State of science

Citation for published version:
Read, GJM, Shorrock, S, Walker, GH & Salmon, PM 2021, 'State of science: evolving perspectives on ‘human error'', *Ergonomics*, vol. 64, no. 9, pp. 1091-1114. https://doi.org/10.1080/00140139.2021.1953615

Digital Object Identifier (DOI):
10.1080/00140139.2021.1953615

Published In:
Ergonomics

Read, G. J. M., Shorrock, S., Walker, G. H. & Salmon, P. M. (In press). State of Science: Evolving Perspectives on 'Human Error', *Ergonomics* [accepted for publication 2 July 2021] https://doi.org/10.1080/00140139.2021.1953615

State of Science: Evolving Perspectives on 'Human Error'

Read, G. J. M.[1], Shorrock, S.[1,2], Walker, G. H.[1,3], & Salmon, P. M.[1]

1. Centre for Human Factors and Sociotechnical Systems, University of the Sunshine Coast, Maroochydore, Australia

2. EUROCONTROL, Brussels, Belgium

3. Centre for Sustainable Road Freight, Heriot-Watt University, Edinburgh, UK

Abstract

This paper reviews the key perspectives on human error and analyses the core theories and methods developed and applied over the last 60 years. These theories and methods have sought to improve our understanding of what human error is, and how and why it occurs, to facilitate the prediction of errors and use these insights to support safer work and societal systems. Yet, while this area of Ergonomics and Human Factors (EHF) has been influential and long standing, the benefits of the 'human error approach' to understanding accidents and optimising system performance have been questioned. This state of science review analyses the construct of human error within EHF. It then discusses the key conceptual difficulties the construct faces in an era of systems EHF. Finally, a way forward is proposed to prompt further discussion within the EHF community.

Practitioner statement

This state-of-science review discusses the evolution of perspectives on human error as well as trends in the theories and methods applied to understand, prevent and mitigate error. It concludes that, although a useful contribution has been made, we must move beyond a focus on individual error to systems failure to understand and optimise whole systems.

Abbreviations: CWA: cognitive work analysis; EAST: event analysis of systemic teamwork; EHF: ergonomics and human factors; FRAM: functional analysis resonance method; HEI: human error identification; HFACS: human factors analysis and classification system; HRA: human reliability analysis; OHS: occupational health and safety; Net-HARMS: networked hazard analysis and risk management system; SHERPA: systematic human error reduction and prediction approach; STAMP: systems theoretic accident model and processes

Keywords: human error, accident analysis, systems thinking, complex systems, future of ergonomics

1. Introduction

Many of us, particularly those working within the safety critical industries, received a fundamental education in human error models (e.g. Rasmussen, 1982; Reason, 1990; 1997) and methods (e.g. Kirwan, 1992a; 1992b). The simplicity of the term 'human error' can be a blessing. We have likely explained our role to those outside the discipline using this term. Indeed, the familiarity of human error within everyday language may have facilitated buy-in for the importance of ergonomics and human factors (EHF)[1], but its simplicity may also be a curse (Shorrock, 2013), with unintended consequences for safety and justice, and for the EHF discipline generally.

Currently EHF finds itself within a shift that is changing the nature of many long-standing concepts, introducing subtleties that are less easy to explain to clients of EHF services, the media, the justice system, and the public. The movement from the 'old view' to 'new view' of human error proposed by Dekker (2006) and from Safety-I to Safety-II (Hollnagel et al., 2013; Hollnagel, 2014) considers many of these challenges, and introduces newer concepts to safety management. Whether considered a paradigm shift (e.g. Provan, Woods, Dekker & Rae, 2020) or more of an evolution in thinking, recent discourse has challenged the practical usefulness of human error methods and theories (Salmon et al., 2017). Underpinning this is a fundamental change in focus in EHF from analysing human-technology interactions to a broader, more holistic form of thinking that acknowledges various aspects of complexity science (Dekker, 2011; Salmon et al., 2017; Walker et al., 2010).

Sociotechnical systems comprise human and technical components that work together to achieve a common goal. In complex sociotechnical systems, outcomes (e.g. behaviours, accidents, successes) emerge from the interactions between multiple system components (i.e. humans and technologies). These interactions are dynamic, non-linear (i.e. the strength of a cause is not equivalent to its effect) and non-deterministic (i.e. uncertain and difficult to predict). People act locally, without knowledge of the system as a whole; thus, different perspectives and worldviews exist. Importantly, complex sociotechnical systems are generally open to their environment, and must respond and adapt to environmental changes. These aspects differentiate complex systems from merely complicated systems, which are more closed to their environment and have component relationships which can be analysed accurately. Many traditional engineered artefacts can be conceptualised as complicated systems, such as a jumbo jet or an automobile. These systems can be reduced to their component parts, analysed, and then re-assembled to the whole. However, once a social element (i.e. human

[1] Note, the terms ergonomics and human factors can be used interchangeably, and the discipline is also known by HFE.

interaction) becomes a part of the system boundary, the system becomes complex (Cilliers, 1998; Dekker, Cilliers & Hofmeyr, 2011). Complex systems are indivisible, and therefore the system must be the unit of analysis (Ottino, 2003). Ackoff (1973) described how we must move away from 'machine age' views of the world which assume systems are complicated and can thus be treated in a reductionist manner (i.e. broken into constituent parts, analysed and reassembled to the whole). The reverse of this is systems thinking – a way of thinking of the world in systems, emphasising interactions and relationships, multiple perspectives, and patterns of cause and effect. Here, the system is the unit of analysis and component behaviour should only be considered within the context of the whole. A key implication of systems thinking is that accidents cannot be attributed to the behaviour of an individual component in a complex system (i.e. a human error), instead we must examine how interactions between components failed; that is, how the system itself failed. It is worth noting here, in using the term 'system failure', it is acknowledged that systems themselves only function; outcomes are defined as successes or failures from the perspective of human stakeholders (i.e. whether or not stakeholder purposes and expectations of the system are met).

There have been many influential works outlining this case for change (e.g. Rasmussen, 1997; Leveson, 2004; Dekker, 2002; Hollnagel, 2014). While these debates began over thirty years ago (Senders & Moray, 1991) they remain unresolved. Whilst systems thinking perspectives are experiencing something of a resurgence in EHF (Salmon et al., 2017), this perspective is yet to flow through to the media or justice systems (e.g. Gantt & Shorrock, 2017). This creates a difficult situation in which ambiguous messages from the EHF community surrounding the validity and utility of human error are potentially damaging to the discipline (Shorrock, 2013). This state of science review is therefore timely. This paper aims to summarise the history and current state of human error research, critically evaluate the role of human error in modern EHF research and practice, summarise the arguments for a shift to systems thinking approaches, and provide recommendations for EHF researchers and practitioners to take the discipline forward.

1.1. Consequences of 'human error'

Human error continues to be cited as the cause of most accidents (Woods et al., 2010). Emerging in the 1970s as a focus of accident investigations following disasters such as Three Mile Island and Tenerife, human error explanations came to supplement the previous engineering-led focus on equipment and technical failures in investigations (Reason, 2008). At present, it is commonly stated across the safety critical domains that human failure causes most accidents (in the range of 50-90% depending on domain; e.g. Baybutt, 2002; Guo & Sun, 2020; Shappell & Wiegmann, 1996). The United States National Highway Transport Safety Agency (NHTSA; 2018) assigns 94% of road crashes to the driver, interpreted as “94% of serious accidents are caused by human error” (e.g. Rushe, 2019). In aviation, human error is implicated in the majority of the UK Civil Aviation Authority’s (CAA) ‘significant seven’ accident causes (CAA, 2011), and it is estimated that medical error is the third leading cause of death in the USA (Makary et al., 2016). It is no surprise that figures like these are used to justify the replacement of humans with automation, or the introduction of stricter behavioural controls (i.e. rules or procedures) with punishment to deter non-compliance. However, despite (and sometimes because of) safety controls such as automation, behavioural controls and punishments, accidents still occur. Indeed, we have reached a point in many domains where the decline in accident rates in more developed countries has now plateaued (e.g. in aviation, Weigmann & Shappell, 2003; road, Department of Transport, 2017; NHTSA, 2019; and rail, Walker & Strathie, 2016). The traditional view of human-error has taken us so far, and in doing so, has increasingly exposed the systemic nature of the errors ‘left over’ along with the limitations of existing models and methods (Leveson, 2011; Salmon et al., 2017).

1.2. Origins and use of the term

The human tendency to attribute causation to the actions of another human appears to be part of our nature (Reason, 1990). Yet, efforts to translate this common attribution into a scientific construct have proved difficult. To provide context for the modern usage of the concept of error, we begin by exploring its history in common language and its early development as a scientific construct.

The term ‘error’ has a long history. The latin *errorem*, from which error derives, meant "a wandering, straying, a going astray; meandering; doubt, uncertainty"; also "a figurative going astray, mistake"[2]. Around 1300, the middle English *errour* meant, among other things "deviation from truth, wisdom, good judgement, sound practice or accuracy made through ignorance or inadvertence; something unwise, incorrect or mistaken” or an "offense against morality or justice; transgression, wrong-doing, sin” (Kurath, 1953/1989). The term ‘accident’ is derived from the latin *cadere*, meaning “to fall”[3] which has similarities to the meaning of errorem in that they both imply movement from some objective ‘correct path’. This requires a judgment to be made as to what the ‘correct path’ is and the nature of any transgression from it. Judgement is therefore key. It is “the ability to make considered decisions or come to sensible conclusions” or “a misfortune or calamity viewed as a divine punishment”[4]. Judgement is the noun while ‘to be judged’ is the verb, and herein lies the etymology of human error’s relationship to ‘blame’: blame connotes ‘responsibility’, ‘condemnation’ or even

[2] From https://www.etymonline.com/word/error [accessed 6 June 2020]
[3] From https://www.etymonline.com/word/accident
[4] From https://www.encyclopedia.com/social-sciences-and-law/law/law/judgment [accessed 6 June 2020]

‘damnation’. Individual responsibility is fundamental to Western criminal law (Horlick-Jones, 1996) and the overlapping relationships between error, accident, judgement and blame play out regularly in legal judgements and popular discourse. Something that Horlick-Jones refers to as a ‘blamist’ approach.

Human error is used more precisely, and more extensively, within the EHF literature. This is shown in Figure 1 which is based on a keyword search (for ‘human error’ and ‘error’) within the titles, abstracts and keywords of articles in the top three EHF journals (based on impact factor), and the wider literature indexed by Scopus. Figure 1 shows that the term human error began to be used in the mid-late 1960s, which aligns to the establishment of specialised EFH journals (e.g. *Ergonomics* in 1957). The use of the term has increased somewhat over time within EHF journals but has increased much more sharply within the wider academic literature identified in Scopus. Clearly this is a topic that continues to engage writers, both positively and negatively.

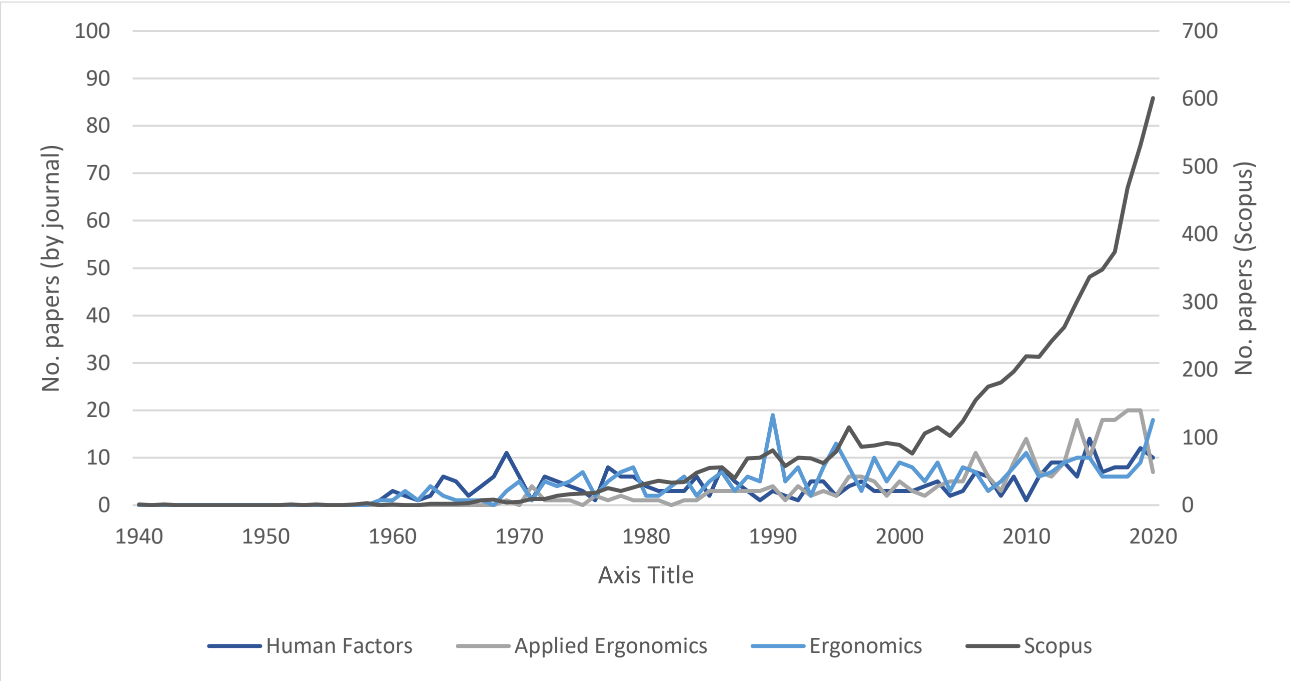


Figure 1. Trends in the use of the term “human error” across journal articles from 1940 to 2020, comparing: Human Factors, Applied Ergonomics, Ergonomics and all records in Scopus (note, Scopus figure refer to secondary y axis due to higher numbers overall).

Figure 1 Alt Text: A graph showing trends between the year 1940 and 2020. All trend lines are close to zero until approximately 1960. Between 1960 and 2020, trend lines for the journals Human Factors, Applied Ergonomics, and Ergonomics appear to rise somewhat. Between 1960 and 2020, the trend line for all records in Scopus rises more sharply.

1.3. The emergence of human error as a scientific concept

The scientific origins of human error begin in the 1940s and span three key time periods.

1.3.1. 1900 to World War II

Human error was a topic of interest to the followers of Freud and psychodynamics, to the behaviourists, and those from the Gestalt school of psychology. Freud saw the unconscious as playing a role in behaviour (e.g. the eponymous 'Freudian slip'; Freud & Brill, 1914). In early psychophysics research the causes of error were not studied per se, but errors were used as an objectively observable measure of performance (Amalberti, 2001; Green & Swets, 1966). Behaviourism research shared an interest in observable indicators of errors (Watson, 1913), with some limited interest in phenomena such as negative transfer of training (e.g. Singleton, 1973). Errors of perception were a common subject of study for those within the Gestalt school (Wehner and Stadler, 1994). Systematic errors in interpreting visual illusions, for example, can be seen as representing early human error 'mechanisms' (Reason, 1990). An interesting exception to the more dominant behaviourism and psychophysical schools was Bartlett's (1932) schema theory, which focused on the role of internal scripts in creating behaviour. To an extent, Bartlett's work presaged the coming the 'cognitive revolution'.

1.3.2. World War II – A turning point

The equipment and technologies deployed during World War II created an imperative to understand and address human error. In 1942 a young psychology graduate, Alphonse Chapanis, joined the Army Air Force Aero Medical Lab as their first psychologist. He studied the controls of the Boeing B-17, an aircraft which had been over-represented in crash landings. Chapanis identified that the flaps and landing gear had identical switches which were co-located and operated in sequence. He determined that during the high-workload period of landing, pilots were retracting the landing gear instead of the flaps. Chapanis solved the design issue by installing a small rubber wheel to the landing gear lever and a small wedge-shape to the flap lever in what we would now call analogical mapping (Gentnor, 1983) or shape coding. Fitts and Jones (1947) built on Chapanis' work by showing that many other so-called 'pilot errors' (quotation marks were used by Fitts and Janes) were in fact problems of cockpit design: *"Practically all pilots of present day AAF aircraft, regardless of experience or skill, report that they sometimes make errors in using cockpit controls. The frequency of these errors and therefore the incidence of aircraft accidents can be reduced substantially by designing and locating controls in accordance with human requirements"* (p.2). As a result of this pioneering work, not only did these (so-called) pilot errors virtually disappear, but more formalised work into human error was initiated.

1.3.3. Post-World War II to 1980s

Following World War II, the language and metaphors of new fields of enquiry such as cybernetics and computing found expression in new concepts of human error in the emerging field of cognitive psychology. Information processing models, such as Broadbent's (1958) stage model of attention, made explicit the idea that different cognitive information processing units perform different functions which enable humans to process information from the environment and to act on it. One of the first information processing-based approaches to human error was Payne and Altman's (1962) taxonomy. This categorised errors associated with sensation or perception as 'input errors'; errors associated with information processing as 'mediation errors', and errors associated with physical responses as 'output errors'. Not all errors, therefore, were the same.

The pioneering work of Chapanis and Fitts also revealed that those committing an error were often as perplexed as the psychologists as to why it occurred. This required an increase in attention to the psychological underpinnings of error and this, in turn, started to challenge 'rational actor' or utility theories tacit in human error thinking to this point, first proposed by Daniel Bernoulli in the 1700s and popular in economics around the 1950s (Tversky, 1975). The rationality of decisions or actions must be defined from the local perspective of the person acting in the particular situation, taking into account their knowledge, their goals and the environmental constraints under which they are operating (Gibson, 1979; Woods & Cook, 1999). This led to interest in local or bounded rationality (Simon, 1957). Far from seeking to rationally optimise an outcome, humans were observed to operate as 'satisficers' (Simon, 1956) who make use of 'fast and frugal' heuristics within an adaptive toolbox of strategies (Gigerenzer, 2001). Rasmussen and Jensen (1974) demonstrated the benefits of studying normal performance and adaptability during real-world problem-solving noting that the processes employed by people quite often differ from what "the great interest of psychologists in complex, rational problem solving leads one to expect" (Rasmussen & Jensen, 1974, p. 7). In another early paper, while still focused on the categorisations of errors, Buck (1963) demonstrated the benefits of observing normal performance in train driving, as opposed to review of failures alone.

1.3.4. 1980s onwards

In the early 1980s, Norman proposed the Activation-Trigger-Schema model of error (Norman, 1981), drawing on Bartlett's schema theory from the pre-WWII era. Only a few years later, as the cognitive systems engineering field began to grow, the human error approach began to be questioned. The

NATO Conference on Human Error organised by Neville Moray and John Senders provided a key forum for discussion. A position paper submitted by Woods (1983) called for the need to look 'behind human error' and to consider how design leads to 'system-induced errors' (Weiner, 1977) rather than 'human errors'. Hollnagel (1983) questioned the existence of human error as a phenomenon and called for a focus instead on understanding decision making and action in a way that accounts for performance variability. In the 1990's these views gathered support but were far from mainstream (we will return to them later) and human error remained largely in step with a rational view on human behaviour. Pheasant (1991), for example, defined error as "an incorrect belief or an incorrect action" (p. 181), and Sanders and McCormick (1993) referred to error as "an inappropriate or undesirable human decision or behaviour (p. 658)". Clearly there was a growing tension between a deterministic view of error and one grounded in the role of bounded rationality and the environment within which errors occur, as exemplified by concepts such as situated planning (Suchman, 1987) and distributed cognition (Hutchins, 1995).

Once the environmental and other systemic factors were admitted into the causation of errors, the natural next step was to focus on the dynamic aspects of complex systems within which errors take place. Rasmussen's (1997) model of migration proposed that behaviour within a system is variable within a core set of system constraints, with behaviours adapting in line with gradients toward efficiency (influenced by management pressure) and effort (influenced by individual preferences), eventually migrating over time towards the boundaries of unacceptable performance. More recently, resilience engineering has emerged to consider "the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions" (Hollnagel, 2014). This development led to a re-branding: Safety-I thinking (i.e. focus on preventing accidents and incidents) to Safety-II (understanding of everyday functioning and how things usually 'go right'). While it has been emphasised that these views are complementary rather than conflicting, Safety-II advocates a much stronger focus on normal performance variability within a system, especially at the higher levels (e.g. government, regulators) who traditionally take a Safety-I view (Hollnagel, Leonhardt, Licu & Shorrock, 2013).

From the etymology of the word error, the growing reference to human error in scientific literature, and the origins of different theoretical bases, it is clear human error – as a concept - is central to the discipline, yet not fully resolved. The origins of the concept in EHF rapidly alighted on the fact that human error is very often 'design induced error', and there has been a tension ever since between that and a more mechanistic, colloquial view of error and 'blamism'. This tension permeates the current state-of-science. We next consider how human error is viewed from different perspectives in a more general way, including by practitioners outside of the EHF discipline. We propose a broad set of perspectives on human error and its role in safety and accident causation (acknowledging that safety is not the only relevant context for discussions of human error). Then, we explore how 'human error' has been defined, modelled, and analysed through the application of EHF methods.

2. Perspectives, models and methods for understanding human error

2.1. Perspectives on safety and human error

In this section we propose a set of four perspectives to synthesise our understanding of human error: the mechanistic perspective, individual perspective, interactionist perspective and systems perspective. The key aspects of each perspective are summarised in Table 1.

The perspectives somewhat represent the evolution of safety management practices over time. While the concept of human error and blame has been prevalent in society throughout history, formal safety approaches such as accident investigation commenced from an engineering perspective (Reason, 2008), followed by the introduction of psychology and EHF, and later the adoption of systems theory and complexity science within EHF. The perspectives could also be seen to fit along could be viewed as fitting along a continuum between Dekker's (2006) 'old view' and 'new view' of human error with the mechanistic and individual perspectives representing the old view, the interactionalist view tending towards the new view, and the systems perspective representing the new view.

2.1.1. The mechanistic perspective

The mechanistic perspective focuses on technology and views human behaviour in a deterministic manner. Underpinned by engineering principles and Newtonian science, this perspective suggests that human behaviour can be predicted with some certainty and that the reliability of human failure can be calculated. A reductionist view, the mechanistic perspective takes a micro view and aligns with Safety-I thinking in relation to preventing failures. It tends to view error as a cause of accidents.

2.1.2. The individual perspective

This perspective can be conceptualised as addressing the 'bad applies' (Dekker, 2006) or bad behaviours. An indication that this perspective is in use may be reference to the 'human factor'. This approach is often associated with 'blamism' and is outdated within EHF. It can, however, still be found within safety practice exemplified in some behaviour-based safety approaches and in the education and enforcement interventions commonly applied to address public safety issues. For example, this approach continues to dominate road safety research and practice whereby driver behaviour is seen as the primary cause of road crashes and driver education and enforcement are common intervention strategies (e.g. Salmon et al., 2019).

2.1.3. The interactionist perspective

Generally applied in a Safety-I context, this perspective still views error as the cause of accidents, while acknowledging that contextual and organisational factors can play a role as contributory factors. Sometimes referred to as 'simplistic systems thinking' (Mannion & Braithwaite, 2017) it does consider system influences on behaviour but often in a linear or mechanistic fashion, and limited to the organisational context. Unlike the mechanistic and individual perspectives, the interactionalist perspective does not connote a negative view of humans, often quite the reverse (e.g. Branton's (1916-90) person-centred approach to ergonomics; Osbourne et al., 2012).

2.1.4. The systems perspective

This perspective, underpinned by systems theory and complexity science, takes a broader view of system behaviour across multiple organisations and acknowledges wider societal influences. It can be differentiated from the interactionist perspective in than it takes the system itself as the unit of analysis (often considering elements beyond the boundary of the organisation); it considers non-linear interactions; and it generally views accidents as 'systems failures', rather than adopting the 'error-as-cause' view. Uniquely, this perspective can explain accidents where there is no underlying 'error' but where the normal performance of individuals across levels of the system lead it to shift beyond the boundary of safe operation. Johnston and Harris (2019), discussing the Boeing 737 Max failures explains that "one must also remember that nobody at Boeing wanted to trade human lives for increased profits… Despite individual beliefs and priorities, organizations can make and execute decisions that none of the participants truly want… (p. 10). The systems perspective is more often aligned with Safety-II, with the aim of optimising performance rather than only preventing failure. It is also compatible with human-centred design approaches whereby user needs and perspectives are considered within a broader understanding of system functioning (e.g. Clegg, 2000).

Table 1. Perspectives on safety management and accident causation

Perspective	**Typical conceptualisation**	**Typical unit of analysis**	**Error-as-cause**	**Safety-I or Safety-II**

	of human-in-system behaviour			
Mechanistic perspective	Complicated	Micro – the human	Often	Safety-I
Individual perspective	Complicated	Micro – the human	Often	Safety-I
Interactionist perspective	Complicated	Meso – the human and broader context, sometimes the organisation	Often	Safety-I
Systems perspective	Complex	Macro – the system as unit of analysis	-	Safety-I or -II

2.2 Human performance models

Stemming predominantly from the individual and interactionist EHF perspectives, human performance models (see Table 2) provide conceptual representations of the mechanisms by which errors arise. The models should be interpreted in the context within which they were developed, with models emerging from changes in the state of science. Structural models, such as information processing approaches, align with the individual perspective. These models are focused on cognitive structures and cognition 'in the head'. In contrast, functional models such as the perceptual cycle model (Neisser, 1976), rewritable routines (Baber & Stanton 1997; Stanton & Baber 1996) and the contextual control model (Hollnagel, 1993; Hollnagel & Woods, 2005), align with the interactionist perspective and consider goal-directed interactions between humans and their environment. Human error, phenomenologically, will appear differently depending on the modelling lens through which it is projected. Some models will illuminate certain perspectives more than others and some will have higher predictive validity, in some situations, than others. The model selected, therefore, affects the human error reality.

Table 2. Overview of human performance and human error models; in order of publication.

Model	**Key aspects / description**	**Human error component / types**	**Theoretical underpinning**
Perceptual cycle model (PCM; Neisser, 1976)	- Cyclical model of perception where top-down processing (schemata) and bottom-up processing (information in the world) drive one another - An active schema sets up the expectations of the individual in a particular context - Expectations direct behaviour to seek certain kinds of information and provide a ready means of interpretation - As the environment is sampled, the information updates and modifies the schema, which directs further search	Errors can occur where: - Inappropriate schemas are activated - Schemas are incomplete due to a lack of previous experience (leading to undesirable search strategies or undesirable actions, or both)	- Schema theory (Bartlett, 1932) - Notion that human thought is closely coupled with interaction in the world
Activation-Trigger-Schema (ATS) / Norman-Shallice model (e.g. Norman, 1981)	- Accounts for action slips – events where an action is not executed as intended. Like the PCM, is based on the activation and selection of schemas - A triggering mechanism requires that appropriate conditions be satisfied for the operation of a schema - Accounts for both external sources of schema activation (i.e. from the world) and internal sources (such as thoughts, associations and habits)	Main sources of action slips include: - Slips during formation of intention (errors in classifying the situation or resulting from ambiguous or incompletely specified intentions) - Mode errors (erroneous classification of the situation) - Description errors (ambiguous or incomplete specification of intention) - Unintentional activation of schemas (schemas not part of a current action sequence become activated for extraneous reasons) - Loss of activation or decay of schema - False triggering (a properly activated schema triggered at an inappropriate time) - Failure to trigger (active schema is not invoked)	- Schema theory (Bartlett, 1932) - PCM (Neisser, 1976)
Skill, rule, and knowledge (SRK) framework	Proposes that information is processed at three levels: - Skill-based: Behaviour is controlled by learnt and stored patterns of behaviour. Rapid, effortless, occurs outside conscious control	Three types of errors, based on the information processing levels: - Skill-based errors relate to variability of force, space or time coordination	- General Problem Solver model, based on means-ends analysis (Newell & Simon, 1972)

(Rasmussen, 1982)	- Rule-based: Rules stored in memory applied in setting a plan for action - Knowledge-based: Conscious planning and problem solving to determine appropriate response in unfamiliar or unusual situations Rasmussen's (1974) decision ladder represents the interaction of the 3 levels identifying information processing stages and resultant states of knowledge from the states between becoming alerted of the need to make a decision and action execution	- Rule-based errors involve incorrect classification or recognition of situations, erroneous associations to tasks, or to memory slips in recall of procedures - Knowledge-based errors relate to goal selection Performance shaping factors acknowledged to influence or contribute to errors	
Execution-evaluation cycle model of human information processing (Norman, 1988)	In *The Design of Everyday Things*, Norman proposed seven stages of action associated with interacting with the world (four stages of execution and three stages of evaluation): 1. Forming the goal 2. Forming the intention 3. Specifying an action 4. Executing the action 5. Perceiving the state of the world 6. Interpreting the state of the world 7. Evaluating the outcome	Errors occur where mis-matches exist between the system and the human: - The gulf of execution relates to the extent to which the system or artefact provides opportunities for action that relate to the user's intentions - The gulf of evaluation relates to the extent to which the actual interaction possibilities of the system or artefact fit what user perceived interaction possibilities	- Ecological psychology (Gibson, 1979) - Distributed cognition (Hutchins, 1995)
Generic error modelling system (GEMS; Reason, 1990)	- GEMS provides a way to classify errors at each level of performance defined in Rasmussen's (1982) SRK model, with addition of intentional behaviours (violations) - Errors can occur at each level of performance or are associated with operating at an inappropriate level of performance	Four categories of unsafe acts: - Skill-based errors (i.e. slips and lapses), usually relating to attention - Rule-based mistakes, involving the misapplication of good rules or the application of bad rules - Knowledge-based mistakes, involving incomplete or inaccurate understanding, or cognitive biases - Violations, involving routine violations, exceptional violations or acts of sabotage	- SRK model (Rasmussen, 1982) - Other psychological models and literature (e.g. Tversky & Kahneman, 1974)
Model of Human Information	- Proposes that environmental stimuli are first processed by the short-term sensory store, then move through stages of	Errors can occur within the different stages of processing: - Perception	Cognitive psychology, linear input-output models such as:

Processing (Wickens & Hollands, 1992)	perception, decision and response selection and response execution - Limited attentional resources are applied to support processing, and both short term and working memory influence perception, decision and response - Feedback loops exist from the response to the environment, which affects the stimuli available for processing	- Memory - Decision making - Response execution	- Broadbent's (1958) stage model of attention - Baddeley and Hitch's (1974) model of working memory
Contextual control model (COCOM; Hollnagel, 1993; Hollnagel & Woods, 2005)	- Focuses on the functions that explain performance within a joint cognitive system - Considers how the joint system (human and technology) acts to achieve its goals while at the same time responding to events in the wider environment - Cyclical model shows how the current understanding of the situation (constructs) directs and controls the action taken, and how the action taken is fed back and along with information received to modify current understanding	Three main constituents: - Competence – possible actions or responses that a system can apply to a situation according to recognised needs and demands - Control – orderliness of performance and the way in which competence is applied (e.g. scrambled, opportunistic, tactical, strategic) - Constructs – what the system knows or assumes about the situation in which the action takes place, similar to schemata they are the basis for interpreting information and selection actions	- Neisser's (1976) PCM, extended to incorporate action and control
Rewritable routines (Baber & Stanton 1997; Stanton & Baber 1996)	- For analysing human-technology interactions - Human-machine interactions follow a sequence of states and the human engages in simple action sequences, based on the affordances of the environment, to move from the current state to the next. Only simple planning is undertaken at each stage, rather than planning the entire interaction through to the final goal state - To move between states, the human holds a record of the interaction in working memory which is modified at each state. Because the record is modified at each state, it is considered 'rewritable'	Errors can occur where: - Working memory decays due to the record not being updated or rehearsed within a certain timeframe - User takes action that appears that it will help them progress to a relevant state although it will not - User incorrectly rejects an action that would have helped them progress to a relevant state	- Newall and Simon's (1972) concept of problem solving as moving through a problem space - Suchman's (1987) situated planning - Neisser's (1976) PCM - Baddeley and Hitch's (1974) model of working memory

2.3. Human error methods

Methods are required to support the application of human error models in practice. Two decades ago, Kirwan (1998a) reviewed and analysed 38 approaches to human error identification, and many more have been developed since. This includes human reliability analysis (HRA), human error identification (HEI) and accident analysis methods. We therefore limited our review (see Table 3) to those methods described in Stanton et al (2013) as representing methods that are generally available and in use by EHF practitioners. To gain an understanding of the relative influence of each method, we used Scopus to identify citations and other article metrics for each seminal publication describing a method. While these citation counts are a coarse measure, and tell us nothing about utilisation in practice[5], they provide some insight into the prominence of the methods within the academic literature.

Broadly, the human error methods reviewed fall within two classes: retrospective and prospective; however, a sub-set span both classes.

Retrospective human error analysis methods provide insight into what behaviours contributed to an accident or adverse event, with most methods encouraging the analyst to classify errors (e.g. slip, lapse or mistake) and performance shaping factors (e.g. time pressure, supervisor actions, workplace procedures). The Human Factors Analysis and Classification System (HFACS; Wiegmann & Shappell, 2001) was the only method identified as retrospective only. Originally developed for aviation, HFACS has been adapted for several safety critical domains (e.g. maritime, Chauvin et al., 2013; mining, Patterson & Shappell, 2010; and rail, Madigan et al., 2016). HFACS provides taxonomies of error and failure modes across four organisational levels (unsafe acts, preconditions for unsafe acts, unsafe supervision, and organisational influences).

Prospective human error methods are used to identify all possible error types that may occur during specific tasks so that design remedies can be implemented in advance. Generally applied in conjunction with task analysis, they enable the analyst to systematically identify what errors could be made. The Systematic Human Error Reduction and Prediction Approach (SHERPA; Embrey, 1986, 2014), for example, provides analysts with a taxonomy of behaviour-based ‘external error modes’, along with ‘psychological error mechanisms’, with which to identify credible errors based on operations derived from a task analysis of the process under analysis. Credible errors and relevant performance influencing factors are described and rated for probability and criticality before

[5] Contemporary metrics such as social media engagements and mentions in news articles or blog posts were not available for these publications given most were released some time ago.

suitable remedial measures are identified. Prospective methods can also support HRA, whereby tasks are assessed for the quantitative probabilities of human errors occurring taking into account baseline human error probabilities and relevant performance shaping factors. Examples include the Human Error Assessment and Reduction Technique (HEART; Williams, 1986) and the Cognitive Reliability and Error Analysis Method (CREAM; Hollnagel, 1998). As can be seen in Table 3, CREAM was the most frequently cited of the methods reviewed.

To varying extents, all human error methods rely on an underpinning mechanistic, individual or interactionist perspective. There is no doubt such approaches are useful, but caution is recommended. The accuracy of HRA approaches has been challenged given uncertainties surrounding error probability estimates (Embrey, 1992) and overall methodological reliability and validity (Stanton et al., 2013). It has also been suggested that categorising and counting errors supports a worldview of humans as unreliable system components that need to be controlled, or even replaced by automation (Woods & Hollnagel, 2006). Both Embrey (1992) and Stanton (2006) suggest that methods should focus on the identification of potential errors rather than their quantification, with the benefit being the qualitative insights gained and the opportunities to identify remedial measures which can inform design or re-design. As may be expected, none of the methods reviewed adopt a systems perspective.

Table 3. Human error methods and taxonomies

Method	**Purpose**	**Human error taxonomies applied**	**Theoretical / empirical underpinning**	**Example application areas**	**Total citations**	**Citations last 15 years**	**Citations last 15 years (excl. self-citations)**
Retrospective methods							
Human Factors Analysis and Classification System (HFACS; Weigmann & Shappel, 2001)	Accident analysis	Failure taxonomy with four levels – unsafe acts (errors and violations), preconditions for unsafe acts, unsafe supervision, organisational influences	- Reason's (1991) GEMS model	- Rail (Madigan, Golightly & Madders, 2016) - Aerospace (Alexander, 2019) - Healthcare (Igene & Johnson, 2019)	252	238	228
Prospective methods							
Human error assessment and reduction technique (HEART; Williams, 1986)	HEI and HRA	Generic task types, taxonomy of error-producing conditions and taxonomy of remedial measures	- Engineering reliability analysis and engineering experience - HTA based on theory of goal-directed behaviour	- Chemical tanker maintenance (Akyuz & Celik, 2015) - Process control maintenance (Noroozi, Khan, MacKinnon, Amyotte & Deacon, 2014) - Surgery (Onofrio & Trucco, 2020)	140	106	106
Human Error HAZOP (Kirwan & Ainsworth, 1992)	HEI	Guidewords focused on human error (e.g. not done, less than, more than, repeated)	- Engineering reliability analysis - HTA based on theory of goal-directed behaviour	- Shale gas fracturing (Hu, Zhang, Wang & Tian, 2019) - Chemical storage (Yang, Liu, Wang, Zhao, & Khan, 2019)	n/a[6]	n/a	n/a

[6] No seminal paper could be identified for human error HAZOP.

Human error identification in systems tool (HEIST; Kirwan, 1994)	HEI	Behaviour categories (e.g. activation / deactivation, observation / data collection, identification of system state, goal selection), global PSFs (e.g. time, procedures, task complexity), EEMs, PEMs, error reduction guidelines	- Rasmussen's SKR framework - HTA based on theory of goal-directed behaviour	- Oil and gas control room (Nezhad, Jabbari & Keshavarzi, 2013) - Aviation (Stanton et al., 2010)	12[7]	11	11
TAFEI (Baber & Stanton; 1994)	HEI	None	- General systems theory – interactions between components (humans and machines) rather than the components themselves - Theory of Rewritable Routines (Baber & Stanton 1997; Stanton & Baber 1996) - HTA based on theory of goal-directed behaviour	- Ticket vending machines (Baber & Stanton, 1996) - IT maintenance (Barajas-Bustillos et al., 2019)	50	34	18
Systems for Predicting Human Error and Recovery (SPEAR; CCPS, 1994)	HEI	EEMs (e.g. action errors, checking errors, communication errors) and PEMs (e.g. spatial misorientation, misinterpretation)	- Similar to SHERPA	- None identified	2[8]	2	2
Human Error and Recovery Assessment (HERA) Framework (Kirwan, 1998a; 1998b)	HEI	EEMs and PSFs	- Toolkit approaches to EFE - EEM taxonomy adapted from SHERPA & THERP - Integration of HEIST	- Nuclear power (Kirwan, 1998b)	51	41	41

[7] The citation search was refined to papers also including the words "HEIST" OR "Human error identification in systems tool", given that the seminal publication contains broader content beyond describing the HEIST method.
[8] The citation search was refined to papers also including the words "SPEAR" OR "Systems for Predicting Human Error and Recovery", given that the seminal publication contains broader content beyond describing the SPEAR method.

			- Integration of Human Error HAZOP				
Technique for Human Error Assessment (THEA; Pocock et al., 2001)	HEI and HRA	Types of cognitive failure relating to goals, plans, actions, perception / interpretation / evaluation	- Execution-evaluation cycle of information processing (Norman, 1988) - HTA based on theory of goal-directed behaviour	- Flight deck re-design (Pocock et al., 2001) - Security sensitive computer interfaces (Maxion & Reeder, 2005)	37	26	15
Human Error Template (HET; Stanton et al, 2006)	HEI (predictive)	Checklist of 12 error modes (e.g. failure to execute, task execution incomplete, task executed in wrong direction)	- Error modes selected based on existing HEI methods and study of pilot errors - HTA based on theory of goal-directed behaviour	- Aviation (Stanton et al., 2006) - Industrial manufacturing (Tajdinan & Afshari, 2013)	41	41	25
Systematic Human Error Reduction and Prediction Approach (SHERPA; Embrey, 1986)	HEI (predictive)	EEMs (e.g. action errors, checking errors, communication errors,) and PEMs (e.g. spatial misorientation, misinterpretation) PIFs (e.g. noise, lighting, fatigue)	- PEMs based on Rasmussen's SRK framework - HTA based on theory of goal-directed behaviour	- Aviation (Harris et al., 2005) - Petrochemical processing (Ghasemi, Nasleseraji, Hoseinabadi, & Zare, 2013)	151	135	127
Both retrospective and prospective methods							
Cognitive reliability and error analysis method (CREAM; Hollnagel, 1998)	HEI, HRA and accident analysis	Common performance conditions (similar to PSFs; e.g. adequacy of organisation, work conditions), error modes (phenotypes; e.g. action at wrong time), person-related genotypes (e.g. false observation), technology-related genotypes (e.g. software fault), and organisation-related genotypes (e.g. design failure)	- COCOM model - HTA based on theory of goal-directed behaviour	- Submersible diving (Chen, Fan, Ye & Zhang, 2019) - Maritime (Wu, Yan, Wang & Soares, 2017)	1,149	1,119	1,101

Technique for the Retrospective and Predictive Analysis of Cognitive Errors (TRACEr; Shorrock & Kirwan, 2002)	HEI and accident analysis	Task error, Information, Performance shaping factors, EEMS, Internal error modes, PEMs, Error detection, Error correction	- Information processing model (Wickens & Hollands, 1992) - EEM classification adapted from THERP (Swain & Guttmann, 1983)	- Air traffic management (Shorrock & Kirwan, 2002) - Rail (Baysari, Caponecchia & McIntosh, 2011)	192	186	178

Note, EEM = External error mode; PEM = Psychological error mechanism; PSF = Performance shaping factor; PIF = Performance influencing factor. Citations from Scopus, to end 2020, self-citations and citations from books included unless otherwise noted.

3. Use, misuse and abuse of human error

It is clear the current state-of-science comprises multiple perspectives on human error, multiple theoretical frames, and multiple methods. All have played a role in raising the profile of EHF within the safety critical industries. Similarly, from the perspective of public awareness of EHF, the idea of human error is widespread within media reporting and tends to be used synonymously with EHF (Gantt & Shorrock, 2017). It has been shown that media 'blamism' influences people to more readily agree that culpable individuals deserve punishment and less readily assign responsibility to the wider organisation (Nees, Sharma & Shore, 2020). Indeed, there is a tendency for the media to modify the term to assign blame to the person closest to the event (e.g. pilot error, driver error, nurse error). The unresolved tension between human error and design error leads to unintended consequences. Even using such terms encourages a focus on the 'human factor' rather than the wider system (Shorrock, 2013). To paraphrase Parasuraman & Riley (1997) on automation; the term human error is 'used', 'misused' and 'abused'. A discussion under these headings helps to further explore the current state-of-science.

3.1. Use and utility of human error

3.1.1 An intuitive and meaningful construct

Human error is "intuitively meaningful" (Hollnagel & Amalberti, 2001, p2). It is easy to explain to people outside of the EHF discipline or those new to EHF. Framing actions as errors, being unintentional acts or omissions as opposed to intentional violations, may help to reduce unjust blame. Most people will agree with the phrase "to err is human" (Alexander Pope: 1688-1744), at least in principle. Indeed, trial-and-error is an important part of normal learning processes and errors help keep systems safe by building operator expertise in their management (Amalberti, 2001). Singleton (1973, p. 731) noted that skilled operators make "good use of error" and "are effective because of their supreme characteristics in error correction rather than error avoidance".

3.1.2. Errors and heroic recoveries emerge from the same underlying processes

Errors also facilitate serendipitous innovation (Reason, 2008). For example, had Wilson Greatbatch not installed the wrong size of resistor into his experimental heart rhythm recording device, he would not have noticed that the resulting circuit emitted electrical pulses, and likely would not have been inspired to build the first pacemaker in 1956 (Watts, 2011). Similarly, had Alexander Fleming kept a clean and ordered laboratory, he would not have returned from vacation to find mould growing in a petri dish, leading to the discovery of penicillin (Bennett & Chung, 2001). In line with Safety-II arguments, these events demonstrate that the processes underpinning human error are the same as those which lead to desirable outcomes. Consequently, the elimination of error via automation or strict controls may also act to eliminate human innovation, adaptability, creativity and resilience.

3.1.3. Learning opportunities

Human error models and taxonomies have been an important tool in EHF for decades. They have facilitated learning from near misses and accidents (e.g. Chauvin et al., 2013; Baysari, McIntosh & Wilson, 2008), the design of improved human-machine interfaces (e.g. Rouse & Morris, 1987), and the design of error tolerant systems (e.g. Baber & Stanton, 1994).

3.2. Misuse and abuse of human error

There are several ways in which human error has been misused, abused, or both. These relate to a lack of precision in the construct and the way it is used in language. It also relates to an underlying culture of 'blamism', simplistic explanations of accident causation, a focus on frontline workers, and the implementation of inappropriate fixes.

3.2.1. A lack of precision in theory and use in language

A key scientific criticism of the notion of human error is that it represents a "folk model" of cognition. That is, human error is a non-observable construct, used to make causal inferences, without clarity on the mechanism behind causation (Dekker & Hollnagel, 2004). This notion is supported by the fact that there is more than one theory of human error. Hollnagel and Amalberti (2001) suggest that misuse of human error can be related to it being regarded as a cause, a process or a consequence.

1. **Human error as a cause**: When human error is used as a cause or explanation for an adverse event, it represents a stopping point for the investigation. This hinders learning, with broader factors that played a contributory role in the accident potentially being overlooked. "H*uman error is not an explanation of failure, it demands an explanation*" (Dekker, 2006, p. 68). Framing error as a cause of failure generates several negative outcomes such as reinforcing blame, recommendation of inappropriate countermeasures (i.e. re-training, behaviour modification) and a failure to act on the systemic issues that are the real underlying causes of failure. Reason (2000) provided the analogy of swatting mosquitoes versus draining swamps. We can either swat the mosquitoes (by blaming and re-training individuals); or drain the swamps within which the mosquitoes breed (by addressing systemic factors). However, the actions of those at the frontline remain a focus of official

accident investigations. For example, the investigation into the 2013 train derailment at Santiago de Compostela by Spain's Railway Accident Investigation Commission focused on the error of the train driver (Vizoso, 2018) and the investigation by France's Bureau of Enquiry and Analysis for Civil Aviation Safety into the Air France 447 crash focused on the pilots' failure to control the aircraft (Salmon, Walker & Stanton, 2015).

2. **Human error as a process**: When human error is discussed as a process or an event, the focus is on the error itself, rather than its outcomes. One of the issues with this approach is that error is often defined as a departure from a 'good' process (Woods & Cook, 2012); some 'ground truth' in an objective reality. But this raises further questions such as what standard is applicable and how we account for local rationality. It also raises questions around what it means for other deviations, given that these are relatively common and occur with no adverse consequences (and indeed often lead to successful consequences).
3. **Confounding error with its outcome**: The final usage of human error discussed by Hollnagel and Amalberti (2001) is where error is defined in terms of its consequence, i.e. the accident event. Human error used in this sense is confounded with the harm that has occurred. As an example, although the seminal healthcare report *To Err is Human* (Kohn, Corrigan & Donaldson, 1999) called for systemic change, and was a catalyst for a focus on patient safety, it also confounded medical harm with medical error, thus placing the focus on individual healthcare worker performance. Referring to human error also had the unintended consequence of the healthcare sector believing that they could address the issues being faced, without assistance from other disciplines, and thus failing to keep up with modern developments in safety science (Wears & Sucliffe, 2019).

As Hollnagel and Amalberti (2001) note, for each of the ways that human error is used in language, it carries negative connotations. This is important, because although many people in everyday life and workplaces will acknowledge that "everyone makes mistakes", we are biased cognitively to assume that bad things (i.e. errors and associated adverse outcomes) happen to bad people – the so called 'just world hypothesis' (see Furnham, 2003 for a review; also Reason, 2000). Similarly, the "causality credo" (Hollnagel, 2014) is the idea that for every accident there must be a cause, and generally a bad one (as bad causes precede bad consequences), that these bad causes can be searched back until a 'root cause' (or set of causes) can be identified, and that all accidents are preventable by finding and treating these causes. This mindset reinforces a focus on uncovering 'bad things'. However, as we have seen, contemporary thinking posits that people act in a context of bounded and local rationality, where adaptation and variability are not only common, but in fact necessary to maintain system performance. The contribution of 'normal performance' to accidents is now widely accepted (Dekker, 2011; Perrow, 1984; Salmon et al., 2017).

3.2.2. Blame and simplistic explanations of accidents

The causality credo aligns with the legal tradition of individual responsibility for accidents, where the bad cause is either a crime, or a failure to fulfil a legal duty. Dekker (2011) describes the shift in societal attitudes from the pre-17th century when religion and superstitions would account for misfortunates, to the adoption of what was considered a more scientific and rational concept of an 'accident'. Accidents were seen as "merely a coincidence in space and time with neither human nor divine motivation" (Green, 2003, p. 31). It was not until the advent of large-scale industrial catastrophes such as Three Mile Island in 1979 and Tenerife in 1977 that a risk management approach came to the fore, with the corresponding reduction in the public acceptance of risk to "zero tolerance of failure" (Dekker, 2011, p. 123). This has brought additional pressure to identify a blameworthy actor and ensure they are brought to justice.

There is also an outcome bias whereby blame and criminalisation is more likely when the consequences of an event are more severe (Hendriksen & Kaplan, 2003; Dekker, 2011). Defensive attribution theory (Shaver, 1970) proposes that more blame is attributed in high severity, as opposed to low severity accidents, as a high severity event evokes our self-protective defences against the randomness of accidents. Increased blame then provides a sense of control over the world. Attributing blame following an adverse event, particularly a catastrophic one, restores a sense of trust in 'experts' (who may be individuals or organisations). This repair of trust is needed for the continuation of technological expansion, and suggests a "fundamental, almost primitive, need to blame…" (Horlick-Jones, 1996, pp. 71).

Hindsight also plays an important role (Dekker, 2006). Post-accident, it is easy to fall into the trap of viewing events and conditions leading up the event as linear and deterministic. Hindsight bias in combination with attribution bias helps to strengthen the beliefs of managers, judges, investigators and others who review accidents to "persuade themselves… that they would never have been so thoughtless or reckless" (Hudson, 2014, p 760). Criminal law in Western countries tends to be predicated on the notion that adverse events arise from the actions of rational actors acting freely (Horlick-Jones, 1996). This "creeping determinism" (Fischoff, 1975) also increases confidence in our ability to predict future events, meaning that we may fail to consider causal pathways that have not previously emerged.

3.2.3. Focus on frontline workers

Human error models and methods, almost universally, fail to consider that errors are made by humans at all levels of a system (Dallat, Salmon & Goode, 2019). Rather, they bring the analysts focus to the behaviour of operators and users, particularly those on the 'front-line' or 'sharp-end' such as pilots, control room operators, and drivers. This is inconsistent with contemporary understandings of accident causation, which emphasise the role that the decisions and actions from other actors across the work system play in accident trajectories (Rasmussen, 1997; Salmon et al., 2020). Further, decades of research on safety leadership (Flin & Yule, 2004) and safety climate (Mearns, Whitaker & Flin, 2003) highlight the importance of managerial decisions and actions in creating safe (or unsafe) environments. Nonetheless, it is interesting to note that despite decades of research, our knowledge of human error is still largely limited to frontline workers or users while the nature and prevalence of errors at higher levels of safety critical systems has received much less attention. This may be explained by the fact that the relationship or coupling between behaviour and the outcome affects how responsibility is ascribed (Harvey & Rule, 1978). Key aspects of this include causality and foreseeability (Shaver, 1985). There is often a time lag between decisions or actions made by those away from the frontline and the accident event. Within the intervening time frame, many opportunities exist for other decisions and actions, or circumstances, to change the course of events. Further, where decisions are temporally separated from the event, it is more difficult to foresee the consequences, particularly in complex systems where unintended consequences are not uncommon.

The way in which we define human error may also generate difficulties when considering the role of those at higher levels of the system. Frontline workers generally operate under rules and procedures which provide a normative standard against which their behaviour can be judged. In contrast, designers, managers and such generally operate with more degrees of freedom. It is easier to understand their decisions as involving trade-offs between competing demands, such as a manager unable to employ additional staff due to budgetary constraints or a designer forfeiting functionality to maximise usability. In many ways the copious rules and procedures that we place on frontline workers to constrain their behaviours masks our ability to see their decisions and actions as involving the same sorts of trade-offs (e.g. between cost and quality, efficiency and safety), and helps to reinforce the focus on human error.

Other reasons for the focus on the frontline worker are more pragmatic such as the practical difficulties of identifying specific decisions that occurred long before the event that may not be well-documented or remembered. Even where those involved could be identified, it may be difficult to locate them. Focusing on the frontline worker, or ‘the last person to touch the system’, provides the easy explanation. Simple causal explanations are preferred by managers and by courts (Hudson, 2014), and are less costly than in-depth investigations that can identify systemic failures.

Further, research has shown that simple causal explanations reduce public uncertainty about adverse events, such as school shootings (Namkoong & Henderson, 2014). It is also convenient for an organisation to focus on individual responsibility. This is not only legally desirable but frames the problem in a way that enables the organisation to continue to operate as usual, leaving their structures, culture and power systems intact (Catino, 2008; Wears & Sutcliffe, 2019).

3.2.4. Inappropriate fixes

A human error focus has in practice led to frequent recommendations for inappropriate system fixes or countermeasures. For example, in many industries accident investigations still lead to recommendations focused on frontline workers such as re-training or education around risks (Reason, 1997; Dekker, 2003). Further, the pervasiveness of the hierarchy of control approach in safety engineering and occupational health and safety (OHS) can reinforce the philosophy that humans are a ‘weak point’ in systems and need to be controlled through engineering and administrative mechanisms. This safety philosophy, without appropriate EHF input, can reinforce person-based interventions which attempt to constrain behaviour, such as the addition of new rules and procedures to an already overwhelming set of which no-one has full oversight or understanding. This contrasts with traditional EHF interventions based on human-centred design, improving quality of working life, and promoting worker wellbeing. A contemporary example is the trend towards use of automation in more aspects of our everyday lives, such as driving, underpinned by the argument that humans are inherently unreliable. This embodies the deterministic assumption underlying many approaches to safety and accident prevention which ignores fundamental attributes of complex systems such as non-linearity, emergence, feedback loops and performance variability (Cilliers, 1998; Grant et al., 2018). Thus, fixes are based on inappropriate assumptions of certainty and structure, rather than supporting humans to adapt and cope with complexity. On a related note, the increasing imposition of rules, procedures and technologies adds to complexity and coupling within the system, which can in turn increase opportunities for failure (Perrow, 1984).

4. Systems perspectives and methods in EHF

The appropriate use of human error terms, theories, and methods has unquestionably enabled progress to be made, but the underlying tension between human error and ‘systems failure’ remains unresolved. The misuse and abuse of human error leads to disadvantages which may be slowing progress in safety improvement. In response, perspectives on human error which were considered radical in the 1990's are becoming more mainstream and the current state-of-science is pointing towards a systems approach. Indeed, EHF is regularly conceptualised as a systems discipline (e.g. Dul et al, 2012; Moray, 2000; Wilson, 2014; Salmon et al., 2017). Specifically, when considering failures, it is proposed that we go further than human error, design error (Chapanis, 1947), further even than Weiner's (1977) 'system-induced failure' to instead recognise accidents as 'systems failures'.

Taking a systems approach, whether to examine failure or to support success, requires structured and systematic methods to assist in the practical use of constructs such as systems theory and systems thinking. While methods from the interactionist perspective may be appropriate in some circumstances, for example, in comparing design options or predictive risk assessment, they do not take the whole system as the unit of analysis and consider potential non-linear interactions.

A core set of systems ergonomics methods are now in use (Hulme et al., 2019). These include AcciMap (Svedung & Rasmussen, 2002), the Systems Theoretic Accident Model and Processes (STAMP; Leveson, 2004), Cognitive Work Analysis (CWA; Vicente, 1999), the Event Analysis of Systemic Teamwork (EAST; Walker et al., 2006), the Networked hazard analysis and risk management system (Net-HARMS; Dallat, Salmon & Goode, 2018) and the Functional Analysis Resonance Method (FRAM; Hollnagel, 2012). These methods take into account with the key properties of complex systems (see Table 4). In comparison to human error methods (Table 3), systems methods offer a fundamentally distinct perspective by taking the system as the unit of analysis, rather than commencing with a focus on human behaviour. Consequently, the analyses produced are useful in identifying the conditions or components that might interact to create the types of behaviours that other methods would classify as errors. Importantly, this includes failures but also normal performance. Notably, the boundary of the system of interest must always be defined for a particular purpose and from a particular perspective, and this may be more broad (e.g. societal; Salmon et al., 2019) or more narrow (e.g. comprising an operational team and associated activities, contexts and tools; Stanton, 2014). Boundary definitions will depend upon the level of detail required and a judgement on the strength of external influences but is vital for defining the 'unit of analysis'.

Table 4. Human error defined in relation to complex systems properties

Complex system property	**Description**	**Properties of systems ergonomics methods**

Outcomes emerge from interactions between system components	Interactions between components produce emergent phenomena. These can only be understood by analysing the system as a whole, rather than examining components in isolation (Dekker, 2011; Leveson, 2004)	- A detailed understanding of a human error cannot provide an explanation of an accident, nor can predicting individual behaviours provide an indication of the level of system safety. Analyses should focus on the interactions that lead to behaviour and the situations within which these occur
System and component performance are variable	System components and systems themselves are constantly adapting in response to local pressures and unforeseen disturbances (Hollnagel, 2009). Adaptation and variability ensure survival under ever-changing environmental conditions (Vicente, 1999). The presence of performance variability makes it difficult to predict component behaviour and system performance	- If performance is variable and adaptation is required to make systems work, human error simply describes the unwanted result of trade-offs, which under normal circumstances result in desirable outcomes
Systems are dynamic	Systems involve dynamic processes, such as the transformation of inputs into outputs, and the operation of feedback loops. Through these processes systems evolve over time in response to changing conditions. Tending towards entropy, systems migrate towards a state of increased risk (Rasmussen, 1997; Leveson, 2011) and can drift into failure (Dekker, 2011)	- Human behaviour should be viewed in the context of the dynamic factors and state of the system at the time it occurred
Systems are organised in hierarchical structures	Systems tend to self-organise into hierarchies of systems and sub-systems (Skyttner, 2005). To understand a system, it is necessary to examine each relevant hierarchical level and its relationship with those above and below (Rasmussen, 1997; Vicente, 1999)	- ‘Human errors’ can occur at all levels of a system, from governments to CEOs, to supervisors and frontline workers - Accidents can only be understood by going beyond the immediate work environment to the influences within the management system and broader social and political environment

While systems ergonomics methods have been in use for some time, there continue to be calls for more frequent applications across a range of domains (Hulme et al., 2019; Salmon et al., 2017), from aviation (e.g. Stanton, Li & Harris, 2019), healthcare (Carayon et al, 2014), to nuclear power (e.g. Alvarenga et al, 2014), to addressing the risk of terrorism (Salmon, Carden & Stevens, 2018). They are seeing increasing use in practice as they are better able to cope with the complexity of modern sociotechnical systems. For example, in the context of air traffic management, EUROCONTROL has argued for a move away from simplistic notions of human error, and towards systems thinking and associated concepts and methods, such as those cited in this paper (Shorrock, Leonhardt, Licu & Peters, 2014). Methods such as Net-HARMS (Dallat et al., 2018) have been developed with the express purpose of being used by practitioners as well as researchers.

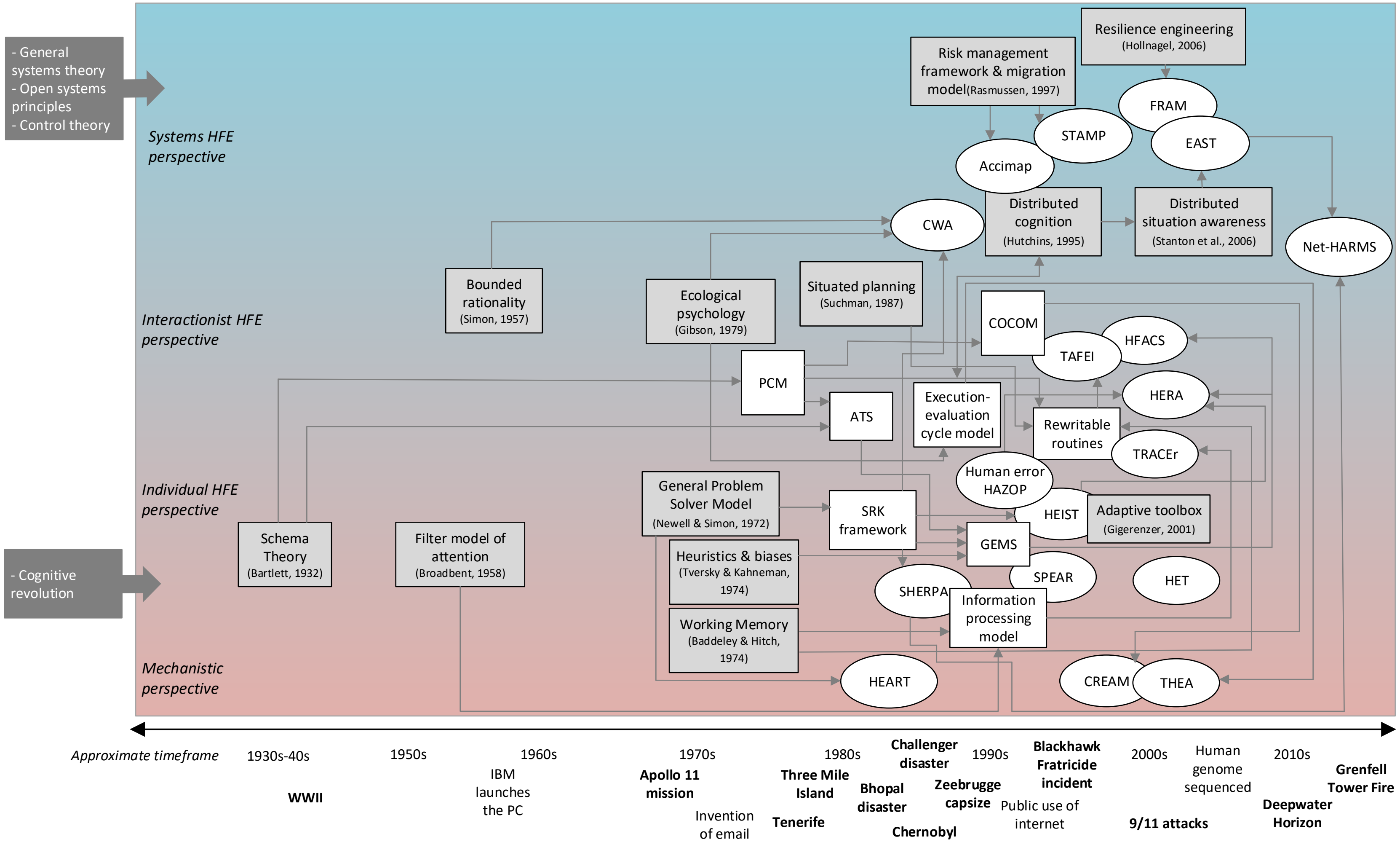
A significant benefit of systems perspectives and methods is to address the issue of blame and simplistic explanations of accidents beyond EHF professionals and the organisations within which they work. As one example, the EUROCONTROL Just Culture Task Force has led to the acknowledgement of just culture in EU regulations (van Dam, Kovacova & Licu, 2019) and delivered, over 15 years, education and training on just culture and the judiciary. This includes workshops and annual conferences that bring together prosecutors, air traffic controllers, pilots, and safety and EHF specialists (Licu, Baumgartner & van Dam, 2013) and include content around systems thinking as well as just culture. Guidance is provided for interfacing with the media, to assist in educating journalists about just culture and ultimately facilitate media reporting into air traffic management incidents that is balanced and non-judgemental (EUROCONTROL, 2008). While it is acknowledged that culture change is a slow process, these activities have a long-term focus with the aim of shifting mindsets over time.

5. A summary of the shift from human performance to systems ergonomics

The rise of the systems perspective in EHF has undoubtedly impacted the popularity and use of human error terms, theories and methods. Figure 2 summarises the development over time of the theories, models and methods that relate to the four perspectives on human error. This highlights the proliferation of model and method development in a relatively short period of time, particularly from the late 1970s to the early 2000s. Hollnagel (2009) attributes this surge in method development, particularly human error methods, to the Three Mile Island nuclear meltdown in 1979. He identifies this at the beginning of the second age of human factors, where the human is viewed as a liability. It is worth highlighting that many of the systems ergonomics methods were developed in parallel to the human error methods, with some notable examples also stemming from the nuclear sector (e.g. CWA). Figure 2 suggests that EHF has spent the past three decades operating under multiple perspectives, rather than necessarily experiencing a paradigm shift. Interestingly, here have been some examples of integration between perspectives, such as the Net-HARMS method drawing from SHERPA, and CWA's adoption of the SRK taxonomy. However, given that EHF is no longer a 'new' discipline, with this journal alone recently celebrating 60 years of publication, it is timely to discuss a clearer way forward.

Figure 2. Overview of the evolution of human performance and systems ergonomics models (*white boxes*), methods (*white ovals*), and underpinning theory (*grey boxes*). Note, alignment against perspectives is intended to be approximate and fuzzy rather than a strict classification.

Figure 2 Alt Text: An illustration of models, methods and theories showing the relationships between them, over time from the 1930s to present day.



6. To err within a system is human: A proposed way forward

The concept of human error has reached a critical juncture. Whilst it continues to be used by researchers and practitioners worldwide, there are increasing questions regarding its utility, validity, and ultimately its relevance given the move towards the systems perspective. We suggest there are three camps existing within EHF (Shorrock, 2013): 1) a group that continues to use the term with 'good intent', arguing that we must continue to talk of error in order to learn from it; 2) a group who continues to use the term for convenience (i.e. when communicating in non-EHF arenas) but rejects the simplistic concept, instead focusing on wider organisational or systemic issues; and 3) a group who have abandoned the term, arguing that the concept lacks clarity and utility and its use is damaging. We predict that this third group will continue to grow. We acknowledge that the concept of error can have value from a psychological point of view, in describing behaviour that departs from an individual's expectation and intention. It might also be considered proactively in system design, however interactionalist methods that focus on all types of performance, rather than errors or failures alone, provide analysts with a more nuanced view. Importantly, however, we have seen how the concept of human error has been misused and abused, particularly associated with an error-as-cause view, leading to unintended consequences including blame and inappropriate fixes.

A set of practical recommendations are offered as a way of moving EHF, and our colleagues within related disciplines, away from a focus on individual, mechanistic, blame-worthy 'human error' and towards a holistic systems conception of system performance. These are shown in Figure 3.

Importantly, we must consider the implications of the proposed changes on other areas of the discipline. A shift away from human error to the systems perspective would fundamentally change how we view and measure constructs such as situation awareness (see Stanton et al., 2017), workload (Salmon et al., 2017) and teamwork, for example. There are also significant implications for areas such as job and work design, where we may see a resurgence in interest in approaches underpinned by sociotechnical systems theory (Clegg, 2000) along with new understandings of how organisations can support workers, for example by promoting agency in responding to uncertainty (Griffin & Grote, 2020). A final note as we draw our discussion to a close, is that a key systems thinking principle relates to maintaining awareness of differing worldviews and perspectives within a complex system. Donella Meadows (1941-2001) highlighted a vital consideration for any discipline when she suggested that the highest leverage point for system change is the power to transcend paradigms. There is no certainty in any worldview, and flexibility in thinking, rather than rigidity, can indeed be the basis for "radical empowerment" (Meadows, 1999, p. 18).

Figure 3. A proposed way forward

Figure 3 Alt text: A list of recommendations for the way forward for the EHF discipline.

- Reject 'human error' as a cause of accidents and adverse events. Focus on how the system failed and interventions to increase the system's capacity to manage disturbances and performance variability
- Reject countermeasures focused on individual behaviour. Advocate for networks of interventions which respond to system-wide issues
- Reject simplistic explanations for accidents. Acknowledge that rationality is bounded and avoid the trap of hindsight
- Acknowledge that humans are assets and problem solvers, operating in imperfect and complex environments, usually with good intentions, and certainly never intend for accidents to occur (by definition)
- Avoid blame-laden terminology. Instead of 'human error' or 'violation', use neutral and factual terms (e.g. decision, action, event, consequence)
- Seek to identify performance variability at the component and system level. Intervene to support positive variability (e.g. adaptation) and reduce negative variability (e.g. drift)
- Continue to apply interactionist HEI methods in a predictive manner to support design, remaining cognisant of wider system influences
- Adopt systems perspectives and embrace systems ergonomics methods. Continue to develop and adapt systems methods to be scalable and usable in practice
- Educate colleagues in other disciplines (e.g. engineering, design, OHS) and broader institutions (e.g. media, the courts, politicians). Introduce complexity science, systems perspectives and systems ergonomics methods
- Incorporate complexity science, systems perspectives and systems ergonomics methods into EHF competency frameworks. Support the next generation of EHF professionals to continue the shift towards systems perspectives

Conclusions

Human error has helped advance our understanding of human behaviour and has provided us with a set of methods that continue to be used to this day. It remains, however, an elusive construct. Its scientific basis, and its use in practice, has been called into question. While its intuitive nature has no doubt assisted EHF to gain buy-in within various industries, its widespread use within and beyond the discipline has resulted in unintended consequences. A recognition that humans only operate as part of wider complex systems leads to the inevitable conclusion that we must move beyond a focus on individual error to systems failure to understand and optimise whole systems. Systems ergonomics theories and methods exist to support this shift, and the current state-of-science points to their uptake increasing. We hope that our proposed way forward provides a point of discussion and stimulates debate for the EHF community as we face new challenges in the increasingly complex sociotechnical systems in which we apply EHF.

Acknowledgements: We sincerely thank Rhona Flin for her insightful comments on a draft version of this paper. We also thank Alison O'Brien and Nicole Liddell for their assistance with proof reading and literature searching. Finally, we are immensely grateful to the anonymous peer reviewers for their valuable comments which have strengthened the paper.

Funding: Gemma Read's contribution to this work was funded through her Australian Research Council (ARC) Discovery Early Career Research Award [DE180101449]. Guy Walker's contribution to this work was funded through the UK Engineering and Physical Sciences Research Council project Centre for Sustainable Road Freight [EP/R035199/1]

References

Ackoff, R. L. (1973). Science in the systems age: Beyond IE, OR and MS. *Operations Research, 21*(3), 661-671.

Akyuz, E., & Celik, M. (2015). A methodological extension to human reliability analysis for cargo tank cleaning operation on board chemical tanker ships. *Safety Science, 75*, 146-155.

Alexander, T. M. (2019). A case based human reliability assessment using HFACS for complex space operations. *Journal of Space Safety Engineering, 6*(1), 53-59.

Alvarenga, M. A. B., Frutuoso e Melo, P. F., & Fonseca, R. A. (2014). A critical review of methods and models for evaluating organizational factors in human reliability analysis. *Progress in Nuclear Energy*, *75*, 25-41.

Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems. *Safety Science, 37*(2), 109-126.

Baber, C., & Stanton, N. A. (1994). Task analysis for error identification: a methodology for designing error-tolerant consumer products, *Ergonomics, 37*(11), 1923-1941.

Baber, C., & Stanton, N. A. (1996). Human error identification techniques applied to public technology: predictions compared with observed use. *Applied Ergonomics, 27*(2), 119-131.

Baber, C., & Stanton, N. A. (1997). Rewritable routines in human interaction with public technology. *International Journal of Cognitive Ergonomics*. *1*(4), 237-249.

Barajas-Bustillos, M. A., Maldonado-Macias, A., Ortiz-Solis, M., Realyvazquez-Vargas, A., & Hernández-Arellano, J. L. (2019). A cognitive analyses for the improvement of service orders in an information technology center: A case of study. *Advances in Intelligent Systems and Computing, 971*, 188-198.

Bartlett, F. C. (1932). *Remembering: a study of experimental and social psychology*. Cambridge: Cambridge University Press.

Baybutt, P. (2002). Layers of protection analysis for human factors (LOPA-HF). *Process Safety Progress 21*(2), 119–129.

Baysari, M.T., McIntosh, A.S., & Wilson, J.R. (2008). Understanding the human factors contribution to railway accidents and incidents in Australia. *Accident Analysis and Prevention, 40*(5), 1750-1757.

Baysari, M. T., Caponecchia, C., & McIntosh, A. S. (2011). A reliability and usability study of TRACEr-RAV: The technique for the retrospective analysis of cognitive errors – For rail, Australian version. *Applied Ergonomics, 42*(6), 852-859.

Bennett, J. W., & Chung, K-T. (2001). Alexander Fleming and the discovery of penicillin. *Advances in Applied Microbiology, 49*, 163-184.

Buck, L. (1963). Errors in the perception of railway signals. *Ergonomics, 6*(2), 181-192.

CAA. (2011). CAA 'Significant Seven' Task Force Reports. West Sussex: UK Civil Aviation Authority.

Carayon, P., Wetterneck, T. B., Rivera-Rodriguez, A. J., Hundt, A. S., Hoonakker, P., Holden, R., & Gurses, A. P. (2014). Human factors systems approach to healthcare quality and patient safety. *Applied Ergonomics*, *45*(1), 14-25.

Catino, M. (2008). A review of literature: individual blame vs. organizational function logics in accident analysis. *Journal of Contingencies and Crisis Management, 16*(1), 53-62.

CCPS (Centre for Chemical Process Safety) (1994). *Guidelines for preventing human error in process safety*. New York: American Institute for Chemical Engineers.

Chauvin, C., Lardjane, S., Morel, G., Clostermann, J.-P., & Langard, B., 2013. Human and organisational factors in maritime accidents: Analysis of collisions at sea using the HFACS. *Accident Analysis and Prevention, 59*, 26-37.

Chen, D., Fan, Y., Ye, C., & Zhang, S. (2019). Human reliability analysis for manned submersible diving process based on CREAM and Bayesian network. *Quality and Reliability Engineering International, 35*(7), 2261-2277.

Cilliers, P. (1998). *Complexity and postmodernism: Understanding complex systems.* London: Routledge.

Clegg, C. W. (2000). Sociotechnical principles for system design. *Applied Ergonomics, 31*, 463-477.

Dallat, C., Salmon, P. M., & Goode, N. (2018). Identifying risks and emergent risks across sociotechnical systems: the NETworked hazard analysis and risk management system (NET-HARMS). *Theoretical Issues in Ergonomics Science, 19*(4), 456-482.

Dallat, C., Salmon, P. M., & Goode, N. (2019). Risky systems versus risky people: To what extent do risk assessment methods consider the systems approach to accident causation? A review of the literature. *Safety Science*, *119*, 266-279.

Dekker, S. (2002). Reconstructing human contributions to accidents: The new view on error and performance. *Journal of Safety Research, 33*, 371-385.

Dekker, S. (2003). The criminalization of human error in aviation and healthcare: A review. *Safety Science, 49*, 121-127.

Dekker, S. (2006). *The field guide to understanding human error*. Aldershot, Ashgate.

Dekker, S. (2011). *Drift into failure: From hunting broken components to understanding complex systems*. Ashgate, Surrey.

Dekker, S. W. A., & Breakey, H. (2016). 'Just culture:' Improving safety by achieving substantive, procedural and restorative justice. *Safety Science, 85*, 187-193.

Dekker, S., Cilliers, P. & Hofmeyr, J-H. (2011). The complexity of failure: Implications of complexity theory for safety investigations. Safety Science, 49(6), 939-945.

Dekker, S. W., & Hollnagel, E. (2004). Human factors and folk models. *Cognition, Technology and Work, 6*, 79-86.

Dul, J., Bruder, R., Buckle, P., Carayon, P., Falzon, P., Marras, W. S., Wilson, J. R., & van der Doelen, B. (2012): A strategy for human factors/ergonomics: developing the discipline and profession. *Ergonomics, 55*(4), 377-395.

Embrey, D. E. (1986) 'SHERPA: a systematic human error reduction and prediction approach' Proceedings of the International Topical Meeting on Advances in Human Factors in Nuclear Power Systems Knoxville, Tennesee.

Embrey, D. E. (1992). *Quantitative and qualitative prediction of human error in safety assessments*. I. Chem. E. Symposium Series No. I30 I. Chem. E., London.

Embrey, D. (2014). SHERPA: A systematic human error reduction and prediction approach to modelling and assessing human reliability in complex tasks. In *22nd European Safety and Reliability annual conference (ESREL 2013)*; CRC Press: Amsterdam, The Netherlands.

Fischoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance, 1(3), 288–299.

Fitts, P. M., & Jones, R. E. (1947). *Analysis of factors contributing to 460 "pilot error" experiences in operating aircraft controls*. Dayton, OH: Aero Medical Laboratory, Air Material Command, Wright-Patterson Air Force Base.

Flin, R., & Yule, S. (2004). Leadership for safety: Industrial experience. *Quality and Safety in Healthcare, 13* (supp. 2), ii45-51.

Freud, S., & Brill, A. A. (1914*). Psychopathology of everyday life*. Macmillan Publishing.

Furnham, A. (2003). Belief in a just world: research progress over the past decade. *Personality and Individual Differences. 34*(5), 795–817.

Gantt, R., & Shorrock, S. (2017). Human factors and ergonomics in the media. In S. Shorrock & C. Williams (Eds.), *Human factors and ergonomics in practice: Improving system performance and human well-being in the real world*. Boca Raton, FL: CRC Press.

Ghasemi, M., Nasleseraji, J., Hoseinabadi, S., & Zare, M. (2013). Application of SHERPA to identify and prevent human errors in control units of petrochemical industry. *International Journal of Occupational Safety and Ergonomics, 19*(2), 203-209.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Hillsdale, NJ: Erlbaum.

Gigerenzer, G. (2001). The adaptive toolbox. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (p. 37–50). The MIT Press.

Grant, E., Salmon, P. M., Stevens, N., Goode, N., & Read, G. J. M. (2018). Back to the Future: What do accident causation models tell us about accident prediction? *Safety Science, 104*, 99-109.

Green, J. (2003). The ultimate challenge for risk technologies: controlling the accidental. In J. Summerton & B. Berner (Eds.), *Constructing risk and safety in technological practice*. Routledge, London. P. 29-42.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.

Griffin, M. A., & Grote, G. (2020). When is more uncertainty better? A model of uncertainty regulation and effectiveness. *Academy of Management Review, 45*(4), 745-765.

Guo Y., & Sun, Y. (2020). Flight safety assessment based on an integrated human reliability quantification approach. *PLOS ONE 15*(4), e0231391.

Harris, D., Stanton, N. A., Marshall, A., Young, M. S., Demagalski, J., & Salmon, P. (2005). Using SHERPA to predict design-induced error on the flight deck. *Aerospace Science and Technology, 9*(6), 525-532.

Harvey, M. D., & Rule, B. G. (1978). Moral evaluations and judgments of responsibility. *Personality and Social Psychology Bulletin, 4,* 583-588.

Henriksen, K., & Kaplan, H. (2003). Hindsight bias, outcome knowledge and adaptive learning. *Quality and Safety in Health Care*, *12*(suppl II), ii46-ii50.

Hollnagel, E. (1983). Human error: Position paper for NATO conference on human error, August 1983, Bellagio, Italy. Available from: https://erikhollnagel.com/

Hollnagel, E. (1993). *Human reliability analysis: Context and control*. London: Academic Press.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method (CREAM)*. Oxford: Elsevier Science Ltd.

Hollnagel, E. (2009). *The ETTO principle: Efficiency-thoroughness trade-off: Why things that go right sometimes go wrong*. Ashgate, Surrey.

Hollnagel, E. (2012). *FRAM: The functional resonance analysis method: Modelling complex socio-technical systems*. Surrey: Ashgate.

Hollnagel, E. (2014). *Safety-I and Safety-II: The past and future of safety management*. Farnham, UK: Ashgate.

Hollnagel, E., & Amalberti, R. (2001). The emperor's new clothes or whatever happened to human error? Invited keynote presentation at 4th International Workshop on Human Error, Safety and System Development. Linköping, June 11–12, 2001.

Hollnagel, E., Leonhardt, J., Licu, T., & Shorrock, S. (2013). *From Safety-I to Safety-II: A white paper*. Brussels: EUROCONTROL.

Hollnagel & Woods. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: Taylor & Francis.

Horlick-Jones, T. (1996). The problem of blame. In C. Hood & D. K. C. Jones (Eds.), *Accident and design: Contemporary debates on risk management*, 61-71.

Hu, J., Zhang, L., Wang, Q., & Tian, B. (2019). A structured hazard identification method of human error for shale gas fracturing operation. *Human and Ecological Risk Assessment, 25*(5), 1189-1206.

Hudson, P. (2014). Accident causation models, management and the law. *Journal of Risk Research, 17*(6), 749-764.

Hulme, A., Stanton, N. A., Walker, G. H., Waterson, P., & Salmon, P. M. (2019). What do applications of systems thinking accident analysis methods tell us about accident causation? A systematic review of applications between 1990 and 2018. *Safety Science, 117*, 164-183.

Hutchins, E. (1995). *Cognition in the Wild*. Bradford: MIT Press.

Igene, O. O., & Johnson, C. (2019). Analysis of medication dosing error related to computerised provider order entry system: A comparison of ECF, HFACS, stamp and AcciMap approaches. *Health Informatics Journal*, *26*(2), 1017-1042.

Johnston, P., & Harris, R. (2019). The Boeing 737 MAX saga: Lessons for software organizations. *Software Quality Professional, 21*(3), 4-12.

Kirwan, B. (1992a). Human error identification in human reliability assessment. Part 1: Overview of approaches. *Applied Ergonomics, 23*(5) 299-318.

Kirwan, B. (1992b). Human error identification in human reliability assessment. Part 2: Detailed comparison of techniques. *Applied Ergonomics, 23*(6), 371-381.

Kirwan, B. (1994). *A guide to practical human reliability assessment.* Taylor and Francis, London.

Kirwan, B. (1998a). Human error identification techniques for risk assessment of high risk systems—Part 2: towards a framework approach. *Applied Ergonomics, 29*(5), 299-318.

Kirwan, B. (1998b). Human error identification techniques for risk assessment of high risk systems—Part 2: towards a framework approach. *Applied Ergonomics, 29*(5), 299-318.

Kirwan, B., & Ainsworth, L. K. (1992). *A guide to task analysis*. London: Taylor & Francis.

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (1999). *To err is human: Building a safer health system*. Institute of Medicine. National Academy Press Washington, DC.

Kurath, H. (1989). *Middle English dictionary*. Ann Arbor, Michigan: The University of Michigan Press. (Original work published 1953).

Leveson, N. (2004). A new accident model for engineering safer systems. *Safety Science, 42*(4), 38-53.

Leveson, N. (2011). Applying systems thinking to analyze and learn from events. *Safety Science, 49*, 55-64.

Licu, T., Baumgartner, M., & van Dam, R. (2013). Everything you always wanted to know about just culture (but were afraid to ask). *Hindsight, 18*, 14-17.

Madigan, R., Golightly, D., & Madders, R. (2016). Application of human factors analysis and classification system (HFACS) to UK rail safety of the line incidents. *Accident Analysis & Prevention, 97*, 122-131.

Makary, M., & Daniel, M. (2016). Medical error—the third leading cause of death in the US. *BMJ*, *353*.

Mannion, R., & Braithwaite, J. (2017). False dawns and new horizons in patient safety research and practice. *International journal of health policy and management, 6*(12), 685–689.

Maxion, R. A., & Reeder, R. W. (2005). Improving user-interface dependability through mitigation of human error. *International Journal of Human-Computer Studies, 63(1-2*), 25-50.

Meadows, D. H. (1999). *Leverage points: Places to intervene in a system*. Hartland, VT: The Sustainability Institute.

Mearns, K., Whitaker, S.M., & Flin, R. (2003). Safety climate, safety management practice and safety performance in offshore environments. *Safety Science, 41*(8), 641-680.

Moray, N. (2000). Culture, politics and ergonomics. *Ergonomics, 43*, 858-868.

Namkoong, J-E., & Henderson, M. D. (2014). It's simple and I know it!: Abstract construals reduce causal uncertainty. *Social Psychology and Personality Science, 5*(3), 352-359.

Nees, M. A., Sharma, N., & Shore, A. (2020). Attributions of accidents to "human error" in news stories: Effects on perceived culpability, perceived preventability, and perceived need for punishment. *Accident Analysis and Prevention, 148*, 105792.

Neisser, U. (1976). *Cognition and reality*. San Francisco: W.H. Freeman.

Nezhad, Z., Jabbari, M., & Keshavarzi, M. (2013). Identification of the human errors in control room operators by application of HEIST method (case study in an oil company). *Iran Occupational Health, 10*(2), 11-23.

NHTSA. (2018). Traffic safety facts: Critical reasons for crashes investigated in the national motor vehicle crash causation survey. DOT HS 812 506, https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812506, accessed 15 July 2019.

NHTSA. (2019). Fatality analysis reporting system (FARS) encyclopedia, https://www-fars.nhtsa.dot.gov/Main/index.aspx, accessed 14 November 2019.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review, 88*(1), 1-15.

Norman, D. A. (1988). *The psychology of everyday things*. New York, NY, US: Basic Books.

Onofrio, R., & Trucco, P. (2020). A methodology for dynamic human reliability analysis in robotic surgery. *Applied Ergonomics*, *88*, 103150.

Ottino, J. M. (2004). Complex Systems. *AIChE Journal, 49*(2). 292-299.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse and abuse. *Human Factors, 39*(2), 230-253.

Patterson, J. M., & Shappell, S. A. (2010). Operator error and system deficiencies: Analysis of 508 mining incidents and accidents from Queensland, Australia using HFACS. *Accident Analysis & Prevention, 42*(4), 1379-1385.

Payne, D., & Altman, J. W. (1962). *An index of electronic equipment operability: Report of development*. Pittsburgh, PA: American Institutes for Research.

Perrow, C. (1984). *Normal accidents: Living with high-risk technologies*. Basic, New York.

Pheasant, S. (1991). *Ergonomics, work and health*. MacMillan Publishers: London.

Pocock, S., Harrison, M., Wright, P. C., & Johnson, P. (2001). THEA: A technique for human error assessment early in design. In M. Hirose (ed.), *Interact01*. Amsterdam: IOS Press.

Provan, D. J., Woods, D. D., Dekker, S. W. A. & Rae, A. J. (2020). Safety II professionals: How resilience engineering can transform safety practice. Reliability Engineering & System Safety, 195, 106740.

Rasmussen, J. (1974). *The human data processor as a system component. Bits and pieces of a model*. Risø-M, No. 1722.

Rasmussen, J. (1982). Human errors. A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents, 4*, 311-333.

Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety Science, 27*, 183-213.

Rasmussen, J. & Jensen, A. (1974). Mental procedures in real-life tasks: A case study of electronic trouble shooting. *Ergonomics*, *17*(3), 293-307.

Reason, J. (1990). *Human error*. New York: Cambridge University Press.

Reason, J. (1997). *Managing the risks of organisational accidents*. Aldershot: Ashgate.

Reason, J. (2000). Human error: Models and management. *BMJ, 320*, 768.

Reason, J. (2008). *The human contribution: Unsafe acts, accidents and heroic recoveries*. London: CRC Press.

Rouse, W. B., & Morris, N. M. (1987). Conceptual design of a human error tolerant interface for complex engineering systems. *Automatica, 23*(2), 231-235.

Rushe, D. (2019, March 10). 'I'm so done with driving': is the robot car revolution finally near? *The Guardian*. https://www.theguardian.com/cities/2019/mar/09/im-so-done-with-driving-is-the-robot-car-revolution-finally-near-waymo [accessed 25 July 2019].

Salmon, P. M., Carden, T., & Stevens, N. (2018). Breaking bad systems: Using work domain analysis to identify strategies for disrupting terrorist cells. *Proceedings of the Ergonomics and Human Factors Society Conference 2018.*

Salmon, P. M., Hulme, A., Walker, G. H., Waterson, P., Berber, E., & Stanton, N. A. (2020). The big picture on accident causation: A review, synthesis and meta-analysis of AcciMap studies, *Safety Science*, 104650.

Salmon, P. M., Read, G. J. M., Beanland, V., Thompson, J., Filtness, A., Hulme, A., McClure, R., Johnston, I. (2019). Bad behaviour or societal failure? Perceptions of the factors contributing to drivers' engagement in the fatal five driving behaviours. *Applied Ergonomics,* 74, 162-171.

Salmon, P. M., Walker, G. H., Read, G. J. M., Goode, N., & Stanton, N. A. (2017). Fitting methods to paradigms: are ergonomics methods fit for systems thinking? *Ergonomics, 60*(2), 194-205.

Salmon, P. M., Walker, G. H., & Stanton, N. A. (2016). Pilot error versus sociotechnical systems failure: a distributed situation awareness analysis of Air France 447. *Theoretical Issues in Ergonomics Science, 17(*1), 64-79.

Sanders, M. S. & McCormick, E. J. (1993). *Human Factors in Engineering and Design* (7th Ed.). Mcgraw-Hill.

Senders, J. W. & Moray, N. P. (1991). *Human error: Cause, prediction, and reduction / analysis and synthesis*. Hillsdale, N.J.: L. Erlbaum Associates.

Shaver, K. G. (1970). Defensive attribution: Effects of severity and relevance on the responsibility assigned for an incident. *Journal of Personality and Social Psychology, 14,* 101-113

Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. New York, NY: Springer-Verlag.

Shorrock, S. T. (2013). 'Human error': The handicap of human factors. *Hindsight, 18*, 32-37.

Shorrock, S. T., & Kirwan, B. (2002). The development and application of a human error identification tool for air traffic control. *Applied Ergonomics, 33*, 319–336.

Shorrock. S., Leonhardt, J., Licu, T., & Peters, C. (2014). *Systems thinking for safety: Ten principles. A white paper*. Brussels: EUROCONTROL.

Skyttner, L. (2005). *General systems theory: Problems, perspectives, practice* (2nd ed.). Hackensack, NJ: World Scientific.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*(3), 129-138.

Simon, H. A. (1957). *Models of man, social and rational*. Oxford, England: Wiley.

Singleton (1973). Theoretical approaches to human error. *Ergonomics, 16*(6), 727-637.

Stanton, N. A. (2014) Representing distributed cognition in complex systems: how a submarine returns to periscope depth. *Ergonomics,* 57(3), 403-418.

Stanton, N. A., & Baber, C. (1996). A systems approach to human error identification. *Safety Science, 22*(1-3), 215-228.

Stanton, N. A., Harris, D., Salmon, P. M., Demagalski, J. M., Marshall, A., Young, M. S., Dekker, S. W. A., & Waldmann, T. (2006). Predicting design induced pilot error using HET (human error template) – A new formal human error identification method for flight decks. *The Aeronautical Journal, 110*(1104), 107-115.

Stanton, N. A., Li, W-C., & Harris, D. (2019). Editorial: Ergonomics and Human Factors in Aviation. *Ergonomics, 62*(2), 131-137.

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2013). *Human factors methods: A practical guide for engineering and design (2nd ed.)*. Aldershot: Ashgate.

Stanton, N. A., Salmon, P. M., Walker, G. H., Salas, E., & Hancock, P. A. (2017). State-of-science: situation awareness in individuals, teams and systems, *Ergonomics,* 60(4), 449-466.

Svedung, I., & Rasmussen, J. (2002). Graphic representation of accident scenarios: mapping system structure and the causation of accidents. *Safety Science, 40*, 397-417.

Swain, A. D. & Guttmann, H. E. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications*. NUREG/CR-1278.

Tajdinan, S., & Afshari, D. (2013). Human errors in Ancoiler device control room of Ahvaz Pipe Mill using SHERPA and HET methods in 1390. *Iran Occupational Health, 10*(3), 69-77.

Tversky, A. (1975). A critique of expected utility theory: Descriptive and normative considerations. *Erkenntnis, 9*(2), 163–173.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Van Dam, R., Kovacova, M., & Licu, T. (2019). The just culture journey in Europe: Looking back and looking forward. *Hindsight, 28*, 47-50.

Vicente, K. J. (1999*). Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Mahwah, NJ: Lawrence Erlbaum Associates.

Vizoso, S. (2018, July 24). 'Five years on, Galicia train crash shines spotlight on security of Spain's rail network'. *El Pais*. https://elpais.com/elpais/2018/07/24/inenglish/1532418197_112660.html [accessed 17 January 2020].

Walker, G. H., Gibson, H., Stanton, N. A., Baber, C., Salmon, P., & Green, D. (2006). Event analysis of systemic teamwork (EAST): a novel integration of ergonomics methods to analyse C4i activity. *Ergonomics, 49*, 12-13.

Walker, G. H., Stanton, N. A., Salmon, P. M., Jenkins, D. P. & Rafferty, L. (2010). Translating concepts of complexity to the field of ergonomics. *Ergonomics, 53*(10), 1175 - 1186.

Walker, G. H., & Strathie, A. (2016). Big data and ergonomics methods: A new paradigm for tackling strategic transport safety risks. *Applied Ergonomics, 53 (Part B)*, 298-311.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*, 158-177.

Watts, G. (2011). Obituary: Wilson Greatbatch. *The Lancet, 378*, (9807), 1912.

Wears, R., & Sutcliffe, K. (2019). *Still not safe: Patient safety and the middle-managing of American medicine*. New York: Oxford University Press.

Wehner, T., & Stadler, M. (1994). The cognitive organization of human errors: A Gestalt theory perspective. *Applied Psychology, 43*(4), 565-583.

Weigmann, D., & Shappell, S. A. (2001). Human error analysis of commercial aviation accidents: Application of the human factors analysis and classification system (HFACS). *Aviation Space and Environmental Medicine, 72*(11), 1006-1016.

Weigmann, D., & Shappell, S. A. (2003). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Aldershot: Ashgate.

Weiner, E. L. (1977). Controlled flight into terrain accidents: System-induced errors. *Human Factors, 19*(2), 171-181.

Wickens, C. D., & Hollands, J. G. (1992). *Engineering psychology and human performance*. New York: Harper Collins.

Williams, J. C. (1986). HEART – A proposed method for assessing and reducing human error. *9th Advances in Reliability Technology Symposium*, University of Bradford.

Wilson, J. R. (2014). Fundamentals of systems ergonomics / human factors. *Applied Ergonomics, 45*(1), 5-13.

Woods, D. D., & Cook, R. I. (1999). Perspectives on human error: Hindsight bias and local rationality. In F. Durso (Ed.), *Handbook of applied cognitive psychology*, New York: Wiley, pp. 141-171.

Woods, D. D. & Cook, R. I. (2012). Mistaking Error. In B. J. Youngberg (Ed.) *Patient safety handbook (second edition)*, Jones and Bartlett, Sudbury MA, pp. 99-110.

Woods, D. D., Dekker, S., Cook, R., Johannesen, L., & Sarter, N. (2010). *Behind human error (2nd Ed.)*. Surrey, UK: Ashgate.

Woods, D.D., & Hollnagel, E. (2006). Prologue: Resilience engineering concepts. In E. Hollnagel, D.D. Woods, & N. Leveson. (Eds.) Resilience engineering: Concepts and precepts (pp. 1-8). Aldershot, UK: Ashgate.

Wu, B., Yan, X., Wang, Y., & Soares, C.G. (2017). An evidential reasoning-based CREAM to human reliability analysis in maritime accident process. *Risk Analysis, 37*(10), 1936-1957.

Yang, P., Liu, Y., Wang, Y., Zhao, D., & Khan, F. (2019). Qualitative and quantitative analysis on human factors under emergency state. *Chemical Engineering Transactions, 77*, 721-726.