IADC/SPE SPE-167983-MS

How can the Petroleum Industry Benefit from Human Reliability Analysis

Yuandan Li, Joseph J. Cho, and Yali Ren, Baker Hughes Inc.

This paper was prepared for presentation at the 2014 IADC/SPE Drilling Conference and Exhibition held in Fort Worth, Texas, USA, 4–6 March 2014.

Abstract

In the petroleum industry, the study and impact of human reliability on safe and efficient operations is receiving heightened awareness. According to recent studies, human errors contributed to an estimated 60 to 90 percent of all accidents across the petroleum, nuclear and aviation industries, etc.

Two possible outcomes of human errors are of concern: safety risks and the cost of failure. In operations where the consequences of accidents may be unbearable, the emphasis on safety drives the study and research of human reliability. This effort has led to a number of human reliability analysis (HRA) methods, and they provide useful references for managing human errors.

This paper examines how the petroleum industry could utilize and benefit from the available HRA methods by applying them to daily operations. Several HRA methods are reviewed in this paper in three aspects: human error identification, human error prediction of given tasks, and factors that contribute to human errors. A good data collection system can substantially support the HRA study qualitatively and quantitatively. This paper also introduces an available practice for human error data collection implemented in the nuclear industry—human event repository analysis (HERA). Suggestions and remarks are provided about developing similar human error data collection systems tailored for the petroleum industry. The expected benefits of HRA include: a better understanding of the causes and influencing factors of human errors in specific operation conditions, and insights into error-reduction measures for reducing both accident risk and the cost of product failure.

1 Introduction

Humans are involved in many system operations and decision-making. Human errors may happen in any processes: design, manufacture, transportation, operation, maintenance, etc. Therefore, human errors can lead to system failures or even safety or environmental accidents. According to recent studies (Gordon, 1998) (Gertman, 2002) (Wiegmann & Shappell, 2012) (Dhillon, 2007), human errors contributed to an estimated 60 to 90 percent of all accidents across industries, and petroleum industry is no exception. In the Deepwater Horizon oil spill accident in 2010, one of the human errors was the incorrect interpretation of the negative-pressure test result (BP report, 2010). In the Piper alpha accident in 1976, one of the direct accident causes was a missing pressure safety valve, which was due to miscommunication among personnel (PatC-Cornell, 1993). While human error is concerned because of its contribution in accidents, it is also a pain due to its consequences in terms of economic loss, e.g. an omission of calibration leads to a failed job and thus incurs rig down time and loose profit.

Human error management is needed in the petroleum industry, in order to reduce the risk of accident and to reduce the cost of failure. The goal is to reduce and prevent human errors by applying lessons learned in the past, sharing lessons across from one product line to other product lines, from one site to other sites, and from one industry to other industries. This goal is achievable by good data collection systems to capture the lessons in the past, and Human Reliability Analysis (HRA) to synthesize the useful empirical / theoretical findings and to apply the lessons to future operations. This paper discusses how the petroleum industry can utilize the available HRA studies and human error databases for managing human errors.

With human error gradually getting more and more attentions, especially after several catastrophic accidents caused by human errors, there arose an assortment of HRA methods over the past three decades. HRA aims to identify possible human errors for given tasks, to assess the human error probability (HEP), and to identify error reduction measures.

Even though many of these HRA methods rooted in the nuclear industry or aviation industry, these methods preserve a good level of generality for being adapted for other industries; in this paper, we provide an overview of the available HRA work and discuss how the HRA methods are applied to or adapted for petroleum industrial operations for managing human errors.

The purpose of HRA or assessment is to support decision-making, and for managing human errors, two major decisions need to be supported by HRA:

- *Prioritize issues*. It is the decision regarding what human errors should be addressed and how to allocate resources and efforts to address each issue and in what order. A quantitative assessment is essential to support this decision. The decision maker needs to know the trend and distribution of the different human errors, the outcomes of the human errors (economic impact, environment impact, social impact, casualty and injuries, etc.), and the HEPs. This assessment could be an informative basis even for the decision of taking no actions.
- *Determine the error reduction measures*. It is the decision regarding how to reduce the human errors and to mitigate the consequences. Qualitative analysis needs to provide the following information for a good understanding of the human errors, which is necessary for identifying effective error reduction measures: the error contexts, the root causes, the factors that influence the human errors, barriers that prevent the human error cascading higher level of failures, and error detection mechanisms.

Figure 1 shows a general framework of a human error management system. The process starts with identifying human errors in the operation tasks of interest. Then, the analysts quantitatively assess the HEPs, in order to prioritize the issues and to allocate human and capital resources. For the selected human error issues, the analysts determine effective measures to reduce human errors, e.g. enhancing training, improving maintenance rules, and optimizing work processes. The implementation of the selected measures will influence the failure incidents, which could be collected in a human error database. The database provides real data insights for the error identification, assessment, and error reduction measures. Monitoring the new failure data can evaluate the effectiveness of the error reduction measures.

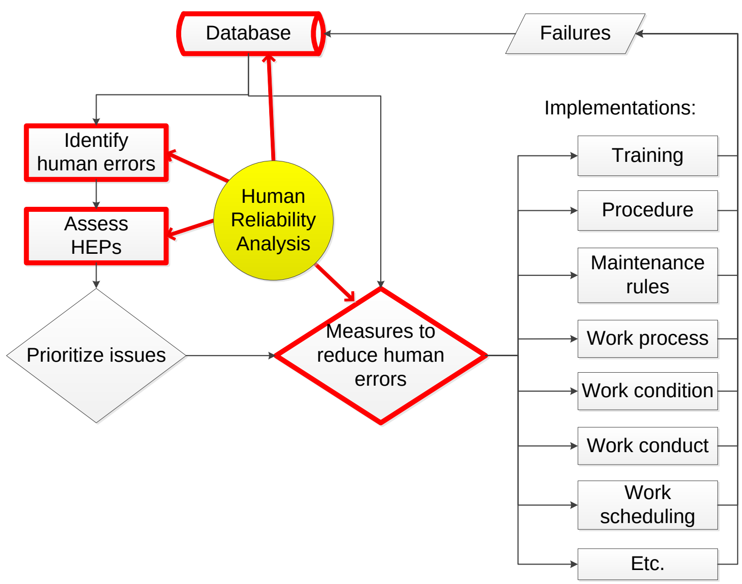

Figure 1 A generic framework of human error management system

This paper discusses how the petroleum industry could utilize the available HRA work for the highlighted processes in the human error management framework shown in Figure 1. Section 2 presents the techniques for human error identification in a prospective way; Section 3 introduces the techniques for human error quantitative assessment and error reduction measures; Section 4 discusses human error database structure and some desirable features; Section 5 suggests some prospective work for the petroleum industry for managing human error; and the last is the conclusion.

2 Human error identification

This section introduces the main steps for human error identification and the available HRA methods useful in Section 2.1 and 2.2. Comments on implementing HRA for error identification in the petroleum industry are presented in Section 2.3.

2.1 Task analysis.

Human errors could be identified in two ways: human error data analysis and/or task analysis. Real data provide valuable information that points to the weak links in the operations, based on the occurrences of human errors in the past. However, task-specific human error data are often scarce and do not cover a good spectrum of human errors. When direct data are not available or not enough for error identification, HRA starts with a task analysis for human error identification.

Task analysis is a descriptive approach to analyze how a task is accomplished and what human operators' functions are in the task. In addition, task analysis requires identifying the following elements of the task: operation goals, human-system interactions, and human-human interactions for achieving system goals, working environment, system conditions, situational contexts, and the resources available for accomplishing the task. The analysis identifies the possible human errors by studying how the human fails in the identified human functions or human activities in the task.

Task analysis requires a good knowledge of specific tasks, which is acquired by observing real operations (e.g., field operation, experiment operation, and simulator operation practice) consulting or interviewing subject matter experts, e.g. engineers/technicians with corresponding operation experiences, design engineers, etc. We often use a walk-through or a talk-through for eliciting information from experts. Some available documents that describe the task of interest are also used as inputs for the task analysis: e.g., standard procedures, manuals, statement of requirement of the tool in use, and work process flow chart. The task analysis process can be coupled with some other approaches: e.g., Process Failure Mode and Effect Analysis (PFMEA), and Hazard and Operability study (HAZOP).

During the task analysis, the analyst usually breaks down the task into subtasks to analyze the detail and to make the analysis organized. This breakdown process continues until a stop-rule applies, resulting in a hierarchy structure of the task. The analyst controls the level of detail, based on the purpose of the analysis and the resource limitation. These are several stop-rules of task decomposition:

- The analyst has a good grasp of the subtask defined at the lowest level and it provides enough information to enable human error identification.
- An HRA has been done for the subtask at the lowest level and it can be used.
- The consequence of human failure in this subtask is trivial.

Figure 2 provides an example of task breakdown of control room operation in a pressurizer power plant. The top goal of this operation is to properly handle a Steam Generator Tube Rupture (SGTR) accident in the control room operation. The task is broken down to four subtasks: correctly diagnose the SGTR accident, isolate the rupture steam generator, cool down the reactor coolant system, and depressurize the reactor coolant system to stop the radioactive reactor coolant leaking to the ruptured steam generator. These four subtasks can be further broken down to smaller items as needed. Figure 3 shows another example of breaking down the task of setting a bridge plug in a completion operation in an oil rig.

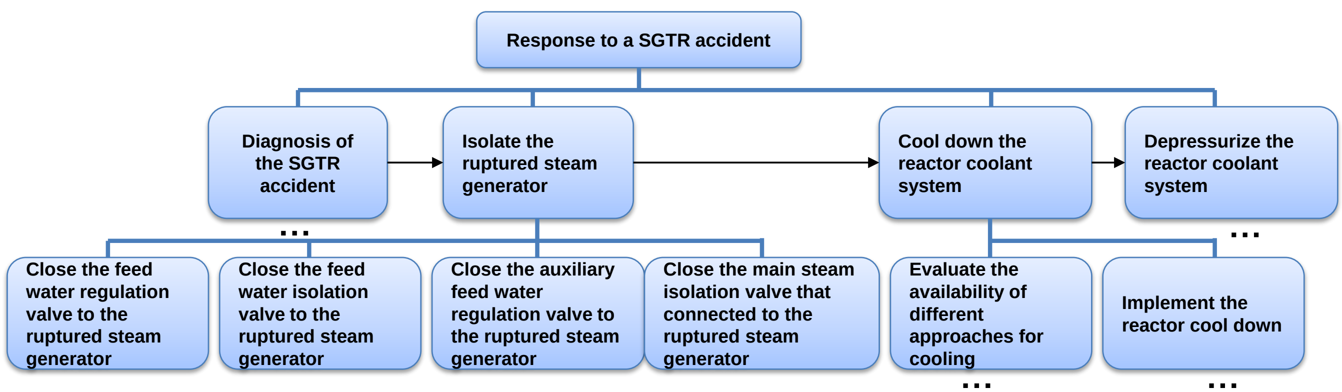

Figure 2 Example: task breakdown of nuclear power plant control room operators responding to a SGTR accident

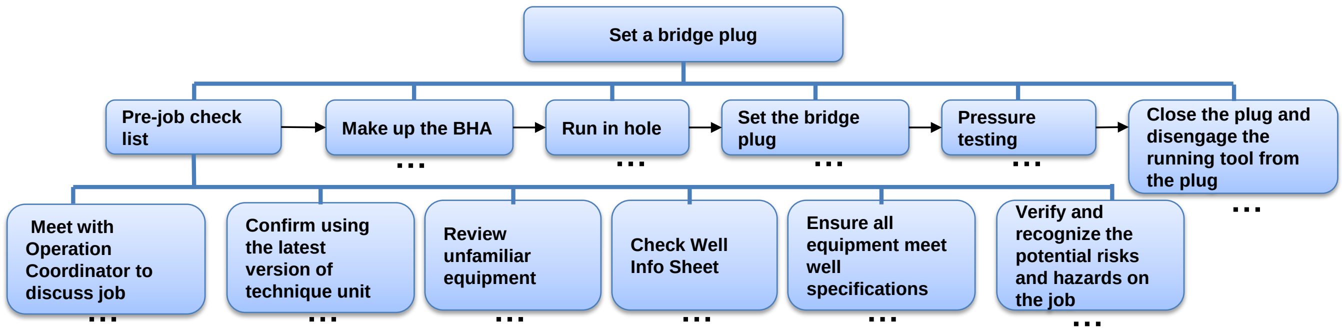

Figure 3 Example: task break down of setting a bridge plug

2.2 Human error contexts and factors.

Some HRA methods[1] provide guidance for the analysts to collect relevant information from task analysis, and summarize general human error contexts and the factors that contribute to the human error. When applying the HRA method to a specific task, analysts map the task of interest to the task categories in the HRA method and instantiate the general discussion to identify potential human errors and its causal / contributing factors. Example HRA methods are ATHEANA (U.S. NRC report NUREG-1624, 2000), THERP (Swain, 1983), HEART (Williams, 1992), and CREAM (Hollnagel, 1998). A few simulation HRA approaches simulate operators' operational behaviors for given task situations, and thus predicts potential human errors by the simulation programs, e.g. ADS-IDAC (Chang & Mosleh, 2007) and CES (Woods, et al., 1987). Even for the simulation HRA methods, a task analysis is still required in the pre-simulation analysis, in order to gather the required information for coding the simulation input, e.g. the procedures that the operator might use, the operator's knowledge that might be used during the operation, the system's initial state and any hardware failure that complicates the situation, etc.. The currently available simulation HRA approaches are dedicated for the applications in nuclear or aviation industries for risk significant task only. Though these simulation HRA methods cannot be directly applied for petroleum industry applications, the simulation techniques can be used to develop similar codes for risk significant operation tasks in petroleum industry.

For HRA methods that rely on the analysts to identify the human errors, the error identification can be carried out by converting the identified human activities to potential human errors. Generally, two types of human errors apply to each human activity: error of omission and error of commission. Error of omission means missing a task or subtask. Error of commission means doing the task incorrectly.

It is not easy to have a good coverage of possible omission errors. An essential step of analyzing the error of omission is to identify the cues that prompt the operator to initiate the subtasks. For example, procedures or manuals explicitly instruct the operator/technician step by step. The subtasks not specified in formal instruction documents are more susceptible to be omitted, because it requires the operator/technician's extra efforts, experience or knowledge to do it, which introduce variation in the reliability of the work. The effectiveness of the reminding cues should also be analyzed. For example, memos and technic bulletins occasionally serve as provisional additions or modifications to the procedure. This adds one more link in the task—the operator/technician needs to remember to use these updates, which is deviating from the familiar routine of using procedures only. The successful use of these updates highly depends on the management efforts, e.g. how well these updates are communicated with the operators/technicians. For error of commission, it is a very open-ended process to specifically describe the error, simply because human can do a task incorrectly in many different ways.

Some HRA methods categorize human tasks into sets of predefined types and provide discussions of how these tasks could fail in each category. Some categorizations are generic, and some are domain specific. Table 1 summarizes the generality of 5 widely used HRA methods in task categorization.

Table 2 shows the categorization of human failures in ATHEANA method as an example. In ATHEANA, human behavior is analyzed in four cognitive phases: detection, situation assessment, response planning, and response. Some common human failures are summarized in ATHEANA based on accident data in the nuclear industry. In the third column of Table 4, we provide some human errors in the petroleum industry to demonstrate that how the ATHENA schema can be applied for human tasks in the petroleum industry.

[1] Some HRA methods are dedicated for quantitative assessment of human errors only, e.g. SLIM.

Table 1 Task Categorizations of several HRA methods

HRA method	Task categorization
THERP	Tasks are summarized for nuclear power plant applications. A lot of them can be translated for tasks in other industries with some adjustments in the descriptions.
HEART	8 generic task categories
JHEDI	9 generic task categories
CREAM	15 generic task categories
ATHEANA	The method decomposes human behavior into 4 generic components. Human error types are summarized for each component from nuclear industry accidents/incidents. They can be translated for tasks in other industries with some adjustments.

Table 2 ATHEANA categorization of human failure types and some examples

Failures in 4 categories	Failures examples from accidents in other industries in ATHENA (U.S. NRC report NUREG-1624, 2000)	Petroleum industry human error examples
Detection failure	• Operators unaware of actual plant state, its severity, and continued degradation in conditions.	• The operator fails to detect the high vibration level exceeding the specification. • Fail to monitor the mud quality; debris in mud flow damages the tool. • Fail to detect a part has moved away from its designated position during assembly. • Fail to detect a damage induced during shipping. • Fail to detect a crack during inspection.
Situational assessment failure	• Operators develop wrong situation model (or cannot explain) plant state and behavior. • Operators unable to distinguish between results of their own actions and accident progression. • Operator misinterpret information or are misled by wrong information, confirming their wrong situation model • Operators reject evidence that contradicts their wrong situation model.	• Incorrectly interpret the negative pressure test results for cementing evaluation. • Fail to recognize the tool has reached its hour limit and the need to switch to a backup tool. • Fail to recognize the selected tool mode is improper for the well. • Fail to realize that wrong data have been entered to tool maintenance system due to the negligence of data unit conversion. • Fail to tell the signal information is incorrect.
Response planning failure	• Operator select non-applicable plans, which direct operators to take actions that are inappropriate for specific situation • Operators follow prepared plans that are wrong or incomplete • Operators do not explicitly use prepared plans • Prepared plans do not exist, so operators rely upon knowledge-based behavior • Operators inappropriately give priority to one plant function over another	• Risky choice: not maintain the tool due to tool shortages and low revenue market. • Use a procedure that is not appropriate for a cold weather condition. • Incorrectly use a part different from the one specified in the manual as a substitute, due to out of stock of the correct part. • Do not use the procedure to recalibrate the tool for different casing size • Forget to reset tool parameters before use. • Continue to use a tool when it reaches a maintenance point. • Risky decision of delaying corrective actions.
Response implementation	• Important procedure steps are missed • Miscommunication • Equipment failures hinder operators' ability to respond	• Miss small parts (e.g. O-rings and bolts) during assembly. • Damage the O-ring during installation. • Tool is not cleaned thoroughly. • Over torque and damage a tool. • Tools are lost in the hole due to late response. • Poor soldering job.

During the error identification process, a major focus is to identify the contexts that are error prone – error forcing context (EFC). Inevitably, it involves analyzing the causal factors of human errors and the factors that affect the likelihood of human error – performance shaping factors (PSF)[2]. Several HRA methods have summarized sets of EFCs from empirical data and expert judgment. As an example, the EFCs in HEART method are provided in Table 3. When practicing HRA, the analysts can compare the EFCs summarized in HRA methods against the specific task/sub-task of interest, to examine whether the task/subtask matches any EFC, using the task information identified in the task analysis.

[2] It is also called Performance Influencing Factors (PIF) in some literatures. There are slightly different interpretations of these two terms. PIF only influences the chance of human errors but is not the major cause, while the word "shape" renders a stronger meaning than "influence". However, these two terms are used interchangeably in many places.

Table 3 Error Forcing Contexts summarized in HEART method

- Unfamiliarity with a situation which is potentially important, but which only occurs infrequently or which is novel
- A shortage of time available for error detection and correction
- A low signal-to-noise ratio
- A means of suppressing or overriding information or features which is too easily accessible
- No means of conveying spatial and functional information to operators in a form which they can readily assimilate
- A mismatch between an operator's model of the world and that imagined by a designer
- No obvious means of reversing an unintended action
- A channel capacity overload, particularly one caused by simultaneous presentation of non-redundant information
- A need to unlearn a technique and apply one which requires the application of an opposing philosophy
- The need to transfer specific knowledge from task to task without loss
- Ambiguity in the required performance standards
- A mismatch between perceived and real risk
- Poor, ambiguous or ill-matched system feedback
- No clear, direct and timely confirmation of an intended action from the portion of the system over which control is to be exerted
- Operator inexperience (e.g. a newly-qualified tradesman, but not an 'expert')
- An impoverished quality of information conveyed by procedures and person-person interaction
- Little or no independent checking or testing of output
- A conflict between immediate and long-term objectives
- No diversity of information input for veracity checks
- A mismatch between the educational-achievement level of an individual and the requirements of the task
- An incentive to use other more dangerous procedures
- Little opportunity to exercise mind and body outside the immediate confines of a job
- Unreliable instrumentation (enough that it is noticed)
- A need for absolute judgments which are beyond the capabilities or experience of an operator
- Unclear allocation of function and responsibility
- No obvious way to keep track or progress during an activity

PSFs can causally lead to human errors (e.g. *lack of raining, high task complexity*), or influence human error as modifier—enhance or degrade human performance (e.g. *stress*). Using of PSFs gives analysts great flexibility to include what they believe to be relevant and important for the study in addition to the EFC descriptions. EFC might be very situation specific, while PSFs are more general across different tasks. PSFs used in HRA methods cover a wide range of factors. They include one's psychological feeling or state (e.g. *fatigue*), one's personal capability limits (*knowledge/training*), task attributes (e.g. *task complexity*), environmental factor (e.g. *temperature*), ergonomic factors (e.g. *human system interface quality*), organizational factors (e.g. *safety culture*), and team factors (e.g. *communication*). Several motivations that drive the use of PSFs: 1) capturing the situation characteristics that cause or foment human errors; 2) capturing the personal/crew/organization characteristics that cause or foment human errors; 3) reflecting the individual/crew differences.

In the human reliability area, there is no standard PSF set. Human reliability analysts use different sets of PSFs for different HRA methods and applications. Choosing the most effective PSFs and having a good coverage of important PSFs are critical for HRA practice. This selection process can be data-informed or based on the analysts' judgment. Gorden summarized sets of PSFs that were effective in offshore accident in the petroleum industry (Gordon, 1998). Table 4 shows an example of a generic PSF set summarized by Groth in 2012; it covers a wide range of PSFs in five groups: organization factors, team factors, person factors, situational factors, and human machine interface factors.

During practice, the analyst might need to expand some of the core PSFs into more details. For example, in many operations, there are procedures to guide the operator to accomplish the task. Inadequate procedure quality strongly affects the failure chance of the operation. Thus, it is worthy to analyze the PSF *procedure* into more detail. For example, the list below elaborates how the procedure quality can be inadequate in different ways:

- The procedure does not capture a step that is necessary for successful operations. E.g., the procedure does not

include a necessary inspection for a failure prone part.
- The procedure guide is vague and does not provide enough details. E.g., the procedure does not specify how much torque that needs to be applied for installing a tool part.
- The procedure instruction is confusing and difficult to understand. E.g. the procedure contains double negation logic.
- The procedure instruction is misleading or incorrect. E.g. the procedure specifies an inappropriate length of wire for assembly.
- The procedure is not user friendly. E.g. there is no visual picture to assist the technician to identify the correct parts for use.
- The available procedure does not apply to a new situation that is different from the usual operations. E.g. the procedure does not apply to a cold weather working condition.

Table 4 PSF classification (Groth & Mosleh, 2012)

Organization-based	Team-based	Person-based	Situation/stressor-based	Machine-based
Training program	Communication	Attention	External environment	HIS
Availability	Availability	To task	Conditioning events	Input
Quality	Quality	To surroundings	Task load	Output
Corrective action program	Direct supervision	Physical & psychological abilities	Time load	System
Availability	Leadership	Alertness	Other load	response
Quality	Team coordination	Fatigue	Non-task	
Other programs	Team cohesion	Impairment	Passive information	
Availability	Role awareness	Sensory limits	Task complexity	
Quality		Physical attributes	Cognitive	
Safety culture		Other	Execution	
Management activities		Knowledge/experience	Stress	
Staffing		Skills	Perceived situation	
Scheduling		Bias	Severity	
Workplace adequacy		Familiarity with situation	Urgency	
Resources		Moral/motivation/attitude	Perceived decision	
Procedures			Responsibility	
Availability			Impact	
Quality			Personal	
Tools			Plant	
Availability			Society	
Quality				
Necessary information				
Availability				
Quality				

Some early HRA methods' structures or schemas are mainly based on limited empirical data and expert judgment e.g. THERP, HEART. A few HRA methods later have built behavioral/cognitive models for the analysis, e.g. ADS-IDAC, CREAM, etc. There has been a need to make connections between HRA methods and the available behavioral / cognitive / social psychological theoretic findings. U.S. Nuclear Regulatory Commission has pushed a research effort of synthesizing psychological theories into a general cognitive framework, in order to provide a cognitive basis for HRA methods for improving the HRA quality. The product of this effort was documented in the report NUREGU-2114 "Building a psychological foundation for human reliability analysis". In this framework, it discusses the human activities in 5 components: detecting and noticing, understanding and sense-making, decision-making, actions, and team coordination. For each of the 5 components, this document summarizes its cognitive mechanisms, the ways how it can fail, and PSFs for this specific cognitive component. Table 5 shows a sample discussion of a type of cognitive failure excerpted from NUREG-2114.

Table 5 Incorrect integration of data, frames, or data with a frame (adapted from NUREG-2114)

Failed cognitive component	Understanding and sense-making
Failure mode	incorrect integration of data, frames, or data with a frame
Mechanism	Perceived external information not properly recognized, classified, or distinguished
Discussion	As information is perceived, more meaning is applied to it. This is the blurry boundary between Detecting and noticing and Understanding and Sense-making. This mechanism refers to classification or categorization of information at the recognition level, before the information is more deeply or consciously processed. Once a scene is identified (perceived), top-down influences from existing knowledge on identification of aspects of the situation begin. Operators can get the "gist" of a situation very quickly, and that "gist" may be inaccurate if information has not been properly classified. If information is improperly categorized, as irrelevant, for example, or as oversimplifications, then operators will not make proper use of the information, and understanding of the situation will suffer. This misclassification can occur due to

	misperceiving the information, or misinterpreting the information based on existing knowledge.
Example	For example, if a particular alarm has a history of sounding spuriously, operators may be likely to dismiss it as a false alarm if it activates in the case of a real problem. Another example is the case of operators misreading procedures. Crews attend to the procedures while performing actions. During this time the crew will judge whether the strategies embodied in the procedures are appropriate to the situation. The crew may make a mistake and inappropriately misjudge the procedures as inapplicable due to misreading the procedures.
Relevant PIF(s)	• Knowledge/ experience/ expertise • Training • HSI Output • Procedure quality

2.3 Comments on implementing HRA for error identification in the petroleum industry.

Error identification could be performed by using applicable human error data, HRA, or a combination of data and HRA. HRA is helpful for human error identification especially when data are not available or partially available, for some tasks of which the failure can cause catastrophic accidents: e.g., operation of safety equipment like BOP, and operators' response to emergent situations.

Section 2.1 and 2.2 reviewed how HRA methods guide the analysts to identify human errors through task analysis. A task analysis breaks down a task into sub-tasks and identifies the required human functions. Then, the taxonomies of HRA methods serves as road maps to assist the analysts in the following steps: matching the task (or sub-task) of interest to a "general" task type defined in the method; translating the identified human functions to potential human failure modes; and gathering relevant information during the task analysis to determine the error enforcing contexts, error mechanisms, and effective PSFs. Even though many HRA methods are nuclear power plant oriented, the taxonomies can be adapted for petroleum industry applications.

It is important to realize that there is no HRA method that provides exhaustive discussion of every possible failure mode, failure cause and failure condition. Different HRA methods vary in their taxonomy structures (task type, context information, and PSFs), coverage, resolution of detail, and complexity to use. For example, the THERP method has a rough discussion of diagnosis errors; CREAM provides more cognitive detail for analyzing diagnosis errors; ATHENA offers in-depth process for identifying human errors in different cognitive stages. Review and comparison of different HRA methods can be found in these literatures: (Kirwan, 1996), (Forester, et al., 2006), (U.S. NRC report NUREG/IA-0216, 2012), (Spurgin & Lydell, 2002), etc.

For the error identification phase, the analysts do not have to rely on only one HRA method. One might find it is desirable to compile useful information from several HRA methods and the available data to form a new taxonomy scheme (task type, human error modes, error contexts, error causes and PSFs) based on the specific application. The governing rules are ensuring a good coverage of error identification, and keeping the analysis effort affordable based on the available resources (e.g. available HRA expert in house, time and cost for the HRA task). Once a structured taxonomy scheme is determined, the analysts use it as road map for identifying potential human errors for the tasks of interest. The same taxonomy scheme can serve as an index for root cause analysis of human errors in a data reporting system.

3 Human error quantification assessment and error reduction measures

3.1 When is the assessment of Human Error Probability (HEP) needed?

For those tasks that possibly result in catastrophic consequences (e.g., core damage accident of nuclear power plant, blowout of oil well), probabilistic risk assessment (PRA) of the accident scenarios is needed to inform the prospective and quantitative risk level. When human interactions are identified that they contributes to accident scenarios, a quantitative assessment of human errors is necessary. The PRA model considers human errors, hardware component / system failure, and software failure in the accident contexts, and calculates the accident probability based on the hardware failure probability, software failure probability, and HEP. Fault tree, event tree, and binary decision diagram are widely used approaches for quantitative PRA.

In addition to the necessary PRA for significant accident scenario, quantitative assessment of human errors helps prioritizing resources to tackle different human errors.

3.2 HRA methods for assessing HEP.

HEP is defined as the ratio of the number of human errors over the number of human error opportunities. When data are not available, some HRA methods provide an approximate assessment. There are several styles to calculate HEPs:

- The method provides base HEP for pre-classified task types/human error modes/situations. The analysts assess the level of each PSF (or error-producing conditions) in the pre-established set, then use the PSF assessment to adjust the base HEP to get the task HEP. (E.g. THERP, HEART, SPAR-H (U.S. NRC report NUREG/CR-6883, 2005), CREAM extended method.)
- HEP is calculated based PSF assessment and contextual information with specific rules. E.g. CREAM basic method.
- Use decision tree to match the task to a pre-determined HEP. E.g. CBDT (U.S. NRC report NUREG/CR-1984, 1984).
- Use available HEP data to calibrate equation and extrapolate the HEP for another specific task of interest. E.g. SLIM (EMBREY, et al., 1984).

When use these quantitative methods to the petroleum industry applications, several things should be noted. If the methods provide base HEP, the analysts should verify the generality of the task description and its intended application, to determine whether it is applicable for the task of interest. Secondly, if the method uses a predefined PSF set or contexts categorization, the analysts should verify whether the important PSFs or likely error-prone contexts are included in the set.

3.3 Human error reduction measures.

Devising human error reduction measures is driven by human error identifications and investigation of human error mechanisms and causes. Generally, human errors can be controlled in these three ways: 1) eliminating the causes and factors that induce human errors, e.g. fixing the confusions in a procedure, and simplifying an assembly process by design changes; 2) adding error detection mechanisms, and correcting the errors before failure happens; 3) mitigating the consequence of human errors, e.g. adding safety barriers. Some examples are provided in Table 6.

Table 6 Human error reduction measure examples

Eliminate the human error cause or reduce the error chance	Increase the error detection rate before failure happens	Mitigate the consequence of human error
• Update procedure to provide accurate instruction; • Enhance technician training; • Revise the design to make a human error impossible; • Improve housekeeping and inventory management; • Improve compliance to procedures.	• Increase awareness of possible errors in a process and improve the self-check by lesson sharing among technicians; • Have a second person to check the work.	• Have filters/screens to catch possible debris in the circulation to avoid jamming the turbine.

A common pitfall of human error management is resolving human error locally and failing to apply the learned lessons broadly. Once the cause and mechanism of one human error is identified, questions should be asked to identify situations where similar error might happen. For example, an assembling error might repeat in other maintenance shops, and in other product lines; a field operation error might repeat at other operation sites; etc. The impact range of the same error mechanism and cause should be taken into consideration for error reduction measures. For an organization or a company, it can be very beneficial to build a knowledge base that summarizes the likely human errors in different work processes, to facilitate sharing lessons across different branches of the organization and to accumulate experience over time. Some HRA literatures provide some general solutions corresponding to different human error contexts patterns, which are also applicable to the petroleum industry. For example, Barry summarized a set of human error mechanisms and ideas of error reduction measure (Kirwan, 1994).

4 Human error database design

A good human error database is particularly useful for human error management. Real data contains first-hand information for the qualitative analysis and quantitative analysis. It provides better understanding of the human errors for the tasks of interest; enables trend analysis for highlighting the weak links in human functions; supports developing error reduction measures; and assists lesson learning and sharing to prevent similar error from recurring.

4.1 Main data content and data structure.

An important question for designing a human error database is what event should be considered and recorded as a human error. Several terms are used interchangeably with human error in literature: human failure event, unsafe act, mistake, etc. For the purpose of collection information and improving task success rate, here human error is defined as human’s action or inaction that fails to achieve the desired function and results in negative impacts. This definition is consequence-oriented only. It does not intend to judge whether the person/team who commits the error should be blamed or not, which is a different topic. As a matter of facts, often the root cause of human error lies deeper in the management or organization, instead of in the front line person. In order to encourage collecting full and accurate information, it is important to note that the purpose of the database is to improve but not to blame.

Though we tend to pay attention to the human errors manifested by its observable negative consequence (e.g. rig downtime, or accidents), there are still a large portion of human errors not detected or recorded, including near misses, latent errors corrected before it takes effect, and some errors not exposed at all. Once these hidden errors are discovered, they are also valuable information sources, especially when the data inventory is low. Lessons from the hidden errors should also be learned to prevent them from recurring and leading to loss.

In most cases, accidents are results of combinations of multiple errors or failures. As demonstrated in Reason's Swiss Cheese Model (Reason, 1990) shown in Figure 4, each hole represents a failure, error, or system bug. An accident happens when multiple holes at different levels align. It is likely that there are more than one human errors contributing to one incident / accident. For example, a tool fails down in the well due to one mechanical component was incorrectly manufactured. There are several opportunities to detect this component fault (inspection at manufacturing site and full dimensional inspection during the assembly process) before the tool was sent to the well. Thus, other than the initial manufacturing error, there are also two other inspection errors. These errors should also be recorded in the database, and the database structure should enable recording multiple human errors for one incident and it is desirable to have them related. In addition to assessing the HEP, this type of data is also useful for evaluating the effectiveness of the current error detection measures.

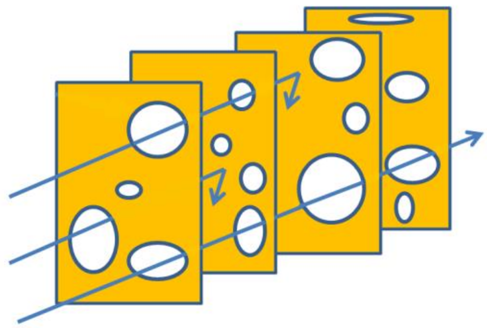

Figure 4 Swiss Cheese Model

For managing human error, the main objectives of a human error database are:

- Track human errors and record information to enable trend analysis
- Capture information needed for developing error reduction and error prevention measures

To achieve these objectives, the database should record the following basic information for each human error:

- Description of the human error, which explicitly describes what and how the human activity or inactivity was deviating from a desirable response;
- Causal and influencing factors: the human error context, which defines the challenges to the human performance in the specific task context; and other factors, which affect the human performance and hinder one from meeting the challenges;
- The consequence of the human error, which describes how the human error affects the system;
- If any, description of the human error recovery—under what circumstance this human error was detected and recovered.

Human error investigation is similar to the human error prediction; one is retrospective process while the other is prospective. Like a good HRA taxonomy structure (categorization of tasks, human error contexts, and PSFs) can provide a map for the analyst to explore the possible errors and their causes, a similar structure can facilitate the error investigation process by guiding the investigator to verify the possible causes. It provides a predefined framework for examining the human error, which can improve the robustness of the data quality, by reducing the variations among investigators in terms of incomplete information recording.

Accordingly, the taxonomy structure should have a good coverage of the relevant factors for the intended tasks. As mentioned in Section 2.3, there is no taxonomy structure universally adequate for every application, because human errors are very situational and vary a lot among different tasks and contexts. However, existing HRA methods provide a lot of useful example taxonomies and discussion. The selection of taxonomy should be based on the application needs. Modifications and combination of the available taxonomies might be desirable to ensure an adequate fit for the application needs. Also, the usability and efficiency of the taxonomy structure should be taken into consideration. The complexity of the taxonomy should be controlled within an acceptable level. The taxonomy can be designed with different levels of detail. The most effective factors for the application should be elaborated at finer grain of detail. Also, it would be necessary to update and optimize the taxonomy structure with feedbacks from the database usage.

One human error database example is Human Event Repository and Analysis (HERA) (Hallbert, et al., 2006). It is for collecting human performance information from commercial nuclear power plant, developed by the U.S. Nuclear Regulatory Commission and Idaho National Laboratory. Except for documenting the general incident description and some basic information (plant information, location, time, etc.), the database system records each incident by a series of key events in a time line order, including the plant system phenomena and the crew's interactions with the plant system. HERA records the human errors and the context in a timeline style. When the investigator registers a new human error to the database, he (or she) is asked to provide a description of human error, and to classify the human error by selecting the applicable options provided in a pre-defined taxonomy frame, which includes the task type, human error type, corresponding cognitive processes, the cognitive level, and PSFs. The top level PSFs in HERA include: available time, stress & stressor, complexity, experience & training, procedure & reference documents, ergonomics & human machine interface, fitness for duty / fatigue, work processes, planning / scheduling, supervision / management, conduct of work, problem identification & resolution / corrective action plan, communication, environment, team dynamics / characteristics. Each PSF is elaborated at finer grain of detail. We excerpt one PSF's data sheet here as an example, shown in Table 7. When entering data, the user checks the effective PSFs and their contributory factors. The comment column allows the user to write more case specific explanation for each factor.

Table 7 PSF data structure example of HERA database

PSF	Negative Contributory Factor	Source/Inference	Comment
Experience & Training	□ Fitness for duty training missing/less than adequate	□Source □Inferred	
	□ Training less than adequate	□Source □Inferred	
	□ Training process problem	□Source □Inferred	
	□ Individual knowledge problem	□Source □Inferred	
	□ Simulator training less than adequate	□Source □Inferred	
	□ Work practice or craft skill less than adequate	□Source □Inferred	
	□ Not familiar with job performance standards	□Source □Inferred	
	□ Not familiar with tools	□Source □Inferred	
	□ Not qualified for assigned task	□Source □Inferred	
	□ Training incorrect	□Source □Inferred	
	□ Situation outside the scope of training	□Source □Inferred	
	□ Others:__	□Source □Inferred	
	□ None / Not applicable / Indeterminate	□Source □Inferred	

4.2 human error databases for lesson learning and sharing.

One of the most important functions of human error databases is to enable lesson learning and sharing, for preventing same or similar errors from happening again. Many human errors repeat, even across different product lines, due to their commonality in error-driving factors. A desired error reduction measure should be applicable and effective to many similar human errors, instead of solving the problem locally one by one. To achieve this, the challenge is to capture the commonality among different human errors and to discover trends or patterns.

***4.2.1 Use taxonomy to capture commonalities*.**

Commonalities of human errors could lie in many different aspects: similar working environment, similar work process, similar work object, same worker or worker receiving the same insufficient training, same work culture, etc. An important objective of the database's taxonomy is to capture the commonality of errors to an extent that it would get good counts to indicate the trends and also the commonality is not too vague to be useful for identifying the error reduction measures. Without an efficient taxonomy, the necessary information for identifying error commonalities might be recorded in the text description of the human errors; however, it could be very cumbersome for the analysts to dig through each description to try to generalize the commonalities. This is a widespread problem existing in many databases, and it directly leads poor or low efficient use of the database. There are various taxonomies from different HRA methods or HRA applications as discussed earlier in this paper. One should choose an appropriate taxonomy or make up one, in the light of the actual application characteristics. It is difficult to come up with a perfect taxonomy at the beginning of the database design, especially when there is limited understanding of the human errors to be collected. Hence it is desirable to update the taxonomy later with analysis of and feedback from the coming data. Example changes are adding categories to the taxonomy, and subdividing a category to several sub-level categories in order to capture more elaborated detail. Other detail that is not captured in the taxonomy should be recorded as texts in the database: e.g. comments, remarks, or descriptions.

***4.2.2 Use key error attributes to capture the trend*.**

In addition to using taxonomy, the database can also provide some open-ended fields for recording brief summary of some specific aspects that are useful for trend analysis. For example, use one database field to record the directly affected objects (e.g. damaged component, system) as a result of human errors; the benefits are to use the counts of each affected object to find error trends, to summarize repeating human errors and to identify the error-prone task steps. Let us take an example to illustrate the benefit. A human error is technicians improperly assemble tools and damage o-ring. The damaged o-ring leads to seal failure during operation and future causes some electronic system failure. Some companies or organizations have incident databases to track the failed systems, but they do not necessarily record the failure at the lowest component level (o-ring). The damaged o-ring incidents might appear as electronic pad failure in one incident and circuit board failure in another incident, thus these o-ring incidents could not be associated easily for the trend analysis. If the human error database has a separate field to record the directly affected object by human errors, o-ring would be entered in that field for these incidents and the trend can be easily discovered.

4.2.3 Commonality might lie in deep root causes.

When investigating the human errors, the cause could often be traced back to different levels. For example, factor A is the cause of the human error 1, factor B is the cause of factor A, and factor C is the cause of factor B. A common mistake is that the database only records the proximate cause and misses the deeper levels of causes. Often the data do not provide enough information for identifying a good error reduction measures. Another concern is that shallow information might prevent the discovery of trends. Continuing with the above example, factor D is the cause of human error 2, and factor B is the cause of factor D. Without recording the deeper level cause factor B, human error 1 and human error 2 appear to be unrelated in the data system. Table 8 gives a real life example. The human error is that a drilling system was improperly torqued. It leads to system failure and rig down time. The direct cause is there was no proper torque gauge at the rig site to ensure proper torque applied. A deeper cause is lack of communication between the contractor and the client regarding the job preparation (torque gauge). A solution at the direct cause level could be providing a torque gauge at this specific rig site, while a solution at the deeper cause level could be formally including the tool preparation in the communication process between the contractor and the client. The former solution only solves the problem at one rig site, while the later solution can fix the similar problems at multiple rig sites and be more effective in preventing recurring errors.

Table 8 Example of multiple levels of human error causes

Second level cause →	Proximate cause →	Human error →	Consequence
Lack of communication between the contractor and the client regarding the job preparation	A torque gauge is not provided at the client's rig site	A drilling system was improperly torqued	Rig downtime. Pull the tool out of hole and have a second trip.

The database should also capture the influencing among PSFs. An efficient PSF set should make the PSF structure orthogonal. In another word, there is no overlap in meaning between different PSFs at the same level. This eliminates the vagueness and makes it clear for the data recorder to select the effective PSFs, a good example is Katrina's PSF taxonomy. Even though the PSF structure can be orthogonal in terms of PSF meanings, there is still dependency among PSFs. Take the human error provided in Table 8 as an example. If we use Katrina's PSF taxonomy to record this human error, two PSFs should be checked: organization based—resource-tool—availability, and team based—communication—quality. In this case, the poor communication quality causes loss of tool availability. This causal relation should also be captured in the database. Capturing the dependency among PSFs is helpful in extracting error patterns.

4.2.4 Support HEP assessment.

HEP is calculated as the ratio of the number of human errors over the number of human error opportunities. For example: HEP of a specific assembling error = Number of the assembling errors / Total number of the assembly jobs. While a human error database provides the error counts for specific tasks, the total number of human error opportunities also needs to be recorded in order to assess HEP. This information might be available in some other data sources, e.g. Systems Applications and Products (SAP) used by many companies. If a human error database is intended for different tasks in different product lines (e.g. manufacturing, maintenance, field operation; drilling, evaluation, completion, production, refinery, etc.), the database should be designed in a way that the human error data for a specific task can be easily extracted.

4.2.5 Automate the error reduction process.

Another desirable feature is that some business intelligence technologies could be used to partially automate the human error management system, in order to expedite the communication and problem-solving. Some general error reduction measures can be mapped with some known error types, for example, one error type "the procedure is inappropriate to this situation" has an obvious solution—"modify the procedure accordingly"; one error type "the required knowledge is not covered in the training" has an obvious solution—"update the corresponding training course material". Once these error categories are selected for new human errors, an intelligent database system asks the investigator to specify / select the relevant procedure or training course in the database, and automatically sends notification to relevant personnel accordingly (e.g. the ones who are in charge of maintaining procedures, and the ones who are in charge of the training courses) for reviewing this new human error.

5 Suggestions for prospective work in the near future

Most companies in the petroleum industry have their own databases to record historical failures. Among the failures, many are due to human errors, which may result in a significant profit loss. So far, little effort has been taken to systematically tackle this problem. Reducing human errors potentially achieves significant economic benefits, not to mention safety improvement. It is time to face this issue by taking advantage of the available HRA work in the human error management framework presented in this paper: to get a better understanding of the human errors in operations and to prevent them in future.

The petroleum industry has an abundance of operational failure data. Unfortunately, many databases are ill-suited for recording human errors. As a starting step of managing human errors in a company, one can examine the current data quality for analyzing the human errors in the past and extract useful information if possible. After getting a preliminary understanding, next step is to design a dedicated human error database for better data quality and efficiency by applying the suggestions provided in Section 4 of this paper. Then, summarizing error-prone task steps in a knowledge base is highly recommended for learning the lessons from the past and preventing the same or similar human errors in the future. As for the petroleum society, developing a synthesized human error database across companies will benefit the industry as a whole.

6 Conclusion

This paper provides an overview of human error management, which is necessary for reducing the accident risk and failure cost. The human error management is broken down to several major components for discussion: human error identification, human error assessment, human error reduction measures, and a human error database. This paper introduces the available HRA techniques for these major components, and discusses how they can be adapted for applications in the petroleum industry.

Some HRA methods provide guide of task analysis for identifying the possible human errors and the factors that could contribute the human errors. There are various taxonomies of task types, human error mechanisms, error-forcing contexts, and PSFs summarized and discussed in many HRA studies. These taxonomies are derived from empirical data from various industries or psychology theoretical insights. The task analysis discussions and the taxonomies from the available HRA methods are useful for prospective human error prediction, retrospective human error investigation and human error trend analysis. Based on the specific application needs in the petroleum industry, the analysts can choose to use an available taxonomy or customize a new taxonomy based on the available ones. In conjunction with human error identification, techniques of human error assessment and human error reduction are introduced.

This paper also discussed database design for collecting human error information, including the types of information to be captured in a human error database and how can a database supports human error management, in trends analysis, HEP assessment, lesson learning and sharing, and facilitating the error reduction process. The authors are implementing a human error database and improving existing failure reporting systems using HRA specifically for the petroleum industry and will share the benefit and improvement of service reliability as the project progresses.

REFERENCE

BP report, 2010. *Deepwater horizon accident investigation report,* s.l.: s.n.

Chang, Y. & Mosleh, A., 2007. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents: Part 5: Dynamic probabilistic simulation of the IDAC model.. *Reliability Engineering & System Safety,* 92(8), pp. 1076-1101.

Dhillon, B. S., 2007. *Human reliability and error in transportation systems.* Berlin: Springer.

EMBREY, D. E. et al., 1984. *SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgment,* s.l.: US Nuclear Regulatory Commission: Washington DC..

Forester, J., Kolaczkowski, A., Lois, E. & Kelly, D., 2006. *Evaluation of human reliability analysis methods against good practices,* s.l.: Sandia National Laboratories.

Gertman, D., 2002. *Review of findings for human error contribution to risk in operating events.,* DC: US Nuclear Regulatory Commission.

Gordon, R. P., 1998. The contribution of human factors to accidents in the offshore oil industry. *Reliability Engineering & System Safety,* 61(1), pp. 95-108.

Groth, K. M. & Mosleh, A., 2012. A data-informed PIF hierarchy for model-based human reliability analysis. *Reliability Engineering & System Safety,* Volume 108, pp. 154-174.

Hallbert, B. et al., 2006. *Human event repository and analysis (HERA) system, overview,* s.l.: US Nuclear Regulatory Commission.

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

Kirwan, B., 1994. *A Guide to Practical Human Reliability Assessment. CRC Press..* s.l.:CRC Press.

Kirwan, B., 1996. The validation of three Human Reliability Quantification techniques—THERP, HEART and JHEDI: Part 1—technique descriptions and validation issues. *Applied ergonomics,* 27(6), pp. 359-373.

PatC-Cornell, M. E., 1993. Learning from the Piper Alpha Accident: A Postmortem Analysis of Technical and Organizational Factors. *Risk Analysis*, 13(2).

Reason, J., 1990. The Contribution of Latent Human Failures to the Breakdown of Complex Systems. *Philosophical Transactions of the Royal Society of London,* 327(1241).

Spurgin, A. & Lydell, B., 2002. *Critique of current human reliability analysis methods.* s.l., Proceedings of the 2002 IEEE 7th Conference.

Swain, A. D. &. G. H. E., 1983. *Handbook of human-reliability analysis with emphasis on nuclear power plant applications,* s.l.: U.S. Nuclear Regulatory Commission.

U.S. NRC report NUREG/CR-1984, 1984. *General human reliability estimates using expert judgment,* s.l.: U.S. Nuclear Regulatory Commission.

U.S. NRC report NUREG/CR-6883, 2005. *The SPAR-H Human Reliability Analysis Method,* s.l.: U.S. Nuclear Regulatory Commission.

U.S. NRC report NUREG/IA-0216, 2012. *International HRA empirical study.* DC, U.S. Regulatory Commission.

U.S. NRC report NUREG-1624, 2000. *Technical basis and implementation guidelines for a technique for human event analysis,* DC: U.S. Nuclear Regulatory Commission.

Wiegmann, D. A. & Shappell, S. A., 2012. *A human error approach to aviation accident analysis: The human factors analysis and classification system.* s.l.:Ashgate Publishing, Ltd..

Williams, J., 1992. *Toward an improved evaluation tool for users of HEART.* Orlando, Proceedings of the International Conference on Hazard Identification, Risk Analysis, Human Factors and Human Reliability in Process Safety.

Woods, D. D., Roth, E. & Pople, H., 1987. *Cognitive environment simulation: an artificial intelligence system for human performance assessment,* s.l.: U.S. Nuclear Regulatory Commission.