Improving Human Factors Review in PHA and LOPA

aeSolutions Technical Team
10375 Richmond Ave. Suite 800
Houston, TX 77042

Prepared for Presentation at
American Institute of Chemical Engineers
2017 Spring Meeting and 13th Global Congress on Process Safety
San Antonio, TX
March 26 – 29, 2017

Improving Human Factors Review in PHA and LOPA

aeSolutions Technical Team
10375 Richmond Ave. Suite 800
Houston, TX 77042

Keywords: Human Factors, Human Error, PHA, LOPA, Facilitator, Task Analysis, Bias, Cognitive Psychology

Abstract

Human Reliability practitioners utilize a variety of tools in their work that could improve the facilitation of PHA-LOPA related to identifying and evaluating scenarios with a significant human factors component. These tools are derived from human factors engineering and cognitive psychology and include, (1) task analysis, (2) procedures and checklists, (3) human error rates, (4) systematic bias, and (5) Barrier effectiveness using Bow-tie.

Human error is not random, although the absent minded slips we all experience seem to come out of nowhere. Instead, human error is often predictable based on situations created external or internal to the mind. Human error is part of the human condition (part of being a human) and as such cannot be eliminated completely.

For example, a task performed at high frequency (e.g., daily or weekly) develops a highly-skilled operator with an expectation of a low error probability for that task. However, as the operator' skill increases, their reliance on procedures decreases, leaving them open to memory lapses caused by internal or external distractions. The fact that a skilled operator becomes less dependent on procedures is not a conscious decision. It is part of the human condition. Forcing a skilled operator to read the procedure while performing the task they are skilled at, is like asking you to think about what your feet are doing as you walk down a flight of stairs. In both cases a loss of adroitness will occur.

A large portion of this paper will be to describe with practical examples the five tools mentioned above. Task analysis is a talk-through and walk-through exercise of a task (typically focusing on one or two critical steps of a procedure) that is used to identify error likely situations (ELS). Quantitative human error rates can be attached to the ELS depending on if the error is associated with skill, rule, or knowledge (SRK) based performance. Systematic biases produced by Type 1 (fast) thinking cause judgment and diagnosis errors related to response to abnormal situations.

Having a working knowledge of these five tools will improve a PHA-LOPA facilitator's awareness and ability to better evaluate human error related scenarios and Barrier failure. In addition the facilitator will feel confident about recommending the need for a more detailed follow-up study such as an HRA (Human Reliability Analysis).

1 Introduction

This paper is written for facilitators of PHA and LOPA and their managers. The message is simple: Improve human factors review in PHA and LOPA by applying the already established tools of human factors and human reliability. How can we improve something we do not practice ourselves as facilitators? There is a caveat however. Much of the heavy lifting has to be done outside of PHA and LOPA studies. The argument for this is presented below in **Section 2.2**.

The LOPA method itself presents challenges and opportunities for using human factors and human reliability. Any LOPA scenario handling human error as initiating cause, human response, or administrative control barriers is a candidate for a safety critical task analysis. There is ample low-hanging fruit with regard to human factor based scenarios already coming out of PHA and LOPA. One of the tenets of this paper is that any credit given for human response to a critical alarm or other human barrier, or any human error initiating a catastrophic scenario, should undergo a safety critical task analysis, regardless if the gap is closed using generic credits (see **Section 2.4.5** for why).

The following tasks have been successfully treated with the qualitative and quantitative methods and concepts discussed in this paper. It is important to note that each of these tasks were well-defined, bounded, primarily skill and procedurally based tasks involving one or more personnel.

- Specific task performed during normal shutdown
- Specific task performed during start-up
- Safing for maintenance work
- Safing for N2 back-up in Instrument Air
- Response to critical alarm (including diagnosis)
- Restoration of critical equipment following ITPM (inspection, test, preventive maintenance)
- Catalyst drum change-out
- Rupture disk Barrier effectiveness
- Tanker Truck Unloading

Organizational factors are recognized as important human factors. Organizational factors such as social and institutional influences are discussed in this paper qualitatively. The quantitative treatment of organizational factors as they pertain to predicting incidents is complex (some would argue currently intractable) with large uncertainty. This paper is primarily written addressing skill or rule based tasks performed by one or more operators.

2 Why Human Factors and Human Reliability?

"Most industrial accidents are caused by human error: estimates range between 75 and 95 percent. How is it that so many people are so incompetent? Answer: They aren't. It's a design problem. If the number of accidents blamed upon human error were 1 to 5 percent, I might believe that people were at fault." Don Norman [1]

"Saying most accidents are due to human error is a bit like saying falls are due to gravity." Trevor Kletz [2]

2.1 We are Under-predicting Process Safety Incidents

There is published evidence [3] stating that current methods (HAZOP, LOPA, QRA) are predicting only about one-third of actual process safety incidents. The under-prediction is coming from the frequency (or likelihood) part of the risk equation. Identifying hazards and estimating consequences has been comparatively more robust.

Could this observed gap be coming from Human Factors impact?

2.2 The Case for Why PHA and LOPA are Not Enough

PHA and LOPA, and their derivatives including Procedural HAZOP, generic Human Factors checklists (see **Section 9** for a comparison with a Task Analysis checklist), etc. are targeted to evaluating states and conditions (e.g., is the alarm configured as priority, or, what-if this step in the procedure is omitted?) as opposed to behavior (e.g., how will the console operator interpret the alarm, or, how likely is the operator to skip that step and why?). States and conditions are important to validate, however behavior is just as important. The existing tools available to evaluate human (and organizational) behavior are task analysis and human reliability analysis. PHA and LOPA methodology cannot get you there.

Another reason is related to Barrier effectiveness. Barriers are listed as providing protection in a LOPA study. The hardware reliability of a Barrier can be and is typically estimated (or assumed to be a generic value). So to can the human impact on a Barrier be estimated, except it's typically not. Sometimes it's assumed the generic data value for hardware failure includes systematic (human) failures. See **Section 2.4.5** for why this is a myth. Human and organizational factors are at least as important as random hardware failures on a barrier, and they are not being accounted for. Methods outside of PHA and LOPA are needed.

2.3 Human Factors

2.3.1 <u>What is it</u>

A human factor is anything that affects human performance, good or bad. The goal of human factors is to match the task environment and system interface(s) to the abilities of the operator. If the physical or mental requirements of a task exceed the operator's ability, errors will occur.

2.3.2 <u>Affordances, Signifiers, Forcing Functions, and Mapping</u>

When performing a task analysis the facilitator must be able to spot deficiencies in the task environment. This can only come with practice. The following tools are used by human factors specialists [1]:

Affordances are used to afford a location to perform work. A simple example is a valve handle. It provides a place for the operator to place their hands. Affordances can be used improperly, for example, an operator may have to stand on said valve handle (or the piping) to reach a manual valve above him.

Signifiers tell us how to interact with the affordance. A round valve handle turns. A levered valve handle rotates. The best signifiers are intuitive (don't require conscious thought, i.e., Type 2 thinking – see **6.2**). Signs require effort and conscious thought (to see, read, and comprehend), and are therefore not the best use for signifiers (this is not to suggest getting rid of signs, but to understand the limitation). Populational stereotypes are learned signifiers, for example, "righty-tighty, lefty-loosey." Designs that violate populational stereotypes are considered error-likely.

Forcing functions force desired behavior. For example, car-seals, chain locks, and captive key systems placed on valves are recognized as reducing errors of omission and commission related to valve manipulation.

And, **mapping** is used to spatially connect controls with affordances. For example, procedures that use labels and drawings to map said valve to the actual valve in the field.

These concepts are only useful in the context of the operator performing the task, and so task observation of the operator with the equipment and procedures is necessary to be complete a successful analysis.

2.3.3 <u>Human Nature</u>

Slips and lapses (defined in **Section 3.4**) that sometimes occur are part of human nature. This type of human error cannot be trained, punished, rewarded, designed, or otherwise removed out of humans. It is part of being human. Policy, procedures, and good human factors help, but we should still expect the occasional human error. We should design systems to be tolerant of human error or allow for their recovery. If an operator commits an error, we should not look at the operator, but instead look at the task, procedures, environment, equipment interfaces, etc. the operator has to work with, because it is more likely those are the root cause (not the human).

2.3.4 <u>Procedures</u>

Different types of procedures are in use, including,

- Routine Operating
- Preventive Maintenance (e.g., for critical instrumentation)
- Response to an abnormal situation (critical alarm based response)

Each type of procedure is amenable to human factors and human reliability analysis via task analysis.

2.4 Human Reliability

Human reliability study is the quantification of human factors. Human reliability puts numbers to (i.e., attempts to quantify) human factors influences on tasks and can be used to help "sharpen the pencil" for error estimates to use in PHA-LOPA or QRA. Human reliability has a long history (see for example reference [4]) and is part of the renaissance of quantitative methods in the process industries [5]. Any number that is produced has an uncertainty associated with it (see **2.4.2** and **Section 8**).

2.4.1 Tractability

How can we predict human error? Humans are unpredictable (i.e., often irrational).

There are several ways to answer this question. The easiest and most relevant is the fact that there is so much low-hanging fruit associated with improving human (and organizational) factors related to hazard scenarios and Barrier effectiveness that it would take a generation worth of work in this field applied to the Process Industries to overtake the capabilities of the current methods.

The current methods to evaluate human behavior that include the SRK (skill, rule, knowledge) model and many human reliability models (e.g., THERP (Technique for Human Error Rate Prediction) – see ref [12]) have been validated as effective at predicting human error. The issue of human irrationality in the context of plant accidents even has a name: local rationality (what is in the head of the operator at the time they made that catastrophic error) [9].

The catastrophic accidents that post to the CSB (Chemical Safety Board) website at a steady rate, do not involve unpredictable human error.

One of the most incredible human factor case studies is the Browns Ferry fire of 1975. Browns Ferry nuclear plant is located in Alabama. The maintenance team had just installed some new cabling into a control room, and were leak testing the cable seals by looking for air flow in the duct, using the naked flame of a candle. The foam insulation caught fire, and burned through safety and control cables (common duct), and the operations team narrowly avoided a complete meltdown. One might ask how something like this could have been predicted a priori. It would be nearly impossible. But this and other extraneous acts often lie in wait for a number of years before causing an incident. The candle flame method to leak test had been in use for several years and was even recognized as a hazard.

2.4.2 Model Error

As mentioned above, the process industry has seen a "renaissance" in use of quantitative methods including SIL (safety integrity level) calculations, Fault Tree Analysis for likelihood of hazard, Event Trees for distribution of outcomes, and Bow-Ties (combined fault tree and event tree) as a means of communicating hazards and identified barriers to Operations staff, and identifying process safety leading indicators.

In any quantitative model of system performance, the following are typically recognized as sources of error:

- Equipment data (a function of equipment type, service, quality of maintenance, etc.)
- Incorrect model (incorrect logic)
- Human reliability impact
- Wrong assumptions (about plugging for example)
- Unforeseen hazards or ways in which hazards can occur (see **Section 8.1**).

The field of human reliability attempts to account for the uncertainty introduced by the impact of human factors on system performance. This may represent the most significant source of model error, next to bad assumptions.

2.4.3 HRA Methodologies

There are a myriad of HRA methodologies, some public, some proprietary, as evidenced by the veritable alphabet soup of acronyms associated with the methodologies.

The Health and Safety Executive (HSE) in the UK considered that it would be useful to be up to date with developments in the field of quantitative HRA methods and to have knowledge of the capability of the tools and an understanding of their strengths and weaknesses, to improve consistency, and determine acceptability of their use [6].
To sort this out, in 2009 the HSE published a review of 72 HRA methods that were "potentially relevant to HSE major hazard directorates." Of the 72 potential HRA tools, 17 were considered useful for major hazard directorates [6].

THERP (Technique for Human Error Rate Prediction) [12] is one of the 17 methods considered to be useful by the HSE.

2.4.4 THERP

THERP is considered a first generation human reliability model (in the public domain the models only go up to second generation). ATHEANA (A Technique for Human Error Analysis) is the second generation model developed for the USNRC (US Nuclear Regulatory Commission) that is intended to improve THERP, especially related to cognitive behavioral mechanisms, organization, and safety culture, in the context of response to abnormal events. But ATHEANA would benefit from further development and ultimately more widespread industry acceptance [6].

THERP is a perfectly acceptable methodology, especially for the low-hanging fruit coming out of current LOPA's. Some of the benefits of THERP:

- It's easy for someone with a PHA-LOPA-QRA-facilitator background to use
- It includes its own extensive human error data set
- It incorporates Task Analysis to uncover error mechanisms
- It has a very nice dependency model built-in
- It uses "recovery factors" such that unrecovered human error is the metric of interest
- It flanges up nicely with LOPA, QRA, and Bow-tie

THERP does take practice to get good at, and it does require an eye for human factors, and some study of cognitive behavior.

2.4.5 The Myth that Generic Equipment Failure Data includes Human Error

There is a belief that generic failure data (e.g., generic data tables for initiating events and protection layers) already includes data on human error. (It can be argued that *all* failure is ultimately human induced, but this is not the point I'm making.) This myth, however, is partly true. For example, failure associated with the quality of maintenance certainly contributes to the failure data of generic equipment. However, the following types of human errors are not included in generic equipment failure data [4]:

- Insertion of an incorrect component or device

- Wrong specification
- Miscalibration
- Left in wrong mode (manual, test, inhibit, etc.)
- Failure to restore system following maintenance (block valves)
- Inadvertent bypass during normal operation

But even if the above myth were true, it misses the point entirely. Hardware failure is considered random (statistical). Human error is systematic (i.e., deterministic). The two types of error differ only by the amount of information known about each. Behavioral models and task analysis provide information of why human error occurs or might occur. We can use this information to improve work processes, work situations, policy, competency, staffing, communication, contractor roles, etc. so that we might prevent the error from occurring. If we bury human error as a generic random variable, we miss this opportunity. Claiming that a generic value of human error is 1/10 and that it is conservative, is misleading and dangerous, because of the nature of determinism, the error likelihood (for an identified cause) is either 1 or 0. The same point could be made when discussing the numbers in **Section 5.2**. The way out of this problem is to *always* couple the quantitative analysis with the qualitative Task Analysis.

3 Tool 1: Task Analysis

3.1 Introduction

Human factors and human reliability share a technique called 'safety critical task analysis' or 'task analysis' for short. As the name implies, task analysis is a systematic analysis of a task (or barrier) that can help identify problems and reduce the likelihood of human error associated with that task (or barrier). Task analysis looks at all the ways that humans touch a system, how failure can occur, and ways to improve it. Task analysis is used to ferret out potential human errors related to a specific task, and make recommendations for improvement as well as estimate likelihood of occurrence.

Task analysis could look at the critical steps of a procedure (usually one or two steps of a large procedure), responding to a priority alarm, or restoring a critical system following maintenance. Task analysis can also be conducted on the escalation factors and escalation factor controls related to a Barrier. Task analysis is conducted as a team exercise, and includes talk-through and walk-through activities. Typically one task (or barrier) takes one day of team time to perform the task analysis, with additional follow-up reporting and some iteration.

You don't have to go far to find a task to analyze. They are right there in your PHA-LOPA study. Any catastrophic hazard scenario that has a significant human factors component including human barriers is a candidate for task analysis.

Task analysis can be picked up quickly by an experienced PHA-LOPA facilitator. The caveat is the facilitator needs to have an eye for human factors issues and an understanding of cognitive bias (see **Section 6.1**). It is recommended that a facilitator experienced in the methodology be engaged as a mentor.

The Functional Safety Assessment required by ANSI/ISA 84 as an audit of the Safety Instrumented System is really just a task analysis.

See **Section 9** for sample Task Analysis checklists.

3.2 Error-Likely Situations

A primary goal of task analysis is to identify error likely situations. This is partly accomplished by talking to operators that perform the task, reviewing procedures, and actually observing (real or a simulated walk-through) the task being performed. Here is where an eye for human factors is necessary. Much of human factors is common sense. If you talk to an operator who has been on the job for 30 years and he tells you this is the first time someone has ever come out to look at how he works, the fruit will be lying on the ground.

A lot of error likely situations are obvious, for example,

- Reading errors caused by missing valve handles or indicators (valve position), or instrumentation that is not back-lit.
- Omission errors caused by equipment that is not accessible.
- Commission errors caused by equipment layout.
- Non-response errors because alarms are buried in a DCS graphic, or the audible alarm volume is turned down too low.
- Interpretation error (of a signal) caused by failure to trend that signal on a DCS graphic (or use of process switches versus transmitters).
- Interpretation error of the state of equipment because there is not enough information for the Operator.
- Not believing an alarm is real because of so many other spurious alarms.
- Not knowing how to respond to an alarm because there is no training or documentation.
- Reliance on short-term memory because of deficient design.
- Spatial disorientation caused by layout (e.g., stairs) of identical equipment trains.

There are other error likely situations related to deficient procedures, work practices, communications, staffing, task load, competency, etc.

To be successful at identifying error likely situations takes practice. But a basic understanding of the types of human error possible is the way to begin.

3.3 Human Error is not a 'Cause'

Human error does not occur randomly (in a statistical sense). Human error is deterministic (i.e., it can be predicted using an appropriate behavioral model and task analysis). Examples of when human error will occur:

- When a person is rushing, distracted, fatigued, or complacent. (These are the Big 4.)
- Spatial disorientation (e.g., working on the wrong but identical equipment)
- Tasks that rely on short term memory (short term memory can hold approximately 5 items at most under ideal conditions)

- Interpreting a signal as false
- Knowledge based tasks (a novel situation) has higher error rates

We can reduce (but not eliminate) the likelihood of human error by evaluating the systems under which it occurs.

3.4 Human Error Classification

The following classification of error types has been established by engineering psychologists (a.k.a. human factors specialists) and has been used for many years [13].

3.4.1 Slips and Lapses

A ‘slip’ is an error of commission (a task done incorrectly).

A (memory) ‘lapse’ is an error of omission (forgetting a task).

Slips and memory lapses primarily occur during Type 1 thinking (see **Section 6.2**).

3.4.2 Mistakes

Mistakes are cognitive, that is they occur primarily during Type 2 thinking (see **Section 6.2**).

Examples of mistakes include:

- PHA errors
- Design errors
- Errors made writing procedures (not typos)
- Mis-diagnosing an abnormal situation

Mistakes primarily occur when faced with a new, ad hoc, or novel situation.

3.4.3 Craft Errors

Lack of adroitness related to tasks performed by maintenance, operations, pipe-fitters, etc. For a skilled craftsmen, these error likelihoods are lower than slips, lapses, and mistakes.

3.4.4 Social and Institutional Influences

These types of errors are the most insidious and can produce the most catastrophic incidents. Coupled with extended time and economic pressures, some hazard scenarios are especially susceptible to extreme outcomes. The added time a scenario may play out (e.g., days) allows multiple people (and therefore personalities) and conflicting objectives (safety v. economic) to creep-in. De facto Company policy, culture, remuneration, and the lines of communication are the organizational factors that are important.

Social and institutional factors can influence specific skill, rule, and knowledge based tasks. Task analysis has the potential to uncover some of these issues (e.g., by asking about non-compliance to a procedure), and can make these issues more discoverable.

3.4.5 Non-Compliance

A violation of a procedure or certain steps of a procedure. Not working to the procedure. Often times, the procedure is wrong, or the operator has found a better way of working. If there is a work process and culture to support operators writing their own procedures with appropriate checks of course, these types of non-compliance will not be an issue. On the other hand, ad hoc or unauthorized bypassing can be a serious non-compliance issue.

3.4.6 Summary of Error Classification

Table 1. Human Error Classification

Type of Error	Example	Most Likely Cause	Recovery Factors
Slip	Opening the wrong valve	Distraction	Feedback (usually immediate)
Memory Lapse	Forgetting a step in a procedure	Distraction	Checklist (see **4.1.5**)
Mistakes	Ignoring a real signal	Bias	Mistakes are hard to detect and correct
Craft	Loose Cam-lock fitting	Maladroitness	Feedback (usually immediate)
Social and Institutional	We've done it this way for many years and never had an accident.	Turkey Trap bias (see **8.1**)	Culture change
Non-compliance	Not following procedure	Poor procedures	Administrative
Knowledge based	Abnormal situation	Novel situation	Planning, Trial and error

4 Tool 2: Using Procedures and Checklists

4.1 How Routine Operating Procedures are Used

"*We have an opportunity before us, not just in medicine but in virtually any endeavor. Even the most expert among us can gain from searching out the patters of mistakes and failures and putting a few checks in place. But will we do it? Are we ready to grab onto the idea? It is far from clear*." Atul Gawande [7]

Procedures are needed for safe operations and reliable maintenance. Working without reference to a procedure creates an error likely situation. Procedures are an important human factors tool.

4.1.1 Referencing Procedures In PHA-LOPA

In a PHA-LOPA, references to specific procedures where human action is involved in the study should be made. For example, operator error as an initiating cause, or operator response to a critical alarm. A laptop should be open in the study and should be able to pull up and reference the specific procedures as needed. At times, a quick read through of the procedure with the team is warranted. If a human action shows up in a PHA-LOPA without a procedural reference and none exists, consideration should be given for the need of one.

4.1.2 Using Procedures to Operate the Plant

Step-by-step routine procedures used by operations to run the plant are a critical administrative control system. Written procedures establish minimum requirements and expectations, which should be communicated via initial operator qualification and periodic refresher training thereafter.

Procedures used in plant practice have been successful at keeping human task error rates lower compared to other fields that do not use any procedures, for example, the error rates of the medical profession as described in reference [7].

In practice, procedures are used or referenced (or not used) in a variety of ways.

- Procedures may not be accessible, or don't exist
- Procedures may be printed and read before a job
- Procedures may be carried in the truck or in the back pocket
- Procedures may be carried by a scribe as part of a checklist
- Procedures may be posted at the location the task is performed
- Procedures may be placed at the location where data readings are taken to record
- Procedures are committed to long term memory
- Operators don't have enough hands to carry and read a procedure while performing the actual task at the same time (walking, bending, carrying tools, turning valves, wrenches, pushing buttons, etc.)

4.1.3 Working from Memory versus Procedure

Experienced operators that perform regular tasks are working primarily from long term memory. New operators rely heavily on written procedures. New operators will often read the procedure just prior to the task (for which they are qualified), while experienced operators will refresh their long term memories on a task performed infrequently. Experienced operators may carry the procedure in their back pocket for reference. New operators will shadow senior operators and be mentored in a task (to be qualified). In all cases procedures should be made available. Procedures should be reviewed annually and refresher training provided per regulatory requirements.

An experienced operator will not perform a task by reading a step in a procedure, performing that step, then reading the next step and performing that step, and so on. Experienced operators work from long term memory as mentioned above. Besides, procedures are insufficiently detailed to be used in that manner, nor should they be made to be. An operator will quickly lose interest in a procedure that spells out every mundane step. All procedures could be improved. Some procedures contain mistakes or critical omissions. If operators actually followed these procedures to the letter, errors would be made every time they were used.

As a skill is perfected over time, Human nature directs us to use Type 1 thinking (see **Section 6.2**), and the tasks we perform are largely done on auto-pilot (i.e., without conscious thought). Potential for slips and memory lapses creep in. This is often why newly qualified operators can often be more reliable than experienced ones (new operators may make more mistakes, but experienced operators have more potential to make the critical errors). New operators tend to spend more time in Type 2 thinking during a task.

4.1.4 Checklist v. Procedure

A checklist can be an effective forcing function (something that forces human behavior in a desired way). The caveat is, a checklist will not be effective if it is not used properly. A checklist will not be used properly if for example it is poorly crafted (e.g., too long) or not used in the correct situation (e.g., where long term memory is good enough). In these cases, checklists become equivalent to procedures, with no added benefit. Many checklists are "pencil-whipped" because it is difficult to check boxes at the same time as holding tools and turning valves. In these cases where a checklist is deemed absolutely necessary, a scribe should be used to initial the steps as they are performed by a different person. The next section discusses improving the use of checklists.

4.1.5 Improving use of a Checklist

Checklists are less useful to experienced operators because much of what an experienced operator can do comes from long term memory, and therefore long checklists are redundant to procedures, tedious, and a waste of good time. Let an experienced operator use their training, experience, intuition, judgment, competence, etc. to perform a task, but supplement this with short (60 to 90 second) "pause points" at which juncture a short check-list is reviewed for the most critical steps.

5 Tool 3: Human Error Rates

Quantification of human error rates provides several benefits beyond having a number to use in LOPA or QRA.

- It forces decomposition of the task (which is more accurate than guessing the whole) see **Section 5.2**
- It forces a deeper dive to understand qualitative factors

There are caveats to using numbers. At times too much effort is spent on the numbers. The uncertainty is too large to justify it. Without qualitative understanding, the numbers are meaningless.

5.1 Failure Modes Distribution

The following is an approximate distribution of error likelihoods for different error types (the more likely being listed first).

- Knowledge based errors. Those that require planning and trial and error.
- Selection errors or slips coupled in time and space (e.g., on the flight deck, selection errors on a DCS graphic, or the myriad of small electronic gadgets that pack multiple modes and functions into a small space). In our personal lives these types of errors may occur *daily*.
- Memory Lapses most typically caused by external interruptions (being pulled to another task) or internal interruptions (what we call 'absent-mindedness', our mind is occupied with something other than the task at hand). The Big 4 that can lead to absent-mindedness are: rushing, distraction, fatigue, or complacency.

- Selection errors decoupled in time and space (e.g., two identical valves 10 feet apart operated 60 seconds or more apart).
- Skill-based errors or craft error based on physical skill.

5.2 Using Human Reliability Data

"*If we break problems down into component questions (What is the chance of A happening? What is the chance of B happening? What is the chance that a man will be in the area when both happen? What is the chance of a fatality if he is there?) answering them with facts where possible and with opinions only when there are no facts available we are more likely to get a correct answer than if we try and guess the answer to the whole problem*." Trevor Kletz [8]

5.2.1 Quantifying Human Error

The formula for calculating human error probability (HEP) is simple:

$$HEP = \frac{\text{Number of errors occurred}}{\text{Number of opportunities for error}}$$

The application of the equation is difficult for several reasons, and is why there exists a sparsity of human performance data.

Firstly, the numerator is hard to get because self-reporting of errors or reporting someone else's error has obvious problematic connotations.

Second, the denominator is hard to get because often we just don't know an accurate estimate of the number of opportunities for a given error.

5.2.2 Being better than 1/10

We are conditioned by LOPA that 1/10 is a typical and acceptable human performance value to use as a generic credit. That's fine (assuming we ignore the deterministic nature of human error, see **Section 2.4.5**). But human factors psychologists have recommended for us to consider error rates in our own personal lives that are as good as 1/1,000 and even under certain circumstances as good as 1/10,000. If you do this, you will convince yourself very easily that better than 1/10 is entirely possible and reasonable.

How good can humans be? Recognized and published human performance limiting values can be found in the literature, for example in Kirwan [4] as follows,

- Single operator carrying out task(s) on plant .. 1e-4
- Operators carrying out task on plant.. 1e-4 to 1e-5
- Control room based team.. 1e-5

These are HEP (human error probabilities). These are performance limits. Often human error rates for a specific task are found to be much higher than this.

To put this in perspective, one would need data for three years performing a daily task error free to demonstrate a HEP of 1e-3. One would need 30 years of data performing a daily task error free to demonstrate a HEP of 1e-4. For this reason, as a human reliability practitioner, one would be hesitant claiming as low as 1e-4 for a task performed by the same person even with good (self) recovery factors and good human factors.

For typical slips and memory lapses under Type 1 thinking (see **Section 6.2**) human error probabilities between 1e-3 and 1e-4 (for example, per step of a procedure) are perfectly reasonable, depending of course on the human factors.

5.2.3 The Relationship between Task Frequency and Human Error Rate

Operator tasks that are performed frequently (e.g., every shift) are especially problematic to handle quantitatively because the opportunity for error is so great, yet slips and lapses will occur at the same rate (there is no data to say they get lower). To convert the HEP (human error probability) into a frequency, one simply multiplies by the number of opportunities per time. Still, this issue probably has more to do with the fact that the HEP assumed are on the conservative side. Use of recovery factors, human redundancy, and inherent redundancy in the procedure itself are ways to handle this issue.

6 Tool 4: Understanding Systematic Bias

6.1 Bias and Heuristics

The following are some ways a bias can be described:

- A trap
- A distortion (the difference between what is seen and what is there)
- A systematic error (not statistical)
- Has the potential to be corrected for if it can be identified

A heuristic is a shortcut that bypasses cognition (i.e., thinking). Different biases provide the short-cut mechanism. We expect cognitive thinking to be less prone to bias (see **Section 6.2**).

Some examples of bias that have relevance to predicting and controlling major accident hazards are:

<u>Confirmation bias and belief perseverance</u> – we naturally look for corroboration for what we already know or believe, and we hold on to it. Rarely do people look to disconfirm. In the context of major accident hazards, what we don't know is far more relevant than what we do know. Why do we struggle with distinguishing between severity x or y when the hazard is already identified? Is the issue of taking multiple BPCS credits really that controversial when every barrier is already coupled by biased thinking (and we don't even realize it)? Are the consequences of summing versus not summing LOPA causes meaningful when accident sequences never play out in a linear fashion? I'm not arguing to ignore these issues, but to understand the context and prioritize. We already know these things, and the uncertainty is small. The uncertainty is very large for what we do not know, and worse if we don't look for it.

<u>Hindsight bias</u> – our inability to correctly remember or interpret past events, and our need to find linear causality (simplification) for explaining past events, produces a powerful distortion. If you've ever watched an accident video and thought, "How could they let that happen," you've fallen victim to

hindsight bias. If you've ever read multiple accident reports from different sources on the same accident, and each report states a different cause or causes, this demonstrates hindsight bias at work.

Information bias – more information produces better results. Not always. What more information does do is make us more (falsely) confident in our results. This bias runs rampant in quantitative analysis. We drill down and sharpen certain variables in a calculation only to ignore other variables whose uncertainty far exceeds what we are attempting to calculate.

Tunneling bias – ignoring or missing sources of uncertainty. This bias may be the umbrella for the others. Any number or calculation produced should have an uncertainty associated with it.

Commitment bias and capture error – these manifest themselves in many accident sequences where an operator has gotten on the wrong equipment, or followed the wrong procedure or action.

Independence trap – simply becoming conscious of an IPL (independent protection layer) or barrier results in a loss of independence of that barrier. This is the most insidious of all traps and has been written about by McLeod [9] and Hopkins [10]. How does this trap work? Say I have two IPLs, x and y, and I am aware of both. Any decision made for one will be influenced by the other. For example, if x is bypassed we still have y. If x fails we still have y. If x works we don't need y. And, it would be nearly impossible for x and y to fail at the exact same time. While this reasoning may be technically and logically true, it can produce a gradual and subtle increase in risk and worst-case it can defeat all of the safety barriers for a scenario.

Normalization of deviation – There have been many examples of this since the term was first coined, and is sometimes why accidents take years to manifest themselves.

Group-think – the presumption that when decisions are made in small groups, everyone will (or must) agree with the decision. This trap has been extensively studied, and can result in bad decisions.

The Turkey trap – see **Section 8.1**.

6.2 How Humans Think

Psychologists, notably Daniel Kahneman [14], have written extensively about how humans think. Two distinct styles are recognized:

- Type 1 – Fast, intuitive, efficient. Not consciously thinking. Zoned out. Prone to bias and irrationality. If you've ever driven to work and don't remember doing it, you've experienced Type 1 thinking. Actually, humans spend the vast majority of their day in Type 1 thinking, while performing tasks, etc. most of the time very reliably.
- Type 2 – Slow and inefficient. Hard. Careful and rational. Conscious thought. Taking exams, writing procedures, performing calculations, require Type 2 thinking. Try to maintain Type 2 thinking *continuously* for more than a few minutes. It can become exhausting.

Type 2 thinking is not error free (see the definition of mistake in **Section 3.4.2**) however, in certain contexts it can be superior to Type 1 thinking. Obviously, because your brain is turned ON, which we assume is better for most people when performing tasks and making decisions.

It is impractical to expect any one to spend their day in Type 2 thinking. The so far elusive goal of Human factors Engineering is this:

> To jolt the operator out of Type 1 thinking into Type 2 thinking, just before he is about to make a catastrophic mistake.

For example, say an operator is staring blankly at a sign that states 'Reactor Valve A' getting ready to open that valve, but that the operator intended to be on 'Reactor Valve B'. As the operator reaches for the handle a small jolt of electricity or heat or something shocks him, and an audio output declares "You are on the wrong valve!"

To achieve this takes some serious human factors engineering and technologies that probably do not exist yet, or not available to general consumer.

7 Tool 5: Improving Barrier Effectiveness using Bow-tie

Bow-tie drawings (the left side) may be used as an effective means to communicate how barriers may fail due to human or organizational factors. Task analysis would be used to identify escalation factors (something that can cause a barrier to fail) as well as escalation factor controls (that prevent interference of the escalation factor). The same analysis can identify failures that could impact multiple barriers.

These escalation factors on the bow-tie drawings represent leading indicators of process safety and could be monitored for failure. The bow-tie can be quantified to determine barrier probability of failure. The bow-tie is the best way to show how barriers can fail and the controls to prevent failure.

Common escalation factors related to Barrier failure:

- Errors made by Stores
- Lack of technical competence
- Contractors (where they touch the barrier)
- Poor Maintenance procedures and practices
- Improper restoration following maintenance
- Lack of Change Management
- Poor Work processes
- Lack of Communications
- Wrong specification

8 Handling Uncertainty

Typically we treat uncertainty in a quantitative way by performing something like a sensitivity analysis (what-if this input value is changed to this) or by acknowledging data uncertainty intervals (i.e., confidence limits) for the data we use. Typically this handles uncertainty associated with random hardware failure.

There are at least two other types of uncertainty (see **Section 2.4.2** for others),

- That which we know about but do not account for, i.e., systematic human error, the subject of this paper
- That which we don't even know about (i.e., referred to as a Black Swan) addressed below

8.1 Black Swan Theory

I won't attempt to explain or even summarize Black Swan theory [11]. I can't do it justice. Further, I would not listen to or read any second hand explanation of it, other than from the book itself, reference [11]. It is an important book and should be read by every practitioner of process safety risk management.

Psychologists have recognized a fundamental flaw (a bias) in humans thinking that we can project (predict) the future based on the past. This is obviously a trap that PHA teams still fall into. Many of the most infamous accidents in Industry only occurred after years of performing the same task or operation "safely." These incidents could look something like **Figure 1**, adopted from Ref [11]. From the perspective of the turkey, the probability of a catastrophic incident is 1e-3 (looking backwards) the day before it strikes. In hindsight, the catastrophic event seems readily predictable (Americans eat turkeys on Thanksgiving), and the events leading up to it can be described in a nice linear fashion (feed, water, repeat, later kill). This is hindsight bias at work.

The best we can do to handle this kind of uncertainty is to maintain a sense of vulnerability and continue to use work processes (e.g., hazard identification, change management, preventive maintenance, etc.) and tools (e.g., databases, procedures, task analysis, reliability analysis, etc.), to manage what we can predict, in hopes of effecting what we cannot predict.

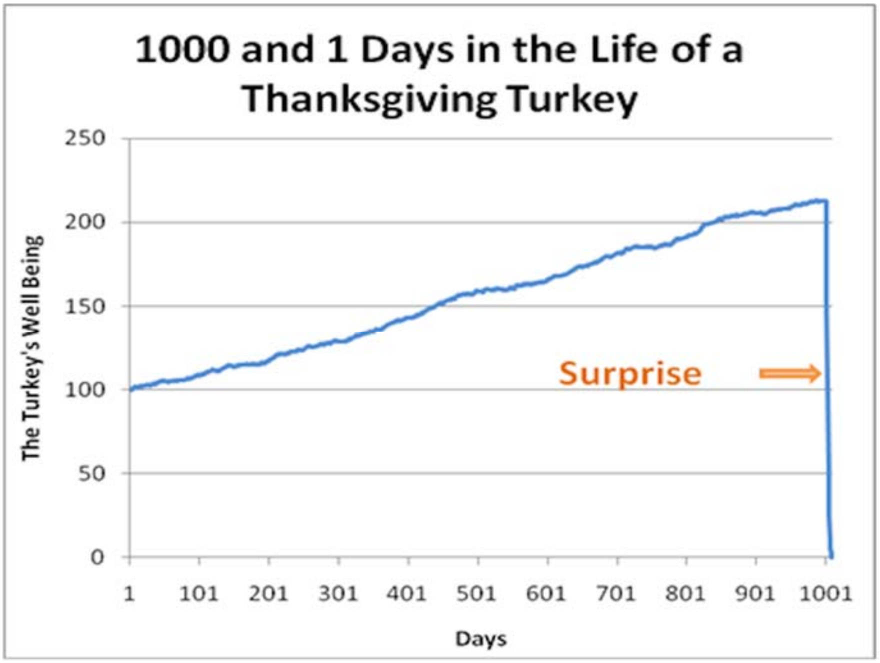

Figure 1. 1e-3 Probability of Catastrophe based on Past History

9 A Short Task Analysis Checklist

The following checklists represents a sample of the types of questions asked during a Task Analysis. A follow-up walk-thru to actually observe the task is always necessary. It is useful to observe how different the below checklists are compared to the standard human factors checklists used in PHA.

For any scenario with a human error initiating event that goes to LOPA, ask the team to answer the following questions:

- What is the cause of the error ("The operator." Why? "Wasn't paying attention." Why? "Gets pulled away during task." Why? etc.)
- What are the administrative controls in place (immediate feedback, procedure, permit, tagging, LOTO, human redundancy, housekeeping audits, etc.) that could allow recovery from the error?
- Is the task skill/ rule/ or knowledge based? Often tasks will be a combinations of the 3.
- How frequent is task performed? How long does the task take to complete?
- What is the qualification for the job?
- What Tools are required for the job?
- Can the task cross shift-change?
- What role do Contractors play in the task? Ask the same set of questions focusing on the Contractor.

For any scenario with a human response to abnormal condition (e.g., critical alarm barrier) that goes to LOPA, ask the team to answer the following questions (some of these can be observed directly during the walk-through):

Focus on the big 4: perceive, interpret, diagnose, respond (the first 3 represent the degree of situational awareness)

- Perceive: How does the Operator become aware of the alarm? Is the alarm prioritized? Is a dedicated alarm screen used? Are gray-scale graphics used?
- Interpret: What are the alarm metrics (how many alarms per hour, false alarms, etc.)? Do the metrics conform to good engineering practice? Are instrument signals trended? Is there a policy that every critical alarm that comes in must be treated as real? Are there periodic (e.g., quarterly) talk-throughs or drills for the alarm?
- Diagnose: Is the scenario known or novel? How many potential causes? How long does each cause take to diagnose? What actions are required?
- Respond: see checklist above.

For any scenario with a human barrier used in LOPA, ask the team to answer the following questions (some of these can be observed directly during the walk-through):

- What are all the human touch points and how can failure occur? E.g.,
    - Supplier
    - Purchasing
    - Stores
    - Maintenance

  - Quality Control
  - Engineering
  - Operations
  - Management
- What are the Contractor roles in the barrier?
- Has the barrier Specification (sizing, set-point, response time, etc.) been independently validated?

For any barrier or IPL that requires good situation awareness (SA*) (e.g., response to abnormal situation - AS), the following items should be looked at:

- Individual factors
  - How good is the operator's 'mental model' of the plant related to the AS (process conditions, equipment status, maintenance activities, etc.)?
  - What is the level of 'Automaticity' (how much cognition does the operator have to engage to resolve the AS)?
  - How well are the goals defined to resolve the AS?
- System factors
  - Interface design (alarm rationalization, use of gray-scale graphics, etc.)
  - System complexity (How many potential causes of the AS are there?)
  - Level of automation (Is there so much automation the operator has lost their hands-on operating skills?)

*SA is the difference between what is actually happening in the plant and what the operator thinks is happening.

These short checklists give you an idea of what task analysis is for each.

10 Summary

Improving Human factors in PHA-LOPA requires a diverse toolkit of methods and theory. Identification, prediction, and control of the human factors role in major accident hazards is currently soft. There is not a quick easy answer, however, improvement especially related to low hanging fruit is readily achievable.

PHA-LOPA facilitators need to begin practicing the methods of human factors and human reliability to improve said practice as it relates to major accident hazards. The skills a facilitator gains by practicing human factors and reliability will directly transfer to all of PHA-LOPA to improve decision making, the ability to write human factors scenarios, and better awareness of when further study is needed.

11 References

[1] Norman, Don. *The Design of Everyday Things*, Revised and Expanded Edition, Basic Books, New York, 2013.

[2] Chemical Safety Board Website. http://www.csb.gov/videos/csb-video-excerpts-from-dr-trevor-kletz/

[3] J. Robert Taylor. *Human Error in Process Plant Design and Operations – A Practitioner's Guide*, CRC Press, Taylor & Francis Group, Boca Raton, 2016.

[4] Kirwan, Barry. *A Guide to Practical Human Reliability Assessment*, CRC Press, Taylor & Francis Group, Boca Raton, 1994.

[5] CCPS. *Guidelines for Engineering Design for Process Safety* 2nd Edition, 2012.

[6] HSE. Review of Human Reliability Assessment Methods, 2009.

[7] Gawande, Atul. *The Checklist Manifesto: How to Get Things Right*, Metropolitan Books, 2009.

[8] Kletz, Trevor. *Hazard Analysis – A Quantitative Approach to Safety*, I.Chem.E. Symposium Series, No. 34, 1971.

[9] McLeod, Ronald W. *Designing for Human Reliability – Human Factors Engineering in the Oil, Gas, and Process Industries*, Gulf Professional Publishing, 2015.

[10] Hopkins, Andrew. *Disastrous Decisions – The Human and Organizational Causes of the Gulf of Mexico Blowout*, CCH, Australia, 2012.

[11] Taleb, Nassim Nicholas. *The Black Swan* 2nd Edition, Random House, New York, 2010.

[12] Swain and Guttmann. NUREG/ CR – 1278. *Handbook of Human Reliability Analysis*. 1983.

[13] Reason, James. A Life in Error – From Little Slips to Big Disasters, CRC Press, Taylor & Francis Group, Boca Raton, 2013.

[14] Kahneman, Daniel. *Thinking Fast and Slow*, Farrar, Straus & Giroux, New York, 2011.