The third age of human factors: From independence to interdependence

Erik Hollnagel

THE THIRD AGE OF HUMAN FACTORS: FROM INDEPENDENCE TO INTERDEPENDENCE

Erik Hollnagel
MINES ParisTech, Crisis and Risk Research Centre, Sophia Antipolis, France

Since its beginning in the mid 1940s, human factors has tried to keep up with the ever increasing demands from technological and societal developments. Looking back, the development of human factors can be described as corresponding to three ages. In the first age, humans were seen as too too imprecise, variable, and slow to allow the full use of the technological potential. In the second age, humans were seen as failure prone and unreliable, hence a challenge to system safety. In both ages, the human was treated as an entity, as a part that could be described independently of the whole. In the third age, humans are recognised as being necessary if work systems are to be safe and productive. Human performance variability is accepted as the necessary basis for effectively coping with the complexity of the work situations and system performance is understood as the non-trivial result of interdependent parts.

1 Introduction

Human factors is today widely accepted as an essential part of industries in practically every domain. The motivation can, of course, vary and the reasons can be as different as regulator demands, safety concerns, efficiency issues, ergonomics, social considerations. competitiveness, etc. The understanding of what human factors actually means is as diverse as the motivations and can range from genuinely acknowledging the need to address human factors issues in system design and operation, to reluctantly accepting it as something necessary to placate a stakeholder.

While human factors by no means is a novel phenomenon, its history is not as long as it might have been. Human factors is often defined as the science of understanding and applying the properties of human capability to the design and development of technology-based systems and services. From that definition it would be reasonable to expect that human factors has existed as long as humans have used technology. That is, however, far from being the case.

Technology has played a role in human life for thousands of years, and the concern for how to design useful artefacts can be found as far back as in ancient Greece, 25 centuries ago (Marmaras et al., 1999). Technology, however, developed relatively slowly until the Industrial Revolution in the second half of the 18th Century. This brought about a fundamental change in the nature of work, specifically that machines became an integral part of work. In the first stage machines were a source of power only, but technological ingenuity soon expanded their role to control and regulation, thereby enabling machines to become independent of humans at the same time as humans became dependent on machines. The industrial revolution thus introduced what we today know as the human-machine system.

Machines were initially used in the mining and manufacturing industries (the first cotton factory driven by steam opened in Manchester in 1789), but the railways soon followed when the first inter-city passenger railway opened between Liverpool and Manchester on 15 September 1830. A train is surely an example of a human-machine system, and one might therefore have expected that human factors was a concern already then. Yet it would take more than a century and one more revolution before human factors became an issue and a separate discipline. By the beginning of the 20th Century, technological developments had brought about a proliferation of sources of power. The result was that work became dependent on many different types of machines, and that systems became larger and more complex. Existing types of work became more specialized and completely new types of work began to appear. This led to the development of large and complicated processes that required specialized training of humans and therefore also specialized work analysis. The clearest example of that is the discipline of Scientific Management (Taylor, 1911). The basic idea of Scientific Management was to analyse tasks to determine the most efficient performance and then select people to achieve the best possible match between task requirements and capabilities.

Although Scientific Management in many ways meet the common definition of human factors, the consideration of human capabilities in the design of work were limited to physical strength and endurance, i.e., to the mechanical aspects of human work. The human was seen as a component in the work process, but not yet as a factor that could influence the work process as a whole. In particular, it was not thought necessary to take psychological issues into account. This changed after the meta-technological developments in the 1940s that gave rise to what today is known as the *information technology revolution*. These developments included the digital computer (ENIAC in 1945), the formulation of the mathematical theory of communication (1949) and cybernetics (1948), the invention of the transistor (1947-48), followed ten years later by the invention of the integrated circuit. The introduction of information technology in work meant that machine capabilities – hence the demands of humans to control them – quickly exceeded what humans could naturally do. This created what is now called the demand-capacity gap and led to the development of human factors engineering as we know it today.

2 The First Age: The Human Factor as a Bottleneck

The information technology revolution created work situations where humans appeared as too imprecise, variable, and slow. Human capacity limitations, in performance and control, were seen as the reason why system performance (e.g., productivity, precision or speed), was below what the technologies made possible. The three main solutions that human factors engineering developed to overcome these limitations were training, design, and automation. *Training*, supported by selection, was used to bridge the gap between what people in general were able to do and the skills. knowledge, and proficiency required effectively to work with the machines or technology. As technology became more sophisticated, training often took longer and longer - in extreme cases several years, e.g., for aircraft pilots or nuclear power plant operators. *Design* was used to ensure a good fit between the system and the users, first with regard to the basic ergonomics (anthropometric characteristics) such as force, posture, reach, size, and shape, and later with regard to the psychological or cognitive characteristics. Today this is very much an issue of display and interaction design, but until the mid 1980s most human-machine interfaces were based on conventional 'knobs-and-dials' technology. Design also covered other issues such as ease of use, comfort, productive, safety, and aesthetics. *Automation*, finally, used technology itself to overcome the problems created by technology, either by directly replacing humans by automation or by compensating for human weaknesses by 'smart' technology. This created an uncomfortable dependency that remains to this day.

2.1 Consequences

Training, design, and automation all required that clear and detailed descriptions of the activity in question were available. Task analysis therefore soon became a *sine qua non* for human factors, not least for how to structure the human-machine interaction (e.g., Miller, 1953). Task analysis made it natural to think of systems and events as being composed of discrete and identifiable components, and therefore to focus on the characteristics of these components.. .

Although the primary concern in the first age of human factors was to ensure the effective use of technology, human factors addressed both the quality of work and the quality of working. During the first age technology was analogue rather than digital, and both vertical and horizontal integration was limited. It was therefore both natural and appropriate to consider problems very much by themselves, as independent rather than as interdependent.

Even after human factors entered the second age, cf., below, the fundamental concerns remained and were perhaps accentuated by the transition to digital technology. Interface design, for instance, became very much an issue of information presentation and control. The continued technological developments made it possible to extend system integration, and also to automate more and more functions. This slowly changed the nature of work from tracking and regulating to monitoring and targeting, with consequences for what human factors was expected to deliver.

3 The Second Age: The Human Factor as a Liability

The second age of human factors was introduced rather abruptly by the accident at the Three Mile Island (TMI) nuclear power plant on March 28, 1979. It was clear even from the first descriptions of the accident that operator actions had played a significant role in how the events developed, hence in bringing about the final outcome. This pointed to the necessity of considering the effects of human actions, and in particular the effects of incorrect human actions, in the design and operation of human-machine systems.

At first, the established means of training, design, and automation were applied. *Training* was used to teach humans not only how to do something but also how to think about it. This introduced a focus on the cognitive functions involved in, e.g., diagnosis, problem solving, decision making, and planning. *Design* was used to ensure that people could cope with an increasing flow of data and a potential information input overload, and also that they could perform the required control actions. *Automation* almost became a panacea in the sense that it was used wherever possible to take over the tasks and activities that humans were unable to accomplished or had failed in doing correctly. Yet it soon became clear that automation could create more problems than it solved (Bainbridge, 1983)

The concern with the human as a liability led to a focus on “human error.” The need to model “human errors,” both quantitatively and qualitative, resulted in a large number of methods, taxonomies, and models. One of the most important of these was the skill-based, rule-based, knowledge-based framework proposed by Rasmussen (1986). In addition to designing systems to overcome or bypass human capacity limitations, it also became necessary to ensure that human errors were either prevented or that steps were taken to limit or contain adverse outcomes. This led to a focus on human error identification, often in terms of human error quantification, and on human error reduction, for instance by finding ways to enhance human reliability.

While humans initially were described as "fallible" machines, it gradually became obvious that humans differed significantly from technological artefacts, and in particular from information processing systems, in the sense that humans neither worked with very specific inputs and outputs, nor were limited to one or a few functions. Technical systems and components are designed and built to perform with little or no variability, until they have to be replaced. Humans cannot naturally provide the same constant performance, or even be drilled to do so, nor should they rightly be expected to.

3.1 Consequences

The primary concern in the second age of human factors was to ensure that "human errors" were reduced or eliminated, and that the consequences of such "errors" could be contained. The first goal led to an extensive use of automation, and to the "mechanisation" of many simple cognitive functions and tasks. The explanations for "human errors" were initially sought in the human mind and often described as faulty information processing, but after some years it became clear that performance failures could be a product of the working conditions as well as of "human errors." The second goal was pursued through the design of barriers of many types, not least barrier functions such as interlocks or automated recovery. Other means for error prevention were interface and interaction design combined with training, and stricter compliance with procedures and prescriptions in some cases supported by warning or monitoring technologies. The very idea of the "human error" sustained the view of humans as independent rather than interdependent.

4 The Third Age: The Human Factor as an Asset

While the second age of human factors was introduced rather abruptly, the transition to the third age was less obvious. The relentless development of socio-technical systems gradually created situations where the established ways of thinking were powerless. Case after case made it clear that it was impossible to ensure the required safe and efficient system performance by a combination of overcoming human capability limitations and eliminating "human errors." This led to the realisation that humans were not just bottlenecks or possible sources of error, but also the resource that enabled increasingly large and incomprehensible socio-technical systems to function both efficiently and safely.

4.1 Tractability and Intractability

In order to control or manage systems and organisations it is necessary to know what goes on "inside" them. It is therefore important that a sufficiently clear description or specification of the system and its functions can be provided. This requirements must, for instance, be met in order for a system to be analysed, in order for specific tools and solutions to be designed, in order for its risks to be assessed, and in order for safety to be managed. That this must be so is obvious if we consider the opposite. If we do not have a clear description or specification of a system, and/or if we do not know what goes on "inside" it, then it is impossible effectively to control it, to design for it, or to make a risk assessment. These qualities are captured by making a distinction between tractable and intractable systems, cf., Table 1 below.

	Tractable system	Intractable system
Number of details	Description are simple with few details	Description are elaborate with many details
Comprehensibility	Principles of functioning are known	Principles of functioning are partly unknown
Stability	System does not change while being described	System changes before description is completed
Relation to other systems	Independence	Interdependence
Metaphor	Clockwork	Teamwork

Table 1: Tractable and intractable systems

Intractable systems are underspecified in the sense that details may be missing or unavailable (e.g., Clarke, 2000). If a system is underspecified it is clearly not possible to provide precise procedures or instructions. On the contrary, the people working in the system, be it at the sharp end or at the blunt end, must be able to use the available procedures in situations that differ from what was assumed. In other words, it is necessary that people are able to vary or adjust what they do, to ensure that the system functions as required and can achieves its operational goals.

4.2 Performance Variability

A highly regular or constant performance is the ideal for machines and technology. This was also the hidden assumption behind the first and second ages of human factors. Performance variability, in the form of habitual and/or intentional adjustments made during actual work, is nevertheless necessary because performance conditions as a rule are underspecified. Performance variability is more often a strength than a liability, and is probably the primary reason why socio-technical systems work as well as they do. Humans are extremely adept at finding effective ways of overcoming problems at work, and this capability is crucial for both safety and productivity. Human performance can therefore at the same time both enhance and detract from system safety.

There are also other reasons why human performance cannot be constant or invariable:

- Physiological and/or fundamental psychological characteristics (e.g., affecting perception and vigilance).
- Higher level psychological phenomena such as ingenuity, creativity, and adaptability.
- Organizationally induced performance variability, as in meeting external demands (quality, quantity), stretching resources, substituting goals, etc.
- Socially induced variability, as in meeting expectations of oneself or of colleagues, complying with informal work standards, etc.
- Contextually induced performance variability, for instance if the working conditions are too hot, too noisy, too humid, etc.

- Performance variability induced by the unpredictability of the domain, e.g., weather conditions, number of flights, pilot variability, technical problems, etc.

Performance variability is usually noticed on the individual level, where it often, incorrectly, has been labelled "human error." But it may equally well occur collectively, i.e., in the performance of groups and organisations. In these cases it is less easy to observe, and therefore only noticed when it leads to undesired outcomes. While such outcomes may be attributed to organisational failures or accidents, they are in fact due to the interdependence among the performance of individuals. This interdependence comes about because performance adjustments in the long run can be efficient only if the behaviour of others is predictable.

5 The ETTO Principle

Human behaviour - and human performance - can be described as if it was guided by or followed a principle of efficiency-thoroughness trade-off (the ETTO principle; Hollnagel, 2009). In all that they do people are faced with the problem of how to be both efficient and thorough at the same time - or rather, how to be sufficiently efficient while being acceptably thorough. As one might expect, efficiency typically dominates thoroughness, in the sense that people trade off or sacrifice thoroughness for efficiency. They obviously never do that to the extent that they take unnecessary risks or are unsafe, at least in their own understanding of the situation. (It may possibly look different to others.)

In its simplest possible form, the ETTO principle can be stated as follows: In their daily activities, at work or at leisure, people routinely make a choice between being efficient and being thorough, since it rarely is possible to be both at the same time. If demands to productivity or performance are high, thoroughness is reduced until the productivity goals are met. If demands to safety are high, efficiency is reduced until the safety goals are met.

Efficiency means that the level of investment or amount of resources, including time, used or needed to achieve a stated goal or objective are kept as low as possible. For individuals, the decision about how much effort to spend is usually not conscious, but rather a result of habit, social norms, and established practice. For organisations, it is more likely to be the result of a direct consideration - although that choice in itself will also be subject to the ETTO principle. Thoroughness means that an activity is carried out only if the individual or organisation is confident that the necessary and sufficient conditions for it exist, so that the activity will achieve its objective and not create any unwanted side-effects. More formally, thoroughness means that the pre-conditions for an activity are in place, that the execution conditions can be ensured, and that the outcome(s) will be the intended one(s).

5.1 Consequences

The third age of human factors recognises that our socio-technical systems are complex, and that individual and collective behaviour therefore cannot be understood in isolation. All socio-technical systems are underspecified, and underspecified systems can only function if performance is variable, i.e., if it is adjusted to the current conditions. This means that individual human performance depends on what happens in the work environment, hence that there is a mutual dependency. The mission of human factors is therefore not just to compensate for limited capabilities or to reduce or eliminate error. The mission is instead to understand the nature of the interdependence, and to develop ways to sustain or strengthen the individual and collective performance variability that leads to desired or improved outcomes as well as to dampen the performance variability that may put safety at risk.

References

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775-779.

Clarke, S. G. (2000). Safety culture: Underspecified and overrated? *International Journal of Management Reviews*, *2*(1), 65-90.

Hollnagel, E. (2009). *The ETTO principle: Why things that go right sometimes go wrong*. Aldershot, UK: Ashgate.

Marmaras, N., Poulakakis, G. & Papakostopoulos, V. (1999). Ergonomic design in ancient Greece. *Applied Ergonomics*, *30*(4),361-368.

Miller, R. B. (1953). *A method for man-machine task analysis*. Dayton, Ohio: Wright AF Development Center.

Rasmussen, J. (1986). *Information processing and human-machine interaction*. New York: North-Holland.

Taylor, F. W. (1911). *The principles of scientific management*. New York: Harper.