will gain or retain them over time.
	Existing stakeholders will be expected to take on new responsibilities beyond those they currently hold. There is uncertainty whether the individuals expected to hold responsibility for the performance of the automated system will understand and accept those responsibilities, or whether they will have the skills, knowledge and experience or situation awareness to be able to accept them all of the time.						

Themes	Challenges	Definitely does not apply		Not sure		Definitely applies	Implications for human factors (where a 'definitely applies', or 'not sure' response is given)
Ability	The technology will not have the ability fully to perform all of the tasks needed to completely meet the system objectives. The automation will rely on people to assist or perform some tasks.						Functional and task analysis should be conducted to understand the allocation of tasks and functions between the automation and human users. The analysis will need to include consideration of how the human and automation will need to communicate and collaborate. Assessment should evaluate the impact on the human ability to cope with the workload and task demands imposed by the automation, especially during the transition to manual control. Risk assessment and failure mode analysis should take into account the potential for impaired collaboration between the automated and human components of the system.
	The automation will not have the ability fully to perform all four core functions (acquiring information, extracting meaning, making decisions, taking action).						
	There could be circumstances where the technology will not be able to detect when it is reaching the limits of its abilities, or if the external environment is reaching the limits the system was designed for.						The limits of the operating design domain (ODD) where the automated is expected to be able to accept authority to control tasks should be very clearly defined in the early stages of system design. Analysis should be conducted to assess how the combined human and technical system is expected to detect when it is approaching the limits of the ODD. Features should be designed into the system that are effective in ensuring the human users are aware sufficiently early if the system is approaching the limits of the ODD. Assumptions about the ability of users to recognise the situation that the system is reaching the limits of the ODD and respond appropriately should be challenged to ensure they are reasonable.

Themes	Challenges	Definitely does not apply		Not sure		Definitely applies	Implications for human factors (where a ‘definitely applies’, or ‘not sure’ response is given)
Authority	The intention is to give the automation full authority to perform all aspects of one or more of the four core functions (acquiring information, extracting meaning, making decisions, taking action) without any human involvement.						Function and task analysis should include assessment of all situations where the human users will be relied on to perform any of the core functions. Risk assessment should consider the implications for human and wider organisational performance and reliability if the system was unexpectedly unable to perform one or more of the core functions it had been given the authority to perform.
	There are foreseeable circumstances where the system might be required to operate in conditions when it may not have the ability fully to perform one or more of the core functions for which it has been given the authority.						
Control	There are foreseeable circumstances when people would be expected to either take over control of the system in real-time, either permanently or temporarily, or to support tasks that would normally be expected to be fully automated. There are foreseeable circumstances where the system may not be able to give advanced warning of the need for people to get involved.						Scenario analysis, supported by function and task analysis, should be conducted to understand the characteristics of situations where the human may need to take over control from the automation. This should also consider how the human can regain control when operating outside the ODD. Risk assessment should consider the implications if the users are not given adequate advance warning of the need to take control. Assumptions made about the ability of the users to remain sufficiently aware, alert and informed to anticipate the need to take control should be challenged to ensure they are reasonable.

Themes	Challenges	Definitely does not apply		Not sure		Definitely applies	Implications for human factors (where a 'definitely applies', or 'not sure' response is given)
Transitioning to automation	Introduction of the automation will represent a significant change to the organisation.						Change management needs to ensure the impact of the change on the human elements of the combined system are fully addressed and understood.
	The automation will need to be introduced as a 'big bang' one-off event. There is no opportunity for a transition period, where changes can be introduced gradually, and there is opportunity to learn and improve as implementation proceeds.						Significant effort will need to be made to ensure the people involved as part of the system are fully prepared for their new roles. A strategy should be prepared to monitor and support the performance of the humans during the introduction of the automation. The strategy should include preparation for fall back to human performance in the event the automation does not perform as expected.

Summary of principles

1. Understand the potential influence of other elements of the system on the automated components, as well as how the introduction of automation can affect those components. Automation must be seen in the context of the overall socio-technical system it exists in.

2. Recognise that automation nearly always changes, rather than removes, the role of people in a system. Those changes are often unintended and unanticipated. They can make the tasks people need to perform more difficult and can disrupt established relationships, lines of communication and the ability to exert authority.

3. Be clear about which of the four core functions (acquiring information, extracting meaning from it, making decisions and taking action) automation will have the ability to perform for each system task, and under what conditions it will be given the authority to control those functions without human oversight.

4. Be realistic in acknowledging that people, at some level, are going to have to monitor, supervise, and hold responsibility for, the performance of the automation. Design, introduce and support the automation such that those people can maintain awareness of the state of both the automation and the world it operates in.

5. Ensure effective, transparent and unambiguous communication between the automation and the human elements of the system, such that the human is able to remain in the loop and situationally aware at all times.

6. For each task or function an automated system has the ability to perform, be as explicit as possible where the balance between authority, responsibility and control lies. Be clear about what the expectations about responsibility imply for the different stakeholders in the system.

7. Ensure the people relied on to support the automation understand what the system is doing and why. There should be no automation surprises.

8. Avoid making unrealistic assumptions about the ability of people to monitor and effectively intervene in any system where there is little for them to do over sustained periods.

9. Recognise that automated systems can increase the levels of task difficulty and workload imposed on the human elements in the system as well as the level of human reliability needed in the inspection, calibration, maintenance and testing of system components.

5. References

Bainbridge, L. (1982). Ironies of Automation. IFAC Proceedings Volumes, 15(6), 129–135. https://doi.org/10.1016/s1474-6670(17)62897-0

BEA (2012) Final Report on the investigation into the crash of the Air France AirBus A330-230. BEA. Bureau d'Enquetes et d'Analyses pour la securite de l'aviation civile. July 2012. Available from: http://www.bea.aero/en/index.php.

Burgess-Limerick, R. (2020) Human-Systems Integration fir the safe implementation of Automation. Mining, Metallurgy & Exploration. https://doi.org/10.1007/s42461-020-00248-z

CAA (2016) ATM Automation: Guidance on human-technology integration. CAP 1377. Safety and Airspace Regulation Group.

CIEHF (2016) Human Factors in Barrier Management. Chartered Institute of Ergonomics & Human Factors.

CIEHF (2019) Learning from Adverse Events: A White Paper. Chartered Institute of Ergonomics & Human Factors.

CIEHF (2020a) The human dimension in tomorrow's aviation system. A White Paper. Chartered Institute of Ergonomics & Human Factors.

CIEHF (2020b) Understanding misuse of partially automated vehicles – A discussion pf NTSB's findings of the 2018 Mountain View Tesla crash. Chartered Institute of Ergonomics & Human Factors.

CIEHF (2021) Human Factors and Ergonomics in Healthcare AI. A White Paper. Chartered Institute of Ergonomics & Human Factors.

Cummings, M. (2018). Informing Autonomous System Design Through the Lens of Skill-, Rule-, and Knowledge-Based Behaviors. Journal of Cognitive Engineering and Decision Making, 12(1), 58–61. https://doi.org/10.1177/1555343417736461

Endsley, M. R. (2017). From Here to Autonomy: Lessons Learned from Human-Automation Research. Human Factors, 59(1), 5–27. https://doi.org/10.1177/0018720816681350

Endsley, M. et al (2003) 'Designing for Situation Awareness' Taylor and Francis.

Energy Institute (2020) Human Factors Engineering in Projects. Joint publication with the International Oil and Gas Producers Association (IOGP)

Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A.& Beller, J. (2012) Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative situations. Cogn. Tech. Work. 14:3-18.

Global Mining Guidelines Group (2018) Guideline for the implementation of autonomous systems in mining. Global Mining Guidelines Group.

Law Commission (2022) Automated Vehicles: Joint Report. Law Commission of England and Wales and the Law Commission of Scotland (2022)

McLeod, R.W. and Balfe, N. (2022) A human factors approach for analysing highly automated systems. Ergonomics & Human Factors 2022 Conference, CIEHF.

Mindell, D. (2011) Digital Apollo. MIT Press

MoD (2021) Human Factors Integration for Defence Systems. Issue 2 Def-Stan 00-251. UK Defence Standardisation.

Moray, N. & Inagaki, T. (1999) Laboratory studies of trust between humans and machines in automated systems. Trans. Inst. MC 21(4/5), pp203-211)

National Transportation Safety Board (2018). Collision Between a Sport Utility Vehicle Operating With Partial Driving Automation and a Crash Attenuator Mountain View, California. Washington, DC USA: NTSB. NTSB/HAR-20/01 2020

Norman, D. (2013) The Design of Everyday Things. MIT Press.

Norman D. (2011) Living with Complexity. MIT Press

Parasuraman, R. & Riley, V. (1997) Humans and Automation: Use, Misuse, Disuse, Abuse Human Factors, 39(2).

Parasuraman, R., Sheridan, T.B., Wickens, C.D. (2000) ‘A Model for Types and Levels of Human Interaction with Automation’ IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans, Vol 30, No. 3, May.

RAIB (2019) Doors open in traffic between Shenfield and Hockley stations, 22 August 2019. Safety Digest. D10/2019. Rail Accidents Investigation Branch

SAE (2018) Surface Vehicle Recommended Practice: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. J3016. SAE International.

Senders, J. (1964) The human operator as a monitor and controller of multidegree of freedom systems. IEEE Transactions on Human Factors in Electronics. Sept.

Woods, D., and Sarter, N.B. (2000) Learning from Automation Surprises and “Going Sour” Accidents. In N. B. Sarter & R. Amalberti (Eds.), Cognitive engineering in the aviation domain (pp. 327–353). Lawrence Erlbaum Associates Publishers.

Sheridan, T. B., Verplank, W. L. 1978. Human and computer control of undersea teleoperators. Technical Report, Man-Machine Systems Laboratory, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge. Mass.

Chartered Institute
of Ergonomics
& Human Factors

www.ergonomics.org.uk

ciehf@ergonomics.org.uk

Designed for CIEHF by Connect Communications – connectmedia.cc

The Chartered Institute of Ergonomics & Human Factors (CIEHF) received its Royal Charter in 2014 to recognise the uniqueness and value of the scientific discipline and the pre-eminent role of the Institute in representing both the discipline and the profession in the UK. This includes the protected status of “Chartered Ergonomist and Human Factors Specialist” with the post-nominal C.ErgHF awarded to practising Registered Members/Fellows who are among a group of elite professionals working at a world-class level.