Essay-4 human error

Precis: Humans as presumably sentient beings unconditionally feel anxious whenever something unexpected happens, particularly if it also results in unacceptable outcomes. One thousand years ago the Muslim scholar Ibn Hazm (994-1064) wrote that "The chief motive of all human actions is the desire to avoid anxiety." In the 19th century the great German philosopher (Friedrich Nietzsche (1844-1900) elaborated "To trace something unfamiliar back to something familiar is at once a relief, a comfort and a satisfaction, while it also produces a feeling of power. The unfamiliar involves danger, anxiety and care -the fundamental instinct is to get rid of these painful circumstances. First principle – any explanation is better than none at all." In their attempts to find a socially acceptable cause people usually, and involuntarily, make an efficiency-thoroughness trade-off and settle for the cause that represents the default explanation in the current age of safety, cf. Table 1. This is why the preferred cause in the second age of safety became human error. It was first of all convenient, it was also immediately understood and accepted by others. And it finally corroborated with the common experience that we all sometimes unintentiionally act in a way that leads to UUOs.

Professors Andrew Hale and Jan Hovden (Hale & Hovden,1998) defined the second age of safety thinking as the age of human factors(Figure 1 and Table 1 in the third essay).

The original version of the domino model (Heinrich, 1959) included Fault of Person as the second domino piece, defined as in Figure 1 and Table 1.

Heinrich provided a description ofthe five factors or steps in an accident sequence:Second essay Table 2. The second factor "fault of person" was desceibed as follows:"Inherited or acquired faults of a person; such as recklessness, violent temper, nervousness, excitability, inconsiderateness, ignorance of safe practice, etc., constitute proximate reasons for committing unsafe acts or for the existence of mechanical or physical hazards."Today this would be considered psychological weaknesses, rather than human error in the current fashion, this is clear if it is compared e.g., to the <generic Error modelling System '(GEMS) derscribed below.

Ever since human factors were adopted by safety studies as described in the third essay, there has been a bias in the preference for using human error to at least partly explain why work and actions that usually result in EAOs, occasionally, but very rarely produce UUOs. The bias is the assumption that there exists something called human error about which meaningful questions can be asked - and answered! human error thereby gets the status of a concrete phenomenon similar to, for instance, decision making. It is, however, obvious that human error does not refer to something observable, in the way that decision making does. Decision denotes the activity of making a decision as well as the assumed cognitive functions behind that. There are many theories of decision making, ranging from normative decision theory (*homo economicus*) to various descriptive theories, such as Simon's *satisficing*, (Simon, 19xx) Lindblom's *muddling through* (Lindblom, 19xx) and Klein's *naturalistic decision making* (Klein, 1993) as well as mixed approaches (Tversky, 1972). (The activity of making a decision is observable in an everyday meaning of the term. The underrlying cognitive functons must, of couse, be inferred, but since everyone has tried to make a decision, the inference is not very difficult to make. Lindblom (1959) made a distinction between two approaches as seen in Table 2.

Table 1: Two completely different approaches to decision making (Lindblom, 1959, p. 81)

Rational-comprehensive (root) approach	**Successive Limited Comparisons (Branch) approach**
Clarification of values or objectives distinct from and usually prerequisite to empirical analysis of alternative policies.	Selection of value goals and empirical analysis of the needed action are not distinct from one another but are closely intertwined.
Policy-formulation is therefore approached through means-end analysis: First the ends arc isolated, then the means to achieve them are sought.	Since means and ends are not distinct, means-end analysis is often inappropriate or limited.
The test of a "good" policy is that it can be shown to be the most appropriate means to desired ends.	The test of a "good" policy is typically that various analysts find themselves directly agreeing on a policy (without their agreeing that it is the most appropriate means to an agreed objective).
Analysis is comprehensive; every important relevant factor is taken into account.	Analysis is drastically limited: • Important possible outcomes are neglected. • Important alternative potential policies are neglected. • Important affected values are neglected.
Theory is often heavily relied upon.	A succession of comparisons greatly reduces or eliminates reliance on theory.

Compared to decision making, there are practically no theories of human error, and there obviously cannot be a normative theory about the ideal way to make a human error.There are, of course quite a few descriptive theories Best known of these is, without a doubt James Reason's terminological distinction between mistakes, defined as deficiencies or failures in the judgmental and/or inferential processes in the selection of an objective or in the specification of the means to achieve it, irrespective of whether or not the actions directed by this decision-scheme run according to plan. (Reason, 1990) additionally distinguished between slips and lapses, and violations as illustrated by Figure 2. These terminological proposals do, however, not constitue an articulated theory of human error as such. result from a failure in the execution and/or storage stage of an action sequence, regardless of whether or not the plan which guided them was adequate to achieve its objective.where the plan or intention to act is inadequate, even if the action goes exactly as intended and performance deficiencies, where the plan is satisfactory, but the execution fails. With further distinctions between slips, lapses, and violations. These terminological proposals do, however, not constitue an articulated theory of human error. Reason's terminology was later combined with the Skill-bases, Rule-based, Knowledge-based (SRK) framework developed by Rasmussen (19xx), resulting in the following typology

- Skill-based (SB): slips and lapses, usually errors of inattention or misplaced attention
- Rule-based (RB): mistakes, usually a result of picking an inappropriate rule caused by misconstrued view of state, over-zealous pattern matching, frequency gambling, deficient rules
- Knowledge-based (KB): mistakes, due to incomplete/inaccurate understanding of system, confirmation bias, overconfidence, cognitive strain.

The SRK framework, is however itself without any theorerical foundation, so although the combination of the SRK framework with Reason's error types was called a Generic Error modelling System (GEMS), it is actually not even a (descriptive) model of human error, and cannot replace a theory either, impressively looking flowcharts notwithstanding.

Phenotypes and genotypes

The distinction between error phenotypes and possible error genotypes is more systematic, yet not a theory as such. The phenotypes can bebased on the experience from FMEA and FMECA, an therefore have an undisputed empirical basis, even though it is still far from being a theory, let alone a model of human error. somewhat more systematic treatment of human error, yet not a theory as such. What (Reason, 1990) described, the acronym GEMS notwithstanding was neither a theory or model of human error, but essentially a systematic classfication system Figure 1 does not in any way explain how human errors happen, but just how they can or should be categorised. Violations have recent yearws attracted renewed interest, but now go by the name of non-compliance.

Positive and negative connotations.

Most of the terminology used to describe the occurrence of human error unsurprisingly have negative connotations, cf. Heinrich. Terms such as violations, deviations, and non-compliance are all negatively loaded. This is in good agreement with the safety legacy. Human error is mainly used to explain UUOs, and since UUOs have a negative valence, it is only to be expected that the same goes for their putative causes. This is indeed one of the seven biases of safety thinking (Hollnagel, 2026).

Reason used the negatively loaded term violations. But this raises an interesting psychological problem. If a person intentionally tries to do something that differs from rules, procedures, or safetystandards, then to do it is exactlky what the person had in mind and can theregore not in itself be considered an error.

“violations are actions taken by ‘bad’ people.” (Alper & Karsh, 2009, p. 1). A violation definitely has a negative connotation, in the same way as deviation. The very names imply, that such acts were intentional, done on purpose, hence done by ‘bad’ people. Replacing violations with non-compliance offered a few additional nuances, as shown in Table 3.

The distinction between genotype and phenotypes, makes it possible to categorise both the ways in which the UUOs can show themselves and the possible underlying reasons, hence a rudimentary theory of human error. One example seen in Table 2.

Table 2: Phenotypes with generic and specific genotypes (Hollnagel, 1998).

Phenotype	Generic genotype		Specific genotype
Timing / duration	Communicatiphon failure	Inadequate procedure	Earlier omission
	Faulty diagnosis	Inattention	Trapping error
	Inadequate plan	Observation missed	Trapping error
Sequence	Access limitations	Inadequate procedure	
	Communication failure	Inattention	
	Faulty diagnosis	Memory failure	
	Inadequate plan	Wrong identification	
Force	Communication failure		Ambiguous label
	Communication failure		Convention conflict
	Inadequate plan		Incorrect label
Distance / Magnitude	Communication failure		Ambiguous label
	Communication failure		Convention conflict
	Faulty diagnosis		Incorrect label
Speed	Communication failure	Inadequate plan	None defined
	Distraction	Inadequate procedure	
	Equipment failure	Observation missed	
	Faulty diagnosis		
Direction	Communication failure	Inadequate procedure	Ambiguous label Convention conflict Incorrect label
	Faulty diagnosis	Inattention	
	Inadequate plan	Observation missed	
Wrong object	Access problems	Inadequate procedure	Ambiguous label Incorrect label
	Communication failure	Inattention	
	Wrong identification	Performance variability	
	Inadequate plan	Observation missed	

Table 3: Categories of non-compliance (Hudson et al., (2008)

Type of non-compliance	Description/example
Unintended forms of non-compliance	
Unintended Understanding failure	when people have a different understanding of what the procedure is and what they have to do.
Unintended Awareness failure	when people are unaware of the existence of a rule or procedure and therefore work without any reference to it.
Intended forms of non-compliance	
Situational non-compliance	when the situation makes it impossible to do the job and be compliant, e.g., because of insufficient time or resources.
Optimizing non-compliance for company benefit	individuals take short-cuts believing that this will achieve what they believe the company, and their managers, really want;
Optimizing non-compliance for personal benefit	short-cuts taken to achieve purely personal goals;
Exceptional non-compliance	short-cuts if the formal procedures is too difficult to follow under specific, and usually novel, circumstances

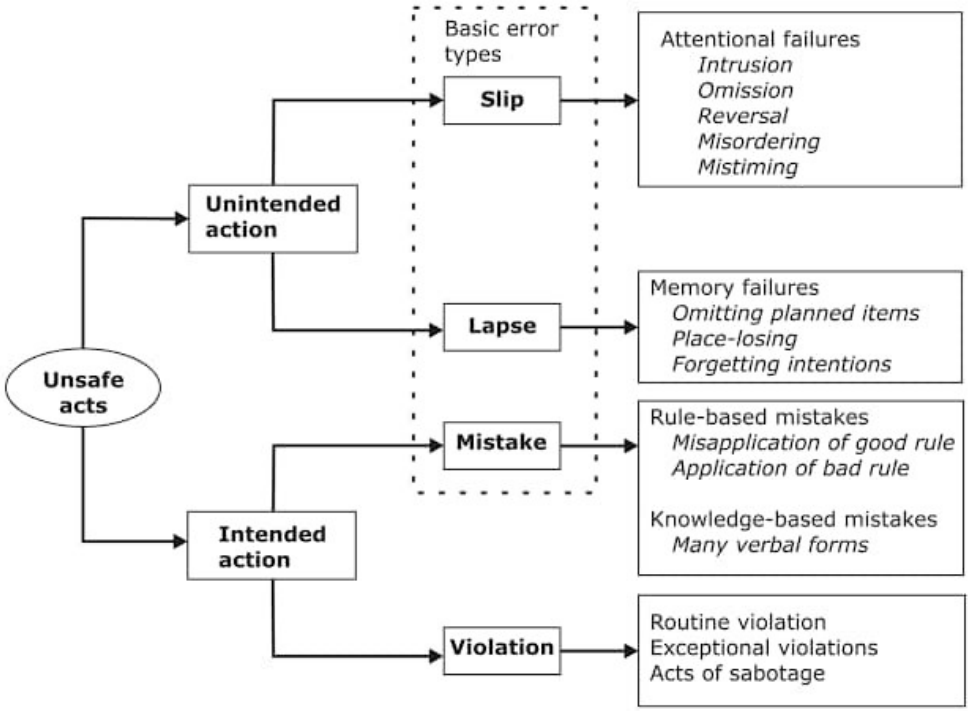

Figure 1: (Reason, 1990) error types.

The main problem with the SRK framework, apart from the lack of clear definition of what Skills, Rules, and Kowledge mean, respectively, is that no one intentionally chooses to work on a specific level, just as no one chooses to think fast or slow according to Kahneman's (2011) theory, if they do it must ironically by definition be by thinking

fast. People adjust their peformance, or mode of working and thinking to the condtitions without much, if any deliberation. If they perceive a risky unusual situation, they may try to trade-off efficiency for thoroughness, but never as an explicit choice not to be thorough. Reason's GEMS provides the following categories of human error:

- **Mistakes**: (e.g., following a bad checklist)These are further broken down into:
  - **Slips**: Observable, unintended actions (e.g., flipping the wrong switch).
  - **Lapses**: Internal memory failures (e.g., forgetting to complete a step), and
  - **Violations**: Intentional deviations from rules, procedures, or safety

In all cases decision making is regarded as a function. As a rule I will assume that functions, such as decision making, can be detected or seen in a straightforward way by an observer - although they may not be directly observable from a more stringent philosophical point of view, or open to introspection.

human error is, however, not a function, but a cause (or, to be precise: an assumed cause). We can use the term in a functional sense, as when we say that someone is making a mistake or an error. But in neither case is the human error an activity, nor the result of an intention. It is simply a contradiction of any reasonable definition to say that a person can make an error intentionally. It would therefore also be meaningless to call it a function. It may be argued that human error characterises the outcome of an action rather than the cause. But classifying an outcome as a human error is a misuse of the terminology. What is meant is rather that the outcome was agreed to be caused by a human error. Neither can the human error be the activity that leads to the outcome. We cannot classify an activity as being a human error, since that would assume that making the error was intentional. As that is not the case, it will be more correct to classify the activity as a failure to accomplish the intended outcome.

As a cause, human error must be inferred from observations rather than observed directly, just as causality, according to David Hume is metaphysical rather than physical. Other examples of such non-observables are "goal", "memory", etc. This requires that it is possible specify the observations from which the inferences are made. These observations will normally be about a particular performance or segment of a performance. We may observe the performance of a person, classify it as being incorrect, and determine the cause to be a human error. But in no case can we observe the human error directly.

Since human error is inferred, it is not necessarily unique. Another way of saying this is by noting that human error is just one explanation out of several possible for an observed performance (or more precisely, a part of an actual performance description, cf. Hollnagel , Pedersen & Rasmussen, 1981). The analysis is normally carried just far enough to find a plausible explanation. If an explanation, which refers to the technological parts of the system, cannot be found the category human error is used instead (cf. Rasmussen, 1981) and also (Hale & Hovden, 1998, and also Chapter 9). It is only when the analysis is carried beyond this point that we may realise that an explanation in terms of human error is insufficient.

Since my major source of experience is operators in control of a complex process (for instance, a nuclear power plant), I will assume that the focus is on what conventionally was called human-machine systems, now referred to as socio-technical systems, or joint cognitive systems (Hollnagel & Woods, 2xxx) we deal with is a Man-Machine System (MMS) that functions as a process control system. By an MMS I mean a system that is composed of one or more operators and one or more machines (usually computers) that are designed to support the control of the process. A particular example of this approach is the Cognitive Systems Engineering (cf. Hollnagel & Woods, 1983; Hollnagel & Woods, 2005; Woods & Hollnagel, 2006).

When the performance at work is observed (and evaluated) a mismatch may be detected between the actual and the intended system states, or between the achieved results and the goal. The detection of this mismatch presumes that a description of the intended system state (or goal) is available. The mismatch is assumed not to be random, and therefore to have an identifiable cause. Finding the cause amounts to accounting for the observed difference betaween expected and actual performance. If faults in the technological parts of the system cannot be found, the solution is generality to assign the variance (or residual variance) to people at the sharp end, hence to use human error as an explanation.

The detection of this mismatch is thus the observational basis for inferring the existence of a human error. It should be noted that if there is no observed mismatch, then there is little or no motivation to look for a cause. Variations in performance do not necessarily lead to undesired outcomes, hence mismatches. In fact they are usually an indispensable part of work that goes well. They may, for instance, be detected and corrected, at an early stage or the environment can be sufficiently friendly and forgiving. There will consequently be many cases of performance variability that remain unnoticed. From the point of view of a theory of human error they are,

however, just as important as the cases where a mismatch is observed, and should therefore be accounted for by it. The opposite of what reason claimed.

The crucial point thus is a mismatch between intended and actual outcomes of an action. If the functional analysis is carried one step further, it will show that the cause of the mismatch can be located either in the selection of the goal for the action (the formation of the intention) or in the execution of the actions designed to achieve that goal. One may even distinguish between a larger number of categories by using one of the models of human decision making, or a theory of human performance. But this actually reduces the need for a specific theory of human error, since the observed discrepancies instead can be explained by referring to, for instance, a performance theory. That may furthermore have the virtue of focusing on the situation and context in which the MMS must function, and the interaction between its inherent characteristics and the environmental constraints.

Observed mismatches in performance are always caused, in the sense that they can be analysed until the necessary and sufficient conditions, at least according to the underlying model or theory for their occurrence have been established. In some cases they may be classified as random, but that just means that the natural performance variability is sufficient to account for the mismatch, hence that no definite “other” cause has been identified or is require by the explanatory framework in use.

Since errors are not intentional, and since we do not need a particular theory of errors, it is meaningless to talk about mechanisms that produce errors. If the insatiable need for causes compel us to look for causes or "mechanisms", we should rather be concerned with whatever may be used to explain everyday work that goes well. If we are going to use the term psychological mechanisms at all, we should refer to possible “faults” in the functioning of how our minds usually function rather than to “error producing mechanisms”. We must not forget that in a theory of action, the very same mechanisms must also account for the correct performance which is the rule rather than the exception.

Even though we do not yet have a sustainable“theory of error, it makes sense to distinguish between endogenous and exogenous causes for UUOs. There are certainly cases where an UUO can be attributed to external causes, such as a bad interface design, interruptions, lack of operational support, incorrect information, time pressure etc. Similarly, there are cases where the causes are of an internal rather than external nature. In most cases the cause is, however, best described as a mixture. Stress, for instance, is often caused by (situationally) unreasonable demands to the operator. And deficiencies in the design of the interface may often be compensated by the adaptability of the operator (cf. Taylor & Garvey, 1959). Replacing a theory of error with a general theory of human action increases rather than reduces the importance of both internal and external causes, and emphasises the need to carry the analysis as far as possible.

To conclude, a theory of error must be a theory of the interaction between human performance variability and the situational constraints.

12.2: Terminology And Taxonomies

Other dimensions can easily be found, and several complete taxonomies are available. One good example is the CSNI taxonomy (cf. Rasmussen et at., 1981), which is an attempt to characterise the situation where a mismatch occurs, rather than the “human errors”. In this taxonomy human error is simply one of the many possible causes for a reported incident. Other taxonomies can rather easily be suggested once a proper theoretical background has been established. The choice of a taxonomy must depend on the purpose of the description, e.g., whether one wants to reduce the frequency of reported incidents, or improve the understanding of human decision making.

Before the key terms are defined, it is important to make sure that they are properly selected. One can, of course, make a pot pourri of terms that are normally used to characterise situations where humans make mistakes or errors, and then define them, e.g., by using a recognised dictionary. But if the definitions are to serve a purpose, it is essential that they have a common basis, for instance a theory. By the same rationale it also is essential that the terms have a common basis.

To repeat what has been said above, I believe we must develop forward a theory of Human Action rather than one forhuman error, and that this should be used for selecting and defining the key terms. Such a theory is not yet available, but I will nevertheless attempt to give a definition of some of the terms the organisers have listed, using intentional action as a basis.

Assuming that we can establish a theory of human action rather than human error, any predictions must be about such actions, rather than their outcomes, either EAOs or UUOs. Theories must specifically be about the variability of human action that can lead to UUOs. We can, of course, make a count of the instances where an

operator makes a mistake, i.e., and attribute the cause of the mismatch to a human error. But that does not mean that it is sensible to attempt to assess the reliability of the operator, even if we refrain from considering the operator in mechanistic terms. Making such a count furthermore assumes that a meaningful measurement has been defined.

12.3: Predictions of error

It is obvious for anyone who has worked with the reliability aspect of people at the sharp end, that the occurrence and frequency of what is called human errors depend more on the actual conditions of work (transitory Performance Shaping Factors, than on any stable inherent characteristic, or failure probability of individuals. Simple quantitative measures, such as error rates, are therefore inadequate and even misleading. Instead we need detailed and consistent descriptions of the conditions where mismatches occur. Such qualitative descriptions may eventually be used as a basis for more straightforward measurements.

With regard to the specific questions relating to prediction, it is at our present state of knowledge only possible to predict the frequency of mismatches and typical conditions at a place of work. We know from experimental psychology, particularly the studies of attention and performance, that there are important regularities, as diurnal variations, situation dependencies, etc. Even though most data come from simplified laboratory situations, there is no reason to assume that they cannot be applied to realistic work situations. This has been confirmed, for instance, by studies of shift-work. It is also highly plausible that there are significant individual differences in how a person relies on making efficiency-thoroughness trade-offs.

To summarise, making predictions requires a clear description of what the predictions are about. Unless frequencies and probabilities are sufficient, one must have a theory, or at least a good set of hypotheses, in order to make the predictions. It is furthermore logical that predictions cannot be about causes, unless we assume a strictly deterministic world with permanently fixed cause-effect relations. Any prediction must therefore be about outcomes, i.e., observed mismatches and the manifestations of the absence of safety, and possibly about whatever actions may have led to them. In the sense that human errors are assumed to be causes, we cannot make predictions of human errors.

From a practical point of view the most important question is less how mismatches can be explained, and more but about how they can be prevented. One clue to this is found in the cases where mismatches do not occur, either because they are detected and corrected by people, or because the system is sufficiently forgiving. It would be reasonable to look further into these possibilities tor prevent mismatches, hence to prevent future UUOs.

There are probably very many ways in which a place of work can be designed to facilitate the detection and correction of unintended performance variability. A good working theory of human action is invaluable in this respect, since it makes it possible to indicate more precisely when and how interventions to change the course of action can be made. It is probably better to design for general detection and correction rather than for specific prevention. The experience from all types of process control clearly shows that Murphy’s law cannot be beaten.

However, even if the best of systems has been designed, there will remain a basic variability in human performance that will lead to mismatches when the circumstances are right (or wrong, rather). If the operator was turned into an automaton (or even replaced by one, see the third essay), we might produce an error-free system, provided the degree of complexity was sufficiently low. But that is the crux of the matter. The mismatches may occur not just because of mistakes made by the operator during operation, but also because of mistakes made by the designer during earlier phases, or of those who are responsible for work management. These mistakes would not be contained unless a theory of human action was applied to literally every aspect of the system.

It is not enough to go **behind** human error (as advised by Woods et al.(2010) because that tacitly acknowledges that human error is a legitimate concept and phenomenon, otherwise one could not go behind it. It is instead necessary to leave the concept of human error completely and go **beyond** it. People have been obsessed with human error since the time of Cicero who in the year 43 BCE in a speech against Marcus Antonius uttered the famous words "errare humanum est", but errare means both to stray and to err. which in this context is quite ironic, because it is staying on the path of simple linear reasoning, rather than straying from it that is the “cause of the error” (yet to argue so unfortunately remains within the paradigm). human error as a result of simplified linear reasoning is, perhaps, the only kind of human error that is real. By continuing to do so we also disregard the second part of Cicero's famous statement, which was "nullius nisi insipienestis in errore perseverare" (Sometimes also rendered as "sed in errore perseverare turpe est"), meaning "but only a fool perseveres."(The way out of this is the solution on the other hand fairly simple as explained in this short poem by the Danish polymath Piet Hein: "The road to wisdom? Well, it's plain and simple to express: err and err and err again, but less and less and less." This short

poem clearly represents a Safety-I and decremental safety perspective. A revised version to represent a Safety-II and incremental safety perspective, could be ” The road to safety -- Well it’s plain as everyone can tell, try and try and try again, and learn from what goes well.”

It is possible to argue that it is being seduced or sucked in by oversimplified cause-effect thinking (unfortunately an inseparable part of most investigation methods) that is wrong – without that human error would not exist! So it is staying on the path rather than straying from it that is the “error”. The worst example is probably the attempt of St. Thomas of Aquinas to prove the existence of God by reasoning backward to the first cause of why we are here in the world. It clearly shows the drawbacks of thinking or reasoning within a constrained paradigm. Staying on the path, hence continuing what is wrong,whether it is going backward or forward is precisely what Cicero warned about, but that part has rather conveniently been forgotten.

References

Adamski, A. J., & Westrum, R. (2003). Requisite imagination: The fine art of anticipating what might go wrong. In Handbook of cognitive task design (pp. 193-220). CRC Press.

Alper, S.J., Karsh, B.-T., (2009).A systematic review of safety violations in industry. *Accid. Anal. Prev.*, doi:10.1016/j.aap.2009.03.013(1-16).

Arbib, M. A. (1987). *Brains, machines and mathematics*. New York: McGraw-Hill.

Ashby, W. R. (1960) *Design for a Brain. The origin of adaptive behavior. (Second Edition)*. London: Chapman & Hall, Ltd.

Bainbridge on performance variability;

Bainbridge, L. (1983a). Ironies of automation *Astomatica, 19*(6) pp. 775 779,

Besnard, D., & Hollnagel, E. (2014). I want to believe: some myths about the management of industrial safety. *Cognition, Technology & Work, 16*, 13-23.

Bloch, A. (2003). *Murphy's law*. Harmondsworth UK: Penguin.

Boole, G. (1854). *An investigation of the laws of thought: on which are founded the mathematical theories of logic and probabilities (Vol. 2)*. Walton and Maberly.

Booth, P. (1989). An introduction to human-computer interaction. London: Erlbaum.

Boring, R. L. (2012). Fifty years of THERP and human reliability analysis INL/ CON-12-25623).Idaho Falls, ID; IdahoNational Lab.(INL), (United States).

Carroll, J.M., Campbell, R.L. (1988). Artefacts as Psychological Theories: The Case of Human–Computer Interaction. User Interface Institute, IBM T.J. Watson Research Centre, Yorktown Heights, NY.

Cojazzi, G., Pedrali, M. & Cacciabue, P. C. (1993). Human performance study: Paradigms of human behaviour and error taxonomies ISEI/TE/2443/93). JRC Ispra, Italy: Institute for Systems Engineering and Informatics.

Dekker, S. W. A. (2006) Resilience engineering: Chronicling the emergence of confused consensus. In E. Hollnagel, D. D. Woods & N. Leveson (Eds.), Resilience engineering: Concepts and precepts. Hampshire: Ashgate.

Dekker, S. W. A., & Woods, D. D. (2002). MABA-MABA or abracadabra? Progress on human—automation co-ordination. Cognition, Technology ¢» Work, 4, 240-244.

Dekker, S. W., & Woods, D. D. (1999). To intervene or not to intervene: The dilemma of management by exception. Cognition, Technology & Work, 1, 86-96.

Dowell, J. & Long, J. (1989). Towards a conception for an engineering discipline of human factors. Ergonomics, 32(11), 1513-1535.

Dubois, D. & Prade, H. ((1989). Handling uncertainty in expert systems – pitfalls, difficulties, remedies. In E. Hollnagel (Ed.) The reliability of expert systems. Chichester: Ellis Horwood Ltd. (64 – 99).

Feigenbaum, E. A., & Feldman, J. (1963). Computers and thought (Vol. 7). McGraw-Hill: New York.

Fitts, P. M. et al. (1951). Human engineering for an effective air-navigation and traffic-control system. Washington D. C: National Research Council

Grøtan, T. O., Storseth, E, Ro, M. H., & Skjerve, A. B. (2008). Resilience,Adaptation and Improvisation—increasing resilience by organising for successful improvisation. In 3rd symposium on resilience engineering (pp. 28-30).

Hirschhorn L. (1997). Quoted in Cook RI, Woods DD and Miller C (Eds.), (1998). A Tale of Two Stories: Contrasting Views on Patient Safety. Chicago IL: National Patient Safety Foundation.

Hollnagel, E. (1983). *The design of reliable HCI: The hunt for hidden assumptions*. In J. L. Alty, D. Diaper & S. Guest (Eds.), People and computers VIII. Cambridge: Cambridge University Press.

Hollnagel, E. (1992). The art of efficient man-machine interaction: Improving the coupling between man and

machine. Invited paper for CADES workshop, St. Prix, France, August 1992. (Also in: J.-M. Hoc, P. C. Cacciabue & E. Hollnagel (Eds.), *Expertise and technology: Cognition & human-computer cooperation*. Lawrence Erlbaum.

Hollnagel, E. (1998) *Cognitive Reliability and Error Analysis Method (CREAM)*. Oxford: Elsevier Science Ltd.

Hollnagel, E. (2013). Is safety a subject for science? *Safety Science*, *67*, 21-24.

Hollnagel, E. (Ed.), (2003). *Handbook of cognitive task design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hollnagel, E. & Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies, 18*, 583-600.

Hollnagel, E. & Woods, D. D. (1983). Cognitive systems engineering. New wine in new bottles. *International Journal of Man-Machine Studies, 18*, 583-600.

Hollnagel, E. & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: CRC Press

Hollnagel, E., Pedersen, O. M. & Rasmussen, J. (1981). *Notes on human performance analysis (Risø-M-2285)*. Roskilde, Denmark: Risø National Laboratory:

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

Hollnagel. E (2017) Can we ever imagine how work is done? *HindSight 25* 10-13

Hudson, P. T. W. et al. (2008). Meeting expectations: A new model for a just and fair culture. In SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production. OnePetro.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux

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

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, et al. (Eds.) *Decision making in action: Models and methods*. Norwood, NJ: Ablex, 138-147.

Laird, J., Newell, A. & Rosenbloom, P. (1987) SOAR: An architecture for general intelligence. *Artificial Intelligence 33*(1), 1-64.

Le Coze, J. C. (2022). The 'new view' of human error. Origins, ambiguities, successes and critiques. *Safety science*, *154*, 105853.

Leplat, J., & Hoc, J.-M. (1983). Task and activity in the psychological analysis of situations. Cahiers de Psychologie Cognitive et Cahiers de Psychologie Marseille , 3 (1), 49-63.

Lerner, A. Y. (1975). *Fundamentals of cybernetics*. London: Plenum.

Lindblom, C. E. (1959). The science of "muddling through". *Public Administration Review*, *19*, 79-88.

Lindsay, P. H. & Norman, D. A. (1972). *Human information processing, An introduction to psychology*. NewYork:Academic Press.

making longer: They do less than they should do, hence may render the system incapable of functioning as planned (e.g. default actions, wrong procedure, too little and too late). The fourth irony is that we try to limit or eliminate the variability of human performance that is needed for systems to function as they were intended to — or at all. This is an irony of automation, but also of human factors and HSE management in general. The ETTO-principle only referenced

March, J.G. & Simon, H.A. (1958). *Organizations*. New York: Wiley.

Mawhin, B., Cabon, P., & Buratto, F. (201 1). Integrating Aircrew Resources Variability into the Design of Future Cockpit. In 16th International Symposium on Aviation Psychology (p. 481).

Meister, D. & Enderwick, T. P. (2001). *Human Factors in System Design, Development, and Testing*. CRC Press, Inc.

Merton, R. K. (1936). The unanticipated consequences of purposive social action. *American sociological review,1*(6), 894-904.

Merton, R. K., & Barber, E. (2011). *The Travels and Adventures of Serendipity*. Princeton University Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review, 63*(2), 81.

Moehlenbrink, C., & Papenfuss, A. (2011, September). ATC-monitoring when one controller operates two airports: Research for remote tower centres. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 55, No. 1, pp. 76-80). Sage CA: Los Angeles, CA: Sage Publications.

Morozow, E. (2013a). The perils of perfection. The New York Times, 2 March 2013.

Motozow, E (2013b). To save everything, click here. New York: The Perseus books group

Newell, A. (1990). *Unified theory of cognition*. Cambridge, MA.: Harvard University Press.

Newell, A. & Simon, H. A. (1963). GPS, A program that simulates human thought. In: E. A. Feigenbaum & J.

Feldman (Eds.) Computers and thought, 279-293. New York: McGraw-Hill Book Company.

Newell, A. & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Nyssen, A. S., & Blavier, A. (2019). Investigating expertise, flexibility and resilience in socio-technical environments: a case study in robotic surgery. In: E. Hollnagel, J. Braithwaite & R. L. Wears (Eds.) R*esilient health care* (pp. 97-110). Boca Raton, FL. CRC Press.

Ombredane, A. and Faverge, J. M. (1955). L'analyse du travail, Paris: Presses Universitaires de France.

Papenfuss, A., & Friedrich, M. (2016, September). Head up only—a design concept to enable multiple remote tower operations. In 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC) (pp. 1-10). IEEE.

Paxton, L. J. (2007). "Faster, better, and cheaper at NASA: Lessons learned in managing and accepting risk. Acta Astronautica, 61(10), 954-963.

Rasmussen, J. (1981). *human errors. A taxonomy for describing human malfunctioning in industrial installations (Risø-M-2304)*. Roskilde, Denmark: Risø National Laboratory.

Rasmussen, J.,et al. (1981). *Classification system for reporting events involving human malfunctions (Risø-M-2240), SINDOC(81) 14*. Roskilde, Denmark: Risø National Laboratory.

Reason, J. (1990). Human error. Cambridge university press.

Reason, J. T. (1990). Human error. Cambridge university press.

Roesler, A.,et al., (2001). *Design is telling stories about the future*. Columbus, OH: Cognitive Systems Engineering Laboratory, Ohio State University,.

Sarter, N.B., Woods, D.D., Billings, C.E., 1997. Automation surprises. In: G. Salvendy, (Ed.), *Handbook of Human Factors and Ergonomics, 2nd Ed*. New York: Wiley

Shorrock. S. (2020, October 28). Proxies for work-as-done: 1. Work-as-imagined. Humanistic Systems.https://humanisticsystems.com/2020/10/28/proxies-for-work-as-done-1-work-as-imagined/

Simon, H. A. (1947). *Administrative Behavior: a Study of Decision-Making Processes in Administrative Organization (1st ed.)*. New York: Macmillan. OCLC 35650

Simon, H. A. (1972). *The sciences of the artificial*. Cambridge, MA.: The M. I. T. Press.

Stassen, H. G. (1986). Decision demands and task requirements in work environments: What can be learned from human operator modeling. In E. Hollnagel, G. Mancini & D. D. Woods (Eds.), *Intelligent decision support in process environments*. Berlin: Springer Verlag.

Strauch, B. (2017). Ironies of automation: Still unresolved after all these years. *IEEE Transactions on Human-Machine Systems, 48*(5), 419-433.

Swain, A. D. & Guttman, H. E. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications (NUREG CR-1278)*. Washington, DC: Nuclear Regulatory Commission.

Taylor E W.(1919). *The principles of scientific management*. New York: Harper & Brothers Publishers

Taylor, F V. & Garvey, W. D. (1959). The limitations of a procrustean approach to the optimisation of man-machine systems. *Ergonomics*, *2*, 187-194.

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. J. of Math, 58(345-363), 5.

Turing, A.M. (1960). Computing machinery and intelligence. *Mind*, *LIX* (236),433-460.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*, 281-299.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on systems, man, and cybernetics, 22*(4), 589-606.

Von Főrster, H., & Pőrksen, B. (2002). *Understanding systems: Conversations on epistemology and ethics.* Heidelberg, Germany: Carl-Auer Verlag.

von Neumann, J. (1958). *The computer and the brain*. Yale University Press

Weick, K. E. 1987. Organizational culture as a source of high reliability. *California Management Review 29*(2), 112-128.

Weizenbaum, J. (1976). *Computer power and human reason. From judgment to calculation*. San Francisco: W. H. Freeman

Wiener, N. (1966). *God & Golem, Inc.: A comment on certain points where cybernetics impinges on religion* (Vol. 42). Boston, MA: MIT press.

Wienet, N. (1988). The human use of human beings: Cybernetics and society (No. 320). Da Capo Press.

Womack J, Jones D, Roos D. *The machine that changed the world*. New York: Simon & Schuster; 1990.

Woods, D. D. & Hollnagel, E. (2006). J*oint cognitive systems: Patterns in cognitive systems engineering*. Boca Raton, FL: CRC Press.

Zadeh, L. A. (2014). A note on modal logic and possibility theory. Information Sciences, 279, 908-913.

Zwetsloot, G. I. J. M., Aaltonen, M., Wybo, J.-L, Saari, J. Kines, P. & Op De Beeck, R.(2013). The case for research into the zero accident vision. Safety Science, 58, 41-48.

Long-term knowledge . An operator who finds out how to control the plant for himself, without explicit training, uses a set of propositions about possible process behavior, from which he generates strategies to try (e.g, Bainbridge. 1981). Similarly an operator will only be able to generate successful new strategies for unusual situations if he has an adequate knowledge of the process. There are two problems with this for machine-minding operators. One is that efficient retrieval of knowledge from long-term memory depends on frequency of use (consider any subject which you passed an examination at school and have not thought about since). The other is that this type of knowledge

Performance variability as putty

When WAI is compared to WAD, there is always a small gap or discrepancy. The two never match completely (Hollnagel, 2017). So just as when bricks are put together to build a wall or a house, something is needed to fill out and cover the gaps. The alternative would be to ensure that the bricks fit perfectly together like Lego blocks or the stones in the walls of the pyramids, Roman bridges or Japanese castles. But it is far easier to use putty to fill out the gaps and in this way disguise the differences. In the context of human-machine systems, human performance

variability plays the role of putty and helps to ensure that the wall works as it should and looks complete particularly when it has been in place for some years.

Conclusion: The irony of human factors

As this paper hopefully has demonstrated,human factors as such therefore comprise five ironies, one for each of training and design plus three for automation. The overall irony for human factors is that it considers human performance variability as a liability and, even though it is an asset without which few systems would work.

Wiener, N. (1966). God & Golem, Inc.: A comment on certain points where cybernetics impinges on religion (Vol. 42).

MIT press.

Hollnagel, E., & D Woods, D. D. (1983). Cognitive systems engineering: new wine in new bottles. International Journal of Human-Computer Studies, 5 1(2), 339-356.

WAI and WAD

Origins, Shorrock WP

Hej, nu kan jeg igen ikke bruge mit dankort. jeg ved godt at der er en maxgrænse, men det ville være fint. hvis det også fremgik af kontooversigten