Human error Analysis: Review of past accidents and implications for improving robustness of system design

R. Moura*, M. Beer, E. Patelli & J. Lewis
Institute for Risk and Uncertainty, University of Liverpool, United Kingdom

F. Knoll
NCK Inc., Montreal, Canada

ABSTRACT: Since the establishment of the high-technology industry and industrial systems, developments of new materials and fabrication techniques, associated with cutting-edge structural and engineering assessments, are contributing to more reliable and consistent systems, thus reducing the likelihood of losses. However, recent accidents are acknowledged to be linked to human factors which led to catastrophic consequences. Therefore, the understanding of human behavioural characteristics interlaced with the actual technology aspects and organisational context is of paramount importance for the safety & reliability field. This study first approaches this multidisciplinary problem by classifying and reviewing 200 major accident data from insurance companies and regulatory authorities under the Cognitive Reliability and Error Analysis framework. Then, specific attention is dedicated to discuss the implications for improving robustness of system design and tackling the surrounding factors and tendencies that could lead to the manifestation of human errors.

1 INTRODUCTION

1.1 *The need for human reliability analysis*

Despite the evolution of engineering methods and fabrication techniques, recent accidents involving complex industrial systems within high-technology industries such as oil & gas, nuclear and aerospace were acknowledged to be consistently interrelated to human factors which led to tragic consequences (e.g. Rio-Paris Flight 447, Macondo and Fukushima). No one can deny the impact of human actions and decisions on the performance of engineering systems, hence justifying the need for evaluating the interaction between humans and systems through a suitable Human Reliability Analysis (HRA).

HRA is mainly a predictive tool, intended to estimate the probability of human errors and assess the human factors contribution to the overall risk through the use of qualitative and/or quantitative methods. Essentially, traditional HRA tools consist of (i) the identification of potential human erroneous actions, followed by (ii) the consideration of internal and external factors that could influence the human performance, finally resulting in (iii) the estimation of human error probabilities. The key step in this approach is defining possible tasks to be performed by a human operator, considered to be an element or component subjected to failure due to inborn characteristics, thus having an "in-built probability of failure". Modifiers known as performance shaping factors, error-forcing conditions, scaling factors or performance influencing factors, depending on the methodology, are then applied to adjust the likelihood of human failure when performing the assessed task.

On the other hand, some contemporary approaches to HRA such as "A Technique for Human Error Analysis" – ATHEANA (Barriere et al. 2000), the Connectionism Assessment of Human Reliability (CAHR) based on Sträter (2000) and the Cognitive Reliability and Error Analysis Method (CREAM) by Hollnagel (1998) were developed around the principle that the fundamental element is, in fact, the context in which the task is performed, reducing previous emphasis on the task characteristics *per se* and on a hypothetical inherent human error probability.

Several HRA methods comprising both "task-centred" and "context-centred" approaches are currently in use. Bell & Holroyd (2009) reported 72 different techniques to estimate human reliability and considered 35 as potentially relevant. Further analysis highlighted 17 of these HRA tools to be of potential use for major hazard directorates. For additional details on some of the existing methods, a summary of five commonly used HRA techniques in the offshore industry and other high-hazard facilities has been presented by Moura (2012), where the main features of the Human Error Assessment and Reduc-

*National Agency for Petroleum, Natural Gas and Biofuels (ANP), Brazil.

tion Technique (HEART), the Technique for Human Error Rate Prediction (THERP), the Standardized Plant Analysis Risk (SPAR-HRA), the Justification of Human Error Data Information (JHEDI) and the Systematic Human Actions Reliability Procedure (SHARP-1) were reviewed.

Regardless of the variety of HRA methods available to enable practitioners to assess the risks associated with human error by estimating its probability, the substantially high uncertainties related to the human behavioural characteristics, interlaced with actual technology aspects and organisational context, turn this kind of evaluation into a very complicated matter, raising reasonable concern about the accuracy and practicality of such probabilities.

1.2 *The data collection problem*

Swain (1990) argues that the most serious problem for HRA is the unavailability of meaningful data (or less-than-adequate data) on human performance that could assist quantitative predictions of human behaviour in complex systems. Posing similar reservations, The International Atomic Energy Agency (1990) report on human error classification and data collection stressed this matter, claiming that a global data collection scheme would be *extremely difficult, if not impossible, to create*, justifying this unenthusiastic view by the amount of data related to each human error that should need to be generated in order to satisfy a wide range of objectives required by applications dependent on human performance databanks.

Indeed, in spite of some recent efforts to gather human performance data and overcome this issue, such as Gibson & Megaw (1999) CORE-DATA, the accessibility and usage of data is still very limited, due to four main reasons. Initially, human performance data collection usually begins within a specific domain, making the classification scheme or taxonomies closely associated with the industry where the issue was first raised. Trying to transpose terms and nomenclatures from one industry to another is challenging, thus human error information are likely to be captured under different frameworks.

Secondly, the insertion of the human performance data in a HRA, which would have to be subsequently combined with an equipment/system reliability analysis to finally reach a meaningful result in the probabilistic safety assessment, is complex and time consuming. Consequently, practitioners tend to use expert elicitation (or a qualitative approach) as a shortcut to include human error data in probabilistic risk/safety assessments, leaving the commitment to collect quantitative data behind.

In addition, the data collection and interpretation process could lead to significantly different results, depending on the data source and the acquisition method (field data, simulator data, HRA modelling, expert elicitation, operation observation, reporting methods, near misses, performance indicators, accident investigation reports etc.).

Lastly, despite the increasing complexity of engineering systems and safeguards reduced the possibility of accidents caused from a single failure, the exact interactions between human behaviour, technology and organisation currently required to produce the specific chain reaction leading to a disaster are very difficult to predict and highly associated with a specific work situation, thus unlikely to be repeatable.

2 CLASSIFICATION METHOD

2.1 *The Cognitive Reliability and Error Analysis Method (CREAM) taxonomy as a common framework to classify accidents*

An innovative use of the CREAM taxonomy was conceived as a common framework to classify accidents from different industries. Other classification methods such as The Human Factors Analysis and Classification System - HFACS (Shappell et al. 2007) and the Error Promoting Conditions (EPCs) from the Human Error Assessment and Reduction Technique (HEART) presented by Williams (1989) revealed good potential to be used as a classification scheme (in fact, it is one of the central functions of HFACS) and were also reviewed. However, the former taxonomy has deep associations with aviation accidents (e.g. crew resource management and personal readiness), and the latter does not differentiate between human, technological and organisation EPCs among the 39 nomenclatures available. In addition, HEART is designed to be used prospectively, i.e. to estimate human error probability, while HFACS main application is in retrospective analysis of accidents in the aviation field. Therefore, the use of both would have implied adjustments to the original taxonomies.

Conversely, the CREAM taxonomy (Figures 1, 2 and 3) offers a very effective and clear division between human erroneous actions (phenotypes) and their possible causes or contributing factors (genotypes), grouped by three expressive categories: man, technology and organisation. Furthermore, the terminologies are generic and easily associated with most industrial sectors.

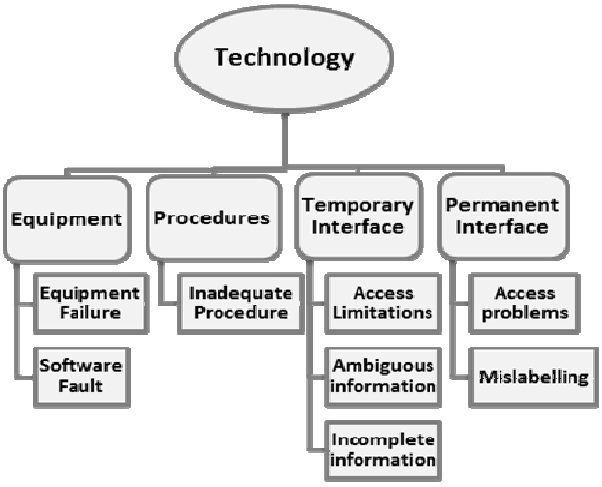

Figure 1. "Technology" taxonomy, adapted from Hollnagel (1998).

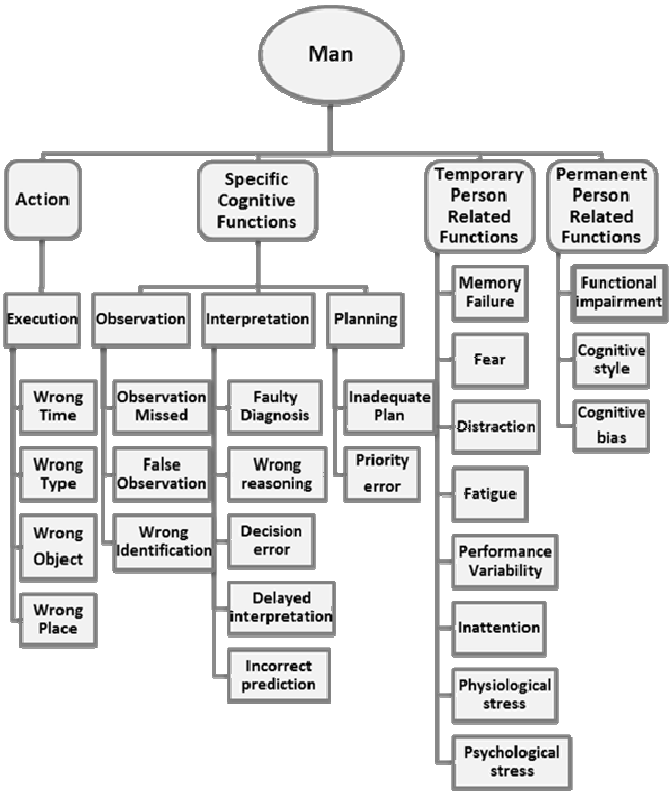

Figure 3. "Man" taxonomy, adapted from Hollnagel (1998).

A noteworthy feature of CREAM's original application as a Human Reliability Analysis approach is the fact that the method was designed for both prospective and retrospective evaluations, whether supporting accident investigations aiming at the search for causes, or predicting events that are likely to occur. This flexibility enables the use of the CREAM HRA method as a data input to the chosen classification scheme, whether prospectively, via simulation, or retrospectively, through the execution of accident analysis.

It is worth to notice that this work applies a "retrospective method" for data collection based on CREAM taxonomy, but certainly not following the procedures developed by Hollnagel to perform acci-

Figure 2. "Organisation" taxonomy, adapted from Hollnagel (1998).

dent analysis. It means that the CREAM retrospective technique is not used whatsoever, considering that the accidents collected were submitted to extensive investigation and the causes and contributing factors are already exposed in the reports, being the current gap the need for interpretation and classification of the data under a common framework.

3 REVIEW OF 200 MAJOR ACCIDENTS DATA

3.1 *Data Collection*

To deal with uncertainties related to data collection and overcome the "less-than-adequate" data problem in this research, factual data from accident investigation reports and case studies were interpreted and then organised to fit the scheme presented in Figures 1, 2 and 3. These accident reports were collected from insurance companies, regulatory bodies, commissions of inquiry, investigation boards and industry experts, such as MARSH Inc., the Australian Department of Industry and Resources (DoIR), the Australian National Petroleum Safety Authority (NOPSA), the Brazilian National Petroleum Agency (ANP), the European Agency for Safety and Health at Work (EU-OSHA), the National Aeronautics and Space Administration (NASA), the Norwegian Foundation for Scientific and Industrial Research (SINTEF), the Petroleum Safety Authority Norway (PSA), the UK Department of Employment, the UK Health and Safety Executive, the US Bureau of Safety and Environmental Enforcement (BSEE), the US

Chemical Safety and Hazard Investigation Board (CSB), the US Department of Energy (DoE), the US Environmental Protection Agency (EPA), the US Fire Administration (USFA), the US Minerals Management Service (MMS), the US National Transportation Safety Board and the US Occupational Safety and Health Agency (OSHA).

These events are listed in Table 1 by industrial activity.

Table 1. Accidents distribution by industry

Industry	Accidents	
	#	%
Aviation	09	4.50
Chemicals Factory	29	14.50
Chemicals Storage	01	0.50
Construction	12	6.00
Crystal Factory	01	0.50
Education	01	0.50
Fiberglass Factory	01	0.50
Fireworks Storage	01	0.50
Food Industry	03	1.50
Gas Processing	09	4.50
Gas Station	01	0.50
Hydroelectric Plant	01	0.50
Ink & Paint Factory	01	0.50
Metal Signs Factory	01	0.50
Metallurgical Industry	06	3.00
Oilfield Waste Disposal Plant	01	0.50
Petrochemicals	25	12.50
Pharmaceutical Industry	01	0.50
Polymers Factory	01	0.50
Power Plant	01	0.50
Refinery	38	19.00
Sterilisation Services	01	0.50
Sugar Factory	01	0.50
Terminals and Distribution	12	6.00
Tyre & Rubber Factory	01	0.50
Upstream (Oil & Gas)	36	18.00
Waste Treatment Plant	03	1.50
Waste Water Plant	01	0.50
Water Supply	01	0.50

3.2 *Relevance of the data sample*

The UK Control of Major Accident Hazards Regulations (1999) define "major accident" as *an occurrence (including in particular, a major emission, fire or explosion) resulting from uncontrolled developments in the course of the operation of any establishment and leading to serious danger to human health or the environment, immediate or delayed, inside or outside the establishment, and involving one or more dangerous substances.* Most of the 200 accidents reviewed fit the United Kingdom's legal definition for major accident.

Forty four of the studied accidents accounted for 1,075 fatalities.

Of the events analysed, 91 reports included some cost information, thus it was possible to calculate a gross estimate (using cumulative inflation rate to update to 2013 values) of £20.25 billion in material losses. Furthermore, it is acknowledged that costs related to environmental recovery, litigation etc. would dramatically increase these figures. For example, in a Wall Street Journal article about the Macondo Oil Spill, which occurred in the Gulf of Mexico in April 2010, Fowler (2013) reported that BP had already paid around US$ 14 billion in spill clean-up; US$ 10 billion to individuals, businesses and governments; US$ 4 billion in a criminal settlement; and US$ 2 billion in environmental restoration and research. Not to mention the impact on the company's reputation/brand, reflected in the stock market: Bell (2012) reported a 35.00% drop of BP's stock price from 2010 to 2012.

Hence, these numbers highlight the significance of this research's data sample.

3.3 *Data collection and classification demonstration method*

To demonstrate exactly how the interpretation and classification of the data was performed, a heat exchanger rupture followed by an ammonia release, which occurred in a Tyre & Rubber Factory, will be scrutinised on Table 2. This is based on the study case publication from the US Chemical Safety and Hazard Investigation Board (2011).

Basically, on 11 June 2008, a heat exchanger containing ammonia has ruptured due to an overpressure, triggered by a failure during the maintenance process. One worker was killed and seven others were injured during this event.

Table 2. Tyre & Rubber factory classification example

Group	Sub-Group	Factor	Justification*
Man	Execution	Wrong Place	Sequence, jump forward. One operator closed an isolation valve between the heat exchanger shell (ammonia cooling side) and a relief valve, to undertake a maintenance job, specifically a burst rupture disk replacement. However, after replacing the rupture disk, the reopening of the closed isolation valve was skipped, due to unknown reasons. On the following day, another operator closed a block valve to clean the pipe, isolating the ammonia pressure control valve from the heat exchange.
Man	Cognitive	Observation missed	Second operator did not observe closed status of isolation valve and

	Function (Observation)		closed block valve, isolating the ammonia pressure control valve from the heat exchange. Afterwards, he started to steam the process line to clean the pipeline. The closed isolation and block valves prevented the increasing ammonia pressure from safely venting through either the ammonia pressure control valve or the rupture disk and relief valve.
Technology	Equipment	Software Fault	Information delay. During the emergency, a malfunction in the computerized electronic badge-in/badge-out system delayed supervisors from immediately retrieving the list of personnel in their area.
Technology	Procedures	Inadequate procedures	Incomplete text/missing instructions. i) ASME Code requirement to have a competent person, capable of releasing the pressure, continuously monitoring pressure vessels with relief devices blocked or when there is a possibility of pressurisation above design limit, was not reflected in operating procedures. ii) Supervisors were to account for their employees using a master list generated from the computerized electronic badge-in/badge-out system, but there were no further instructions on how to compensate the expected absence of some members of the emergency response team. Therefore, the supervisors were unable to detect if there was a missing worker designed as a member of the emergency response team.
Organisation	Communication	Communication Failure	Message not received. Turnover documents and sign-off procedures that should have alerted whether maintenance was completed or not, as well as the current status of the process, were not effective communication channels.
Organisation	Organisation	Maintenance Failure	The relief system was not available due to a maintenance failure. Although maintenance workers had replaced the rupture disk, the valve isolating the rupture disk was not reopened thus the relief valve was not available.
Organisation	Organisation	Inadequate Quality Control	Inadequate quality control procedures. The information that the isolation valve on the safety relief vent remained in the closed position and locked-out was limited to a handwritten note. Although quality control measures (lockout/tag-out procedures to manage the work on the heat exchanger rupture disk were applied), the progress and current state of the maintenance job was not clearly documented. Moreover, there was not a post-maintenance quality check to verify if the service was properly concluded.
Organisation	Organisation	Design Failure	Inadequate emergency system design, poor location of instruments and alarms. Alarm system was not fit for purpose due to a design failure. During emergency, operators were supposed to sound a location-specific alarm using pull-boxes located throughout the production unit. However, the water spray from the automatic water deluge system and the ammonia vapour prevented responders from reaching the alarm pull-box in the affected process unit.
Organisation	Organisation	Inadequate Task Allocation	Inadequate managerial rule (deficient organisation of work due to the lack of safety principles). After the conclusion of the maintenance job (rupture disk replacement), the work order system required the process operator to sign-off upon the completion of service. However, there was no evidence that this occurred (no signal of a signed copy of the work order was found).
Organisation	Organisation	Social Pressure	Group thinking. Although operating procedures required maintenance personnel to obtain production operators' signatures in work orders and keep these documents at production control stations, this was not regularly followed. Therefore, the blatant institutional disregard for procedures shows that the risk perception/understanding of the maintenance worker was clearly guided by the working practice among maintenance personnel.
Organisation	Training	Insufficient Skills	Performance failure. Supervisors failed to perform workers headcount during emergency measures. Some employees had not been fully trained on partial and plant-wide evacuations procedures. Procedures established that plant-wide evacuation and shelter-in-place drills should be conducted at least four times a year, but such drills had not been conducted in the four years prior to the incident. Drills could have indicated problems with the worker headcount procedure.

* Inferred from evidences and descriptions from the US Chemical Safety and Hazard Investigation Board (2011) Study Case.

Therefore, in the above example, the operator failed to perform the maintenance job sequence, skipping the reopening manoeuvre of a closed isolation valve after replacing a rupture disk. The investigation report did not provide further information about any person related function (memory failure,

distraction, fatigue, inattention, stress etc.) that could have led to this specific phenotype (erroneous action). In the following day, another operator overlooked the status of the isolation valve (here we have a clear cognitive function failure – observation) and started to clean the process line. After classifying two man-related factors, one action error (wrong place, sequence, jump forward) and another specific cognitive function (observation missed), further two technology genotypes (equipment and procedures, linked to the escalation of the accident) and six organisation genotypes (communication, maintenance, quality control, design, task allocation, social pressure and training) were identified. The genotypes are the identified factors that caused or contributed to the accident or contributed to its escalation.

3.4 *Data Classification Results*

The examination of the 200 major accidents was similarly performed, and the following tables summarise the outcomes from these interpretations as well as the resulting categorisation.

Table 3. Data Classification results (main groups).

Group	Frequency*	
	#	%
Man	105	52.50
Technology	170	85.00
Organisation	189	94.50

*Number of events where groups appeared.

Table 4. Data Classification results (factors & sub-groups).

Factor	Frequency*		Sub-Group	Freq.*
	#	%		%
Wrong Time	29	14.50	Execution	50.00
Wrong Type	27	13.50		
Wrong Object	06	3.00		
Wrong Place	51	25.50		
Observation Missed	29	14.50	Cognitive	
False Observation	04	2.00	Functions**	45.50
Wrong Identification	07	3.50		
Faulty diagnosis	25	12.50		
Wrong reasoning	23	11.50		
Decision error	17	8.50		
Delayed interpretation	06	3.00		
Incorrect prediction	07	3.50		
Inadequate plan	18	9.00		
Priority error	15	7.50		
Memory failure	02	1.00	Temp Person	
Fear	03	1.50	Related	
Distraction	10	5.00	Functions	11.50
Fatigue	04	2.00		
Performance Variability	02	1.00		
Inattention	01	0.50		
Physiological stress	02	1.00		
Psychological stress	05	2.50		
Functional impairment	00	0.00	Perm. Person	
Cognitive style	00	0.00	Related	
Cognitive bias	12	6.00	Functions	6.00
Equipment failure	123	61.50	Equipment	62.50
Software fault	03	1.50		
Inadequate procedure	86	43.00	Procedures	43.00
Access limitations	01	0.50	Temporary	
Ambiguous information	06	3.00	Interface	13.00
Incomplete information	23	11.50		
Access problems	03	1.50	Permanent	
Mislabelling	03	1.50	Interface	3.00
Communication failure	22	11 .00	Communi-	
Missing information	36	18.00	cation	27.50
Maintenance failure	70	35.00	Organisation	93.00
Inadequate quality control	114	57.00		
Management problem	20	10.00		
Design failure	124	62.00		
Inadequate task allocation	111	55.50		
Social pressure	15	7.50		
Insufficient skills	73	36.50	Training	52.50
Insufficient knowledge	66	33.00		
Temperature	03	1.50	Ambient	
Sound	00	0.00	Conditions	10.00
Humidity	00	0.00		
Illumination	01	0.50		
Other	00	0.00		
Adverse ambient condition	17	8.50		
Excessive demand	10	5.00	Working	
Poor work place layout	06	3.00	Conditions	11.50
Inadequate team support	05	2.50		
Irregular working hours	07	3.50		

*Number of events where factors or sub-groups appeared.
** Cognitive functions detailed on Table 5.

Table 5. Data Classification results (cognitive functions).

Cognitive Function	Frequency*	
	#	%
Observation	37	18.50
Interpretation	63	31.50
Planning	31	15.50

*Number of events where cognitive functions appeared.

Tables 3, 4 and 5 specify the number of appearances of the man-related, machine and organisations phenotypes and genotypes in the major accidents examined. Percentages relate to the total of events (200).

At least one human element was identified in 52.50% of the cases, with 50.00% of direct erroneous actions (phenotypes). Cognitive functions accounted for 45.50%, with the interpretation genotype appearing as the most relevant (31.50%). At least one technology genotype was recognised in 85.00% of the accidents, highlighting equipment failure (61.50%) and inadequate procedures (43.00%) as the foremost factors related to this group. Organisational issues appeared in 94.50% of the accidents, emphasising design failures (62.00%), inadequate quality control (57.00%) and inadequate task allocation (55.50%) as the most significant genotypes within the group.

A segregation analysis of the collected data shows that the prospect of a single group (man, machine or organisation) causing an accident is low. Of the examined events, only 1.00% show an erroneous action with a man-related genotype resulting in an accident. Technological factors were solely responsible for the undesirable outcome in only 4.50% of the cases, while 8.00% of the accidents were due to exclusively organisational factors. Conversely, all combinations involving at least two groups featured significantly in the dataset. A Man-Technology arrangement appeared in 45.50% of the cases, while a Man-Organisation combination performed in 51.50%. The Technology-Organisation pair figured together in 80.50% of the events. In 45.50% of the cases, the three groups appeared together. Table 6 summarises these results.

Table 6. Segregation Analysis (main groups).

Group / Combination	Frequency*	
	#	%
Only Man	02	1.00
Only Technology	09	4.50
Only Organisation	16	8.00
Man-Technology	91	45.50
Man-Organisation	103	51.50
Technology-Organisation	161	80.50
Man-Technology-Organisation	91	45.50

*Events where a single group or combinations appeared.

It is important to notice that the design failure is the most significant single genotype from all three groups, appearing with an incidence of 62.00%, closely followed by equipment failure (61.50%).

There is also a close relationship between the design failure genotype and the man group: 69.00% of the erroneous actions (execution errors) were accompanied by a design failure. In addition, 58.62% of temporary and permanent person related functions and 71.43% of cognitive functions were also connected to design failures. Another interesting feature can be extracted from a tendency analysis along the years, considering the design failure genotype and the man group. Figure 4 suggests an increase on the percentage of the man group contribution along the years, while a stable behaviour of the design failure genotype is observed.

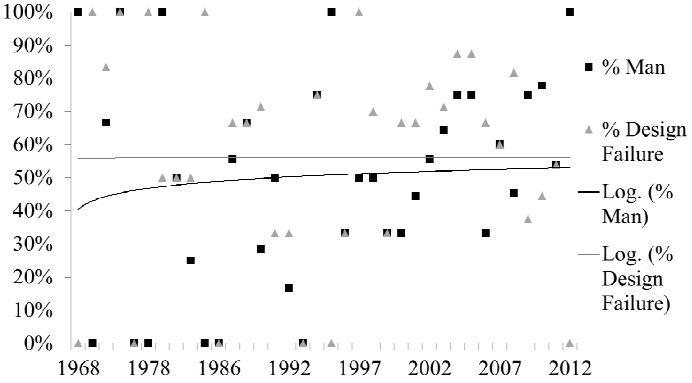

Figure 4. Man and Design Failure contribution along time.

4 DISCUSSION AND CONCLUSIONS

4.1 *Improving robustness of system design*

Even with the notable evolution of standards and regulations, the amount of design failures contributing to accidents does not appear to be reducing through time. The absence of a well-developed data source related to human performance in complex systems in addition to the diversity of human behavioural characteristics might be preventing reference documents and guidelines from containing unambiguous and direct design rules to deal with human interaction. This may, in turn, lead to this apparent inertia resisting a decrease in design failure. It could be explained by the fact that accidents are rare events, thus it is not an easy task to accumulate sufficient data to ascertain tendencies and improve robustness of system design. Therefore, data collection must be a long-term effort and needs to gather substantial evidence for it to be objectively converted into guidelines to be applied by designers.

This paper indicates that the use of taxonomies from a bi-dimensional approach such as CREAM to classify data is a potential solution to produce meaningful information from three different types of source using the same framework: (i) historical data, as demonstrated in this research, plus: (ii) incident investigations and (ii) prospective analysis, as in the original application of CREAM HRA. The common framework to conduct human reliability predictions as well as retrospective analysis of events during Human Reliability Analysis in a specific facility or industry is perfectly able to interface with the proposed classification scheme for past accidents, considering that they basically share the same taxonomy.

The increasing complexity of systems expanded the number of defences or barriers that needs to be breached in order to result in a major accident, explaining why accidents related to a single CREAM group are uncommon. During this study, the peripheral figures of one exact group appearance (1.00%, 4.50% and 8.00% for man, technology and organisation, respectively), contrasting with the statistics of at least two groups (45.50% for man-technology, 51.50% for man-organisation and 80.50% for technology-organisation), showed that the presence of two or more CREAM groups is generally necessary to produce an accident. However, this complexity makes it almost impossible for one individual to anticipate all design nuances and connections with human and/or technological factors that could lead to undesirable outcomes. In addition, the relationship between design failures and the man group, explicitly the connection with cognitive functions (71.43% were associated to design failures) and execution errors (69.00% linked to design failures), shows that design failure and damage tolerance criteria must include (and be tested against) specific

human related features highlighted by this study. This seems to be especially important with a failure in human interpretation of system status (wrong reasoning and faulty diagnosis), a potential observation missed and some execution errors (sequence, timing and type).

Therefore, the figures above indicate that reliability analyses should anticipate that cues, measurements or information originally intended to lead to a human action have a significant probability of being missed. Similarly, if system design allows situations in which some system analysis/diagnosis, interpretation or hypothesis formulation are required before taking an action, it is likely that specific cognitive functions (inferences, generalisations or deductions) will lead to undesirable outcomes. The design should also be tested against direct human execution errors, primarily focusing on: omissions; jumping forward a required action; performing a premature, delayed or wrong action; and performing a movement in the wrong direction, with inadequate speed or magnitude.

4.2 *Future Developments*

The significance of the design failure genotype indicates a well-defined path for future investigation regarding the genesis and perpetuation of human errors.

Further research could also approach other data collection relevant features, due to the significant numbers observed in inadequate quality control (57.00%), equipment failure and maintenance failure (61.50% and 35.00%, respectively), inadequate task allocation (55.50%) and training (52.50%). Quality control guidelines could possibly join a comprehensive design robustness strategy as a failure identification and exclusion tool, while design failure prevention schemes might address both equipment reliability and maintenance efficiency. In addition, effective operational strategies appear to be connected to improved organisation of work by clear rules and principles and individual performance enhancement through workforce skills and knowledge improvements (to handle equipment/perform tasks and increase situation awareness).

Additional effort to expand the data sample is also desirable, whether to confirm or increase the robustness of the results.

5 ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support and insights of Christopher Price-Kuehne (MARSH Global Energy Risk Engineering Team), Gary Munley (Office for Nuclear Regulation), Steve Walker (HSE Energy Division - Offshore) and Nicola Stacey (Health and Safety Laboratory) during the development of this work. This study was partially funded by CAPES (Proc. nº 5959/13-6).

6 REFERENCES

Barriere, M., *et al*. 2000. NUREG-1624 *Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA)*. Washington, DC: US Nuclear Regulatory Commission.

Bell, J. 2012. The Gulf Spill: BP Still Doesn't Get It. In Allen, F. E. (ed), *Forbes, 20 April 2012*. http://www.forbes.com/sites/frederickallen/2012/04/20/the-gulf-spill-bp-still-doesnt-get-it/

Bell J & Holroyd J. 2009. *Review of human reliability assessment methods*. Suffolk: HSE Books.

Fowler, T. 2013. BP Faces New Bout of Spill Liability. *The Wall Street Journal, 18 February 2013*. New York: Dow Jones & Company, Inc.

Gibson W. H. & Megaw T. D. 1999. *The implementation of CORE-DATA, a computerised human error probability database.* Suffolk: HSE Books.

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

International Atomic Energy Agency. 1990. Human Error Classification and Data Collection. *Report of a technical committee meeting organised by the IAEA, Vienna, 20-24 February 1989*. Vienna: INIS Clearinghouse.

Moura, R. 2012. *Improving risk perception during the lifecycle of an offshore oil & gas production facility: the influence of human and organisational surrounding factors in the major-hazard assessment outcomes*. Cranfield: Cranfield University.

Shappell, S., *et al*. 2007. Human Error and Commercial Aviation Accidents: An Analysis Using the Human Factors Analysis and Classification System. *Human Factors* 49(2): 227-242.

Sträter, O. 2000. *Evaluation of Human Reliability on the Basis of Operational Experience*. Cologne: GRS. (English translation of the Report GRS-138: Beurteilung der menschlichen Zuverlassigkeit auf Basis von Betriebserfahrung.)

Swain, A. D. 1990. Human Reliability Analysis - Need, Status, Trends and Limitations. *Reliability Engineering and System Safety* 29: 301-313.

The Control of Major Accident Hazards Regulations. 1999. http://www.legislation.gov.uk/uksi/1999/743/contents/made . Surrey: National Archives.

US Chemical Safety and Hazard Investigation Board 2011. *Case Study 2008-06-I-TX Heat exchanger rupture and ammonia release in Houston, Texas*. Washington, DC: CSB Publications.

Williams, J.C. 1986. HEART - A Proposed Method for Assessing and Reducing Human Error. *Proceedings of the 9th Advances in Reliability Technology Symposium, Bradford, 2-4 April 1986.* Warrington: National Centre of Systems Reliability.