9-1-2017

Retrospective Application of Human Reliability Analysis for Oil and Gas Incidents: A Case Study Using the Petro-HRA Method

Ronald L. Boring
Idaho National Laboratory

Martin Rasmussen
NTNU Social Research

Thomas A. Ulrich
University of Idaho

Harold S. Blackman
Boise State University

Boring, R.L.; Rasmussen, M.; Ulrich, T.A. & Blackman, H.S. "Retrospective Application of Human Reliability Analysis for Oil and Gas Incidents: A Case Study Using the Petro-HRA Method", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1), 1653-1657. Copyright © 2017 (SAGE). Reprinted by permission of SAGE Publications. doi: 10.1177/1541931213601900

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, published by SAGE. Copyright restrictions may apply. doi: 10.1177/1541931213601900

Retrospective Application of Human Reliability Analysis for Oil and Gas Incidents: A Case Study Using the Petro-HRA Method

Ronald L. Boring,[1] Martin Rasmussen,[2] Thomas A. Ulrich,[3] and Harold S. Blackman[4]

[1]Idaho National Laboratory, Idaho Falls, Idaho 83415, USA
[2]NTNU Social Research, 7491 Trondheim, Norway
[3]University of Idaho, Moscow, Idaho 83844, USA
[4]Boise State University, Boise, Idaho 83725, USA

Human reliability analysis (HRA) may be performed prospectively for a newly designed system or retrospectively for an as-built system, typically in response to a safety incident. The SPAR-H HRA method was originally developed for retrospective analysis in the U.S. nuclear industry. As HRA has found homes in new safety critical areas, HRA methods developed predominantly for nuclear power applications are being used in novel ways. The Petro-HRA method represents a significant adaptation of the SPAR-H method for petroleum applications. Current guidance on Petro-HRA considers only prospective applications of the method, such as for review of new systems to be installed at offshore installations. In this paper, we review retrospective applications of Petro-HRA and analyze the Macando Oil Well-Deepwater Horizon accident as a case study.

BACKGROUND

Human reliability analysis (HRA) was originally established in the U.S. to ensure minimal human errors during weapons assembly work, but the approach was quickly adapted for other safety critical applications, particularly control room operations in nuclear power plants. Since the advent of the first HRA method, the Technique for Human Error Rate Prediction (THERP; Swain and Guttmann, 1983), numerous HRA methods have been developed. These methods allow analysts to determine causes of human errors and quantify the human error probability. Yet, much of the focus of HRA has remained in nuclear power, where risk-informed safety regulation mandates the use of HRA in risk analyses and plant licensing.

HRA is nonetheless growing in popularity outside nuclear power. Safety critical domains such as aerospace, military, chemical process control, transportation, and oil and gas are seeing increased interest in HRA to support a broader application of risk and safety analysis in those fields.

HRA is increasingly applied in oil and gas. Boring (2015) has noted there are fundamental differences between the nuclear power and oil and gas domains. Key differences include: the types and configurations of control rooms and operations centers, the types of processes being controlled, the types of technologies being used, the types of hazards and consequences, the specification level of the written procedures, and the safety culture at the facilities. It is therefore reasonable to question many of the operational assumptions underlying different HRA methods and to develop new HRA methods or adapt existing HRA methods to better support the context of oil and gas. From a risk analysis perspective, nuclear power typically features considerably more comprehensive probabilistic risk assessment (PRA) models, which allow easy incorporation of HRA. Lacking such PRA models in many cases, the HRA used for petroleum applications must be much more standalone than its nuclear power counterparts.

One method that's been developed to date to address these differences is the Petro-HRA method, which will be described in the next section.

THE PETRO-HRA METHOD

Developing a New HRA Method for Oil and Gas

Although HRA had been applied in the oil and gas industry historically to a limited extent, a sudden increase occurred after the 2010 oil and gas blowout incident at the Macondo Oil Well, and there were subsequent mandates to look towards the nuclear industry and their incorporation of human aspects in risk analysis through HRA. In many countries, there are new requirements to model HRA, creating a regulatory framework similar to that found in the nuclear power industry.

In 2012, the Research Council of Norway and the Statoil crown corporation funded a new research project to refine or develop an HRA method specific for oil and gas applications for the Norwegian oil shelf. Although primarily involving Norwegian partners (i.e., Norwegian University of Science and Technology, DNV-GL, Institute for Energy Technology, and SINTEF, plus Idaho National Laboratory), the approach would be generalizable to other countries and their oil and gas applications in drilling, production, and transportation.

The four-year research project resulted in the publication of the *Petro-HRA Guideline* (Bye et al., 2017). The Petro-

HRA method was developed to meet the needs of the oil and gas industry through reviewing and adapting existing HRA methods, best practice documents, and research on human performance. Petro-HRA was created to be a complete HRA method including steps on: scenario definition, qualitative data collection, task analysis, human error identification, human error modeling, human error quantification, and human error reduction. This is contrary to most HRA methods where many of these steps are not described. For example, many simplified HRA methods from the nuclear power industry focus primarily on human error quantification. This focus is possible because many other steps are already specified as part of the PRA. Absent the PRA, it was necessary to make these implied steps explicit requirements in the method.

Origins in SPAR-H

After reviewing several HRA approaches, the Petro-HRA project team decided to use the Standardized Plant Analysis Risk-Human Reliability Analysis (SPAR-H; Gertman et al., 2005) method as the basis for Petro-HRA. SPAR-H is designed to be a simplified method for human error quantification. Because SPAR-H was originally built around SPAR PRA models, it came with clearly defined human failure events that minimized the need for extensive model building and qualitative analysis. However, when removed from the SPAR models, SPAR-H lacked clear guidance on building the HRA leading up to quantification. Thus, the Petro-HRA method can be seen as filling in gaps toward a complete SPAR-H HRA method. In addition to filling in gaps, the Petro-HRA method reconsidered the performance shaping factors (PSFs) used in SPAR-H. The PSFs account for the context that primes or decreases human error, and these proved to be different for nuclear power vs. oil and gas.

The original version of SPAR-H (Blackman and Byers, 1995) was developed to support the Accident Sequence Precursor (ASP) program at the U.S. Nuclear Regulatory Commission (NRC). ASP is invoked when there is a reportable incident at a U.S. nuclear power plant. The purpose of this retrospective analysis is to determine the likelihood that a similar incident could happen again or elsewhere. Although the probability of the incident happening is 1.0, because the incident actually happened, the probability of a similar event happening is usually less than 1.0. The goal of ASP is to determine if the risk of recurrence requires corrective action at the plant or similar plant. Is it a given that the same thing would happen again given a similar context? If so, it's necessary to implement corrective actions to reduce the risk of recurrence.

SPAR-H provided a mechanism to support improved human error quantification. One of the struggles that the U.S. NRC was facing in ASP was the ability to quantify a wide range of human errors that resulted from a variety of factors. Any new method needed flexibility. In addition, the validity and reliability of the quantification of human errors had been questioned such that the method needed to have a firm basis in theory as well as an approach that improved reliability. To address these issues the decision was made to base the model on an underlying model of human cognition, namely information processing.

The SPAR-H development team began with an information processing model and then identified factors that would affect each element of the model as well as factors that might influence the execution of a response elected as a result of the information processing. This was done to ensure a level of completeness and transparency in the method in terms of the relevant psychological elements. The next step was to do the same for the operational factors found in the environment of a nuclear power plant. These operational factors were identified and then attached to the appropriate elements of the information processing and response model previously identified. This resulted in a description of how humans make decisions and respond, and the factors that would ultimately impact their ability to successfully do so. In order to produce a workable method, these factors—both psychological and operational—were then examined to produce a set of summary factors for use in the method. These summary factors became the SPAR-H PSFs: available time, complexity, stress and stressors, experience and training, procedures, ergonomics and the human-machine interface, fitness for duty, and work processes (including crew dynamics). The background information generated through the model development and factor identification and definition then provided the information necessary to define the factors in a way meaningful to those involved in the nuclear industry.

The final step was then to devise an approach for quantification. The team believed that by knowing what factors influenced performance that essentially built the context of the situation, they could assess their influence and in turn use that knowledge to modify a nominal error rate. The team used THERP (Swain and Guttmann, 1983) as well as the open literature as a source to determine both the base rates and the multipliers that could be used to assess the impact of the level of the PSFs. By doing so, the team was able to create worksheets that could be used in a straightforward fashion to quantify the human error once the PSFs were assessed. After the original development, the method went through a series of modifications and is currently best described in NUREG/CR-6883 (Gertman et al., 2005).

Retrospective vs. Prospective Analysis

As noted, Petro-HRA is an expansion of the SPAR-H method. Yet, an important distinction between the two variants is that SPAR-H was developed for retrospective analysis, while Petro-HRA was developed for prospective analysis. *Retrospective* analysis refers to investigation of an incident that has already occurred, while *prospective* analysis refers to anticipated performance. Retrospective analysis is commonly associated with accident investigation, although most incidents investigated do not rise to the level of severity of an accident. Prospective analysis is commonly associated with quantifying the safety of new systems. It is not always possible to anticipate every factor that will influence operator performance within a system that hasn't actually been installed yet. As such, prospective analyses tend to be at a higher level with generic or nominal behaviors assumed and modeled. When an incident actually occurs, this may cue a more thorough analysis, in which all details are available.

Some PSFs in SPAR-H are inherently retrospective. In almost no cases, for example, would an analyst assume degraded fitness for duty by those using the system. Fitness for duty is a condition in which the user is not in a mental or physical state to operate the system. Factors ranging from psychological stress to medications or psychoactive drugs may negatively impact performance, but rarely would these be considered prospectively. Fitness for duty is a PSF that can be applied retrospectively in the unlikely event that staff are found to have violated safety protocols related to fitness.

In this paper, we explore the use of Petro-HRA for retrospective analysis. Retrospective analysis is of particular interest to the regulators of oil and gas installations, because it enables the regulators to determine root causes of an incident after the fact and use those root causes to trend vulnerabilities at similar installations and prescribe corrective actions to prevent recurrence. In the next section, we review the Macando Oil Well accident involving Deepwater Horizon. This well-documented event provides a useful example of how retrospective HRA can provide insights into the causes of events and the prevention of similar events in the future.

RETROSPECTIVE ANALYSIS CASE STUDY

Human Actions in the Macando Oil Well Accident

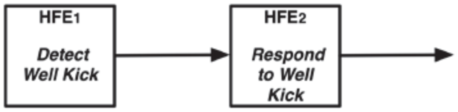
On April 20, 2010, an oil and gas blowout event at the Macondo Oil Well caused an explosion and fire that resulted in 11 fatalities, 17 seriously injured personnel, the sinking of the Deepwater Horizon (DWH) drilling rig, and the devastating release of millions of gallons of oil into the Gulf of Mexico (U.S. Chemical Safety Board, 2014). The accident can be attributed, in part, to a failure to detect the kick and subsequent blowout or uncontrolled release of oil and gas hydrocarbons from the well. The backpressure drove the hydrocarbons through the drilling apparatus to the rig, where they were ignited in an explosion that subsequently set fire to the rig. The rig had finished the exploratory drilling phase of operations and was in the process of performing temporary well-abandonment activities to prepare the well for the production phase of operations that another rig was scheduled to perform.

The well-abandonment activities entail plugging the well with cement, ensuring the integrity of the cement plugs via a negative pressure test, and then retracting the drilling apparatus. The negative pressure test circulates chemically treated mud that serves as the primary barrier to prevent the hydrocarbon from traveling through the well and into the drilling apparatus. The negative pressure created by circulating the mud simulates the low pressure seafloor atmosphere in order to verify the cement plug is properly sealing the well. Pressure and flow indications were available to the drilling team, but due to urgency to finish the drilling phase of operations they went unnoticed until the negative pressure test was performed. A supervising representative from BP overseeing the drilling operation did raise a concern to the driller; however, any concern was alleviated by more experienced drilling team members stating the odd pressure values were not uncommon and did not merit any significant concerns. Operations resumed, though the undetected kick had occurred up to an hour prior and was continuing to worsen over time until ultimately the blowout alarm sounded at 9:47 PM, followed swiftly by the explosion and fire. The order to abandon ship was issued at 10:00 PM.

Human Failure Events

The evolution of a well kick event follows in two phases. First, there is the initial phase prior to the personnel becoming aware of the well kick (which might be called normal operations). Even though the well kick has actually occurred, the personnel involved in the drilling or completion activities have not yet changed their activities to respond to the well kick. The second phase entails response to the well kick after the personnel detect the well kick. At this point, there is a sudden change in the activities of the drill operators and support personnel. We have characterized the event broadly as two Human Failure Events (HFEs) related to well kick (see Figure 1 for a simple graphical depiction).

Figure 1. Example well kick HFEs in sequence.



In reality, the recovery activities consist of many separate HFEs. However, the general context as represented by the PSFs for each of those post well kick activities largely remains the same. Additionally, if there is a failure to detect the well kick, there is obviously little opportunity for recovery actions nor the need to model a second HFE.

Petro-HRA Analysis

The nominal or default human error probability (HEP) in Petro-HRA is 0.01. Petro-HRA, like SPAR-H, uses the nominal HEP to represent basic tasks performed within the HFE. These nominal HEPs are then modified using multipliers corresponding to different levels of influence of the PSFs.

Petro-HRA makes use of nine PSFs: time; threat stress; task complexity; experience/training; procedures; human-machine interface; attitudes to safety, work and management support; teamwork; and physical working environment. This list omits fitness for duty and work processes from SPAR-H and adds attitudes to safety, teamwork, and physical working environment as new PSFs. Some PSFs—threat stress and task complexity—have been relabeled to remove some ambiguity in the original SPAR-H terminology.

Generally, the Petro-HRA PSFs can have three types of effects:

- *Negative:* A negative effect means that the PSF decreases human reliability. For example, to denote the negative effect of time would mean to suggest that there was inadequate time available to complete the task reliably.
- *Nominal:* A nominal effect means that the default applies. Nominal time, for example, suggests that there's adequate time to complete the task without undue time pressure or extra time.
- *Positive:* A positive effect means that the PSF increases human reliability. Positive time means that there is extra time over what is needed to accomplish the task.

For our two example HFEs, the following PSF effects could be noted. For detection of the well kick (HFE_1), the time available will vary considerably from situation to situation. Because there are indicators of an impending well kick such as the negative pressure test, there is generally a window of time to respond and prevent the event. However, if these indicators go unheeded, the available time erodes, and the ability of the drilling crew to respond decreases proportionately to the decreasing time window. It may be assumed that when a well kick is impending, the available time to detect will adversely affect the HEP. The clock is ticking, so to speak, which can only operate negatively on the outcome of the event. Of course, there is considerable task complexity involved, and in the case of DWH, there was degraded equipment (human-machine interface) resulting in poor indicators due to the backlog of maintenance activities. All other PSFs are assumed to be nominal.

The detection of a well kick triggers a change: response actions are needed in order to prevent a blowout (HFE_2). This operational shift will generally result in multiple elevated negative PSFs relative to nominal or normal operations. The time window is closing, but there may also be elevated negative threat stress and task complexity, potentially diminished levels of experience and training for this type of situation, and potentially poor to incomplete procedures. Underlying the situation, negative teamwork and management support factors such as breakdowns in communication, coordination, or command and control may also manifest. Finally, even in the presence of clear well kick indications, there was considerable hesitancy to perform an emergency disconnect due to the extreme cost associated with that action and job penalties for a false alarm. These factors suggest the PSF related to attitudes toward safety, work, and management support was at play.

While detection of the well kick (HFE_1) can be seen as a mostly nominal influence of the PSFs, the transition to emergency operations to prevent blowout (HFE_2) will likely invoke multiple negative PSFs.

The Basic HEP is defined in Petro-HRA as the nominal HEP multiplied by the product of all PSF multipliers:

$$\text{Basic HEP} = \text{Nominal HEP} \times \prod \text{PSF Multipliers}$$

For HFE_1 related to well kick detection, the PSF product is calculated to consider negative effects of time (moderately negative has a multiplier of 10), task complexity (moderately negative has a multiplier of 10), and human-machine interface (moderately negative has a multiplier of 10). All other PSFs are considered nominal, with a multiplier equal to 1:

$$HEP_{HFE1} = 0.01 \times 10 \times 1 \times 10 \times 1 \times 1 \times 10 \times 50 \times 1 \times 1 \approx 1.0$$

According to the laws of probability, this number is, of course, truncated at HEP = 1.0, suggesting that given the circumstances, the well kick detection would almost certainly have been doomed to failure.

The same process applies to HFE_2 related to the response to the well kick. Again, multiple negative PSFs are in effect, including: time (very highly negative has a multiplier of 50), threat stress (very highly negative has a multiplier of 25), task complexity (very highly negative has a multiplier of 50), experience and training (very highly negative has a multiplier of 50), procedures (very highly negative has a multiplier of 50), and management support (very highly negative has a multiplier of 50). Even assuming the remaining PSFs are nominal, the HEP quickly escalates to certain failure:

$$HEP_{HFE2} = 0.01 \times 50 \times 25 \times 50 \times 50 \times 50 \times 1 \times 50 \times 1 \times 1 \approx 1.0$$

A prospective analysis, in contrast, might have assumed nominal PSFs for HFE_1 and primarily weighted time as a factor in responding to the detection for HFE_2. Both would have resulted in HEPs below 1.0, likely HEP = 0.01 for HFE_1 and HEP = 0.1 (considering a moderately negative time PSF multiplier equal to 10) for HFE_2. In other words, a prospective HRA would likely have underestimated the number of simultaneously negative factors contributing to the event. Indeed, the situation at DWH may be considered a worst-case scenario, and it would never be assumed that the confluence of so many negative events would actually occur. Except, they did.

Discussion

What can be learned from the retrospective analysis of the Macando Well accident? The analysis would only seem to confirm that many negative things happened, resulting in the accident. More telling, however, is that the negative PSFs implicated for HFE1 involving detecting the well kick occurred over a prolonged state. The negative PSFs of time, complexity (given the degraded facility conditions), and missing indicators due to a faulty human-machine interface resulted in an ongoing degraded condition at DWH. While it is easy to say in hindsight that these conditions primed failure, the fact that the risks associated with these degraded states and lack of risk-informed decision making regarding decisions to finalize drilling and move quickly to production fatally undermined the safety of DWH. The retrospective analysis shows not only that DWH was doomed to failure but that it had been operating that way for some time. Of course, this assessment is not an indictment on any of the individuals working on DWH. There is no individual blame in HRA. Multiple factors put DWH in jeopardy, and the retrospective analysis reveals factors to watch to avoid similar incidents at other installations.

CONCLUSIONS

While the primary goal of a retrospective analysis is to identify leading indicators to the failure, another goal is to generalize the findings to other installations. Within nuclear power, there are a number of commonalities between nuclear power plants, which makes it relatively easy to generalize findings to the fleet of related plants. Drilling rigs may represent more unique configurations that are increasingly tailored to first-of-a-kind applications. Easy-to-reach oil has now been exhausted, meaning it requires increasingly unique solutions to tap new oil sources. As such, the generalizability of a retrospective analysis from one installation to another may be limited. The value in retrospective analysis may reside mostly as a root cause tool to ensure similar human performance deficits aren't mirrored across different installations.

Several of the PSF multipliers in Petro-HRA are higher than in SPAR-H. It might therefore be concluded that Petro-HRA is more conservative than SPAR-H. This conservatism should be considered when performing a retrospective analysis. It may be desirable to benchmark the Petro-HRA method against other HRA methods like SPAR-H that have been used retrospectively.

In general HRA, methods have a nominal HEP which is increased as contextual factors degrade the situation. While some methods also include the potential to decrease the HEP through positive factors, the emphasis is certainly on negative factors. This can lead to a situation where the more you know about a scenario the higher the HEP becomes (as fewer and fewer factors are assumed nominal). As a retrospective analysis often is done on an incident/accident scenario that has occurred it is likely to produce a rather high HEP. But the high HEP should not be driven by the fact that the quality and amount of information from an event investigation used in a retrospective analysis will likely be much different than that coming from a prospective analysis. The Petro-HRA method should be benchmarked against other HRA methods that have been used in retrospective analyses. Additional guidance should be developed for Petro-HRA to determine the best way to apply the Petro-HRA steps retrospectively. New guidance will ensure that information from event investigations is consistent to inform the HRA.

Retrospective analysis using Petro-HRA promises to be an important tool to the regulator of oil and gas applications. It affords a method to extract the primary human contributions to an event and to mitigate such contributions in the future. This paper has presented a case study of such an application and points to the potential for wider use in the future.

DISCLAIMER

This work of authorship was prepared in part as an account of work sponsored by Idaho National Laboratory, an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights.

REFERENCES

Blackman, H.S., Byers, J.C. (1995). *ASP Human Reliability Methodology Development, INEL-95/0139*. Idaho Falls: Idaho National Engineering Laboratory.

Boring, R.L. (2015). Adapting human reliability analysis from nuclear power to oil and gas applications. *Proceedings of 2015 European Safety and Reliability (ESREL) Conference*, pp. 2853-2860.

Bye, A., Laumann, K., Taylor, C., Rasmussen, M., Øie, S., van de Merwe, K., Øien, K., Boring, R., Paltrinieri, N., Wærø, I., Massaiu, S., & Gould, K. (2017). *The Petro-HRA Guideline*. Halden, Norway: Institute for Energy Technology.

Gertman, D., Blackman, H., Marble, J., Byers, J., & Smith, C. (2005). *The SPAR-H Human Reliability Analysis Method, NUREG/CR-6883*. Washington, DC: U.S. Nuclear Regulatory Commission.

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

U.S. Chemical Safety Board. (2014). *Investigation Report Overview: Explosion and fire at the Macondo Well, 2010-10-I-OS.*