Analysing ``Human-in-the-Loop'' for Advances in Process Safety: A Design of Experiment in a Simulated Process Control Room

*Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)*
*Edited by* Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio

*Published by* Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P481-cd

# Analysing "Human-in-the-Loop" for Advances in Process Safety: A Design of Experiment in a Simulated Process Control Room

Chidera Winifred Amazu
*SAfeR - Department of Applied Sciences and Technology, Politecnico di Torino, Italy.*
*E-mail: chidera.amazu@polito.it*

Houda Briwa
*Technological University Dublin, Dublin, Ireland. E-mail: houda.briwa@tudublin.ie*

Micaela Demichela
*SAfeR - Department of Applied Sciences and Technology, Politecnico di Torino, Italy.*
*E-mail: micaela.demichela@polito.it*

Davide Fissore
*SAfeR - Department of Applied Sciences and Technology, Politecnico di Torino, Italy.*
*E-mail: davide.fissore@polito.it*

Gabriele Baldissone
*SAfeR - Department of Applied Sciences and Technology, Politecnico di Torino, Italy.*
*E-mail: gabriele.baldissone@polito.it*

Maria Chiara Leva
*Technological University Dublin, Dublin, Ireland. E-mail: mariachiara.leva@tudublin.ie*

Control room operators are crucial to ensuring safety in safety-critical environments, as major risk process plants, especially when addressing critical process deviations that could lead to process disruptions or accidents. These operators face increased cognitive loads being more involved in tasks that exert their cognition with less manual or physical engagement. Therefore, process safety analysis should integrate key dynamic elements, including the operators' cognitive states, to allow better predictions. We aim to investigate the impact of the human system interfaces, in this case, two conditions of alarm design (prioritized and non-prioritized) and intervention procedures (paper vs. computer-based), on the operators' cognition (e.g., situational awareness, mental workload) given a set of scenarios and how these impact operators' performance and process safety. We also assess how other performance-shaping factors, such as task complexity, communication, and more during the process operations (alarm handling and intervention), contribute to managing safety. Therefore, we present a design of an experiment and a case study of a simulated formaldehyde production plant with which we plan to investigate the operators' and systems' behavior during abnormal process operation. Results are yet to be obtained from this study. Subsequent work on modeling process safety for early warnings and optimizations can benefit from this experimental design and the data to be collected and, even so, from including the data on operators' cognition during analysis.

*Keywords*: control rooms, process safety, design of experiment, human-in-the-loop, cognitive science

## 1. Introduction

Process industries are major-risk safety critical process plants because of the type of materials being handled and the adverse impact a process safety event can have on the environment, health, safety, and cost. Accident reports in these industries have shown that human error is a major contributor, with statistics for process industries at about 70% Bhavsar et al. (2017). This is mostly associated with control room operators who monitor plant operations and intervene in process upsets. These accidents have led to excluding humans as much as possible from the loop and introducing more process automation.

Efforts to reduce human error in control room operations have included assessing error probabilities during design and installation. However, it is important to consider the impact of human-machine interaction on operators' cognitive states, performance, and safety during normal and abnormal process operations. To improve safety, experts have recommended matching operator support to workload levels and increasing situational awareness while reducing cognitive fatigue. Further suggestions are for decision-makers to evaluate how supporting elements, such as the human-machine interfaces, procedures, and more, contribute to safety in human-in-the-loop configurations.

This study aims to develop an experimental setup that simulates control room operation to test human-in-the-loop conditions during alarm handling and process control of abnormal situations in a formaldehyde production plant. To support the operators in this setup and test the impact of the element of support and interaction, a simulated human-machine interface, procedure, and alarm rationalization support is developed following guidelines HPOG (2021), ISO (2022). This study considers two procedural guides: paper vs. screen-based procedures and a case of alarm vs no alarm prioritization. The impact of these two mediums of procedural guidance, alarm design, and other performance-shaping factors is to be observed through this study. Also, the aim is to see how such a setup can advance human factors analysis and how insight from this can address in-practice activities in process plants, such as during training or emergency drills. The study is set up in a university environment. The operators' psycho-physiological data would be collected with qualitative data from surveys and process data from the simulated plant. Available monitoring tools like eye trackers and electroencephalogram (EEG) caps would be used to acquire the psycho-physiological data.

## 2. Methodology - Design of Experiment

A between-subject design will be adopted in this study to test two key conditions of the human system interfaces: alarm prioritization (with and without) and procedure (screen vs. paper-based), and a within-subject design to test three scenarios (that is, all groups test these three scenarios). Other independent variables are also considered as inputs for this setup.

### 2.1. *Case study*

The present study uses a simulated interface of a formaldehyde production plant with modifications to the actual plant design in Demichela et al. (2017). The plant produces around $10\,000\,\mathrm{kg/hr}$ of $30$ percent formaldehyde solution, operating the partial oxidation of methanol with air:

$$\mathrm{CH_3OH} + \frac{1}{2}\,\mathrm{O_2} \underset{\mathrm{r_1}}{\longrightarrow} \mathrm{CH_2O} + \mathrm{H_2O}$$

Also, a secondary reaction occurs, which completes the oxidation to carbon monoxide and reduces the yield of formaldehyde in the reaction:

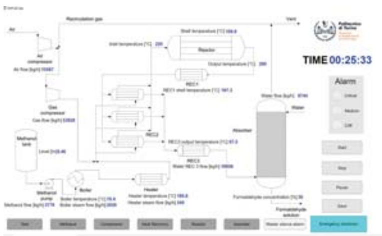
$$\mathrm{CH_3OH} + \frac{1}{2}\,\mathrm{O_2} \underset{\mathrm{r_2}}{\longrightarrow} \mathrm{CO} + \mathrm{H_2O}$$

The modified simulator comprises six sections (Tank, Methanol, Compressor, Heat recovery, reactor, and absorber), as shown in Figure 1a. The developed simulator has 80 alarms, having different prioritization levels: 1 (low - in yellow), 2 (medium - in orange), and 3 (high - in red), including nuisance alarms. Scenario-based simulation of different process safety-related events is an approach taken for this study. Here three scenarios (S), 1 to 3, lasting 15, 15, and 17 minutes, respectively, have been selected to be simulated by the participants. The scenario complexity increases as the participants progress from 1 to 3.

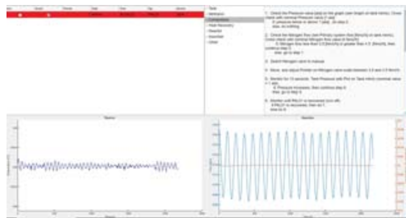
S1: failure of a pressure indicator control (Tank Section). *Hazardous event: Implosion of a methanol storage tank Goal: prevent the activation of pressure switch PSL01 within* $7$ m*inutes after initial alarm*

S2: failure of nitrogen valve primary source (Tank section). *Hazardous event: Implosion of a methanol storage tank Goal: prevent the activation of pressure switch PSL01 within* $7$ m*inutes after initial alarm*

S3: failure of temperature indicator control (Heat Recovery section). *Hazardous event: Over-*




(a) Overview display - middle interface



(b) Support display - right interface

Fig. 1.: Interface 1 and 2

*heating of the reactor. Goal: prevent the activation of temperature alarm high high (TAHH14) within the* 18 m*inutes of the test.*

The interfaces have also been designed to support operators using standard color schemes to provide cues on meters and the color scheme for alarms after prioritization (red color: critical, orange: medium, yellow: low, white: alarm inactive). In addition, the operators will be provided with the intervention procedures designed in this study for handling abnormal situations. The authors further discuss the experimental methodology and subsequent sections in this paper using scenario one as a case study.

#### 2.1.1. *Scenario 1: S1*

Scenario 1 simulates the failure of a pressure indicator controller PIC01 in the tank section. The PIC01 signals the primary nitrogen valve the need to open and close. Therefore, due to the failure of PIC01, the operator has to control the opening and closing of the valve. Figure 2 shows a timeline of events and tasks simulated for this scenario. The process is simulated with a ramp of the primary nitrogen flow, which is to be followed by an alarm on FAL01 (which has been made faulty in this study). To simulate a more process safety situation and observe the support of alarm prioritization, the operator must wait to hear and identify the critical alarm (PAL01), activated with a high noise level alarm (LAH01). The operator is to follow the intervention procedure for PAL01 and recover the process within the available time.

#### 2.1.2. *Assumptions*

- one assumption is that (PIC01) is faulty,

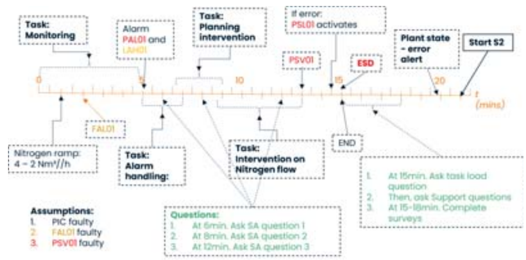



Fig. 2.: Events and task timeline for Scenario 1

- the flow alarm low (FAL01) is faulty,
- the pressure safety valve (PSV01) is faulty, and
- the operator can only start intervention on the first critical alarm.

### 2.2. *Setup*

The experiment is set up in a lab located in a university environment and organized to replicate similar conditions in a process control room as closely as possible. Three interfaces are to be used in this study. Interface 1 is the central monitoring interface with a display of the plant overview. The second interface shows trends of critical processes in the reactor and absorber, specifically the formaldehyde concentration and inlet water into the absorber, including the maximum reactor temperature. It also displays the alarm list and the procedures for each alarm within the different plant sections. The participants are also provided with the operating procedures and intervention procedures on paper which would be used by the control group and a second group that will assess the impact of both alarm prioritization and paper-based procedural support. Finally, the third interface would display any of the mimics of any plant sections opened by the participants during the test.

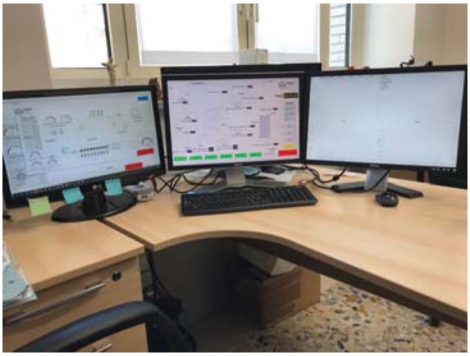


Fig. 3.: Human System Interfaces Setup

### 2.2.1. *Participants and Experiment procedure*

The participants to be recruited for this study are chemical engineering students and researchers at the university. The participants will be divided into three groups; a control group with only paper-based procedures and no alarm prioritization, a second group that tests alarm prioritization and paper-based procedure conditions, and the last group with alarm prioritization and screen-based procedures. Each group tests the different scenarios (S1 - 3) with increasing levels of complexity to effectively capture the impact of the supports as task complexity increases. Initially, the participants will be given an information sheet to be informed about the experiment and the data collection type. This would be followed by a form to sign-up for the experiment and indicate their availability. On the day of the experiment, they will sign the consent form and complete a questionnaire on personal information such as age, vision issues, and dominant hand in use. These are important for collecting objective measures using the eye tracker and EEG. They will be informed that this data will be anonymized and only accessible to the researchers whose emails are also indicated on the information sheet.

This would be followed by training on the operating procedure and simulated interfaces, including the type of available support. After this, the monitoring tools (eye-tracker and EEG) will be worn and calibrated. The participants then test the three scenarios (each 15 - 17 minutes), during which data on their psycho-physiological state are collected. To assess the effects of the procedures on instantaneous situational awareness and task load, the operators, through a think-aloud approach, are asked two questions each at the 6th, 8th, and 12th minute of the 15 to 17 minutes available to assess their situational awareness during the process operation. After each experiment, the operators are asked to rate their overall situational awareness and workload using SART and NASA-TLX. They are also asked to rate their task load and the supports (intervention procedure and alarm design).

This same process is followed for all three scenarios. At the end of all scenarios, based on observation and the need for clarification, the participants are further shown specific interface activities to understand the reason for certain actions at those points. The participants are debriefed after the experiment.

### 2.2.2. *Tasks*

The participant's tasks include following both operating and intervention procedures. Both procedures are integrated based on a control room human-in-the-loop task flow and are broken down into four phases; Monitoring, Alarm Handling, Recovery Planning, and Intervention (Troubleshooting, Control, and Evaluation). This subsequently enables a critical cognitive safety task analysis approach to evaluate the overall procedures or tasks and understand tasks that are bottlenecks to overall performance and safety. The breakdown of tasks and sub-tasks is also organized around macro-cognitive functions (Detection, Understanding, Response planning, and execution). This is important because breaking these procedures into subtasks, specifically the intervention procedures usually classified as action tasks, further reveals a combination of different levels of macro cognitive functions Chang et al. (2014). Figure 3 shows the intervention procedure for PAL01. The failure or success of the participant on the subtasks is based on time and the error (wrong or incomplete diagnoses or actions). The variables taken into consideration in this setup are discussed below:

#### 2.2.3. *Independent variables*

The key independent variables in this study are the alarm design (prioritization), intervention procedures, and task complexity. However, other variables like noise, communication, simultaneous actions/Interactions, and time available are measured in this study.

***Alarm prioritization:*** Alarm rationalization is a crucial practice in control rooms where alarms are identified, assessed, and ranked to minimize their quantity and guarantee that they provide relevant and useful data to operators ISA and ANSI (2009). This simulation study will assess the impact of two alarm configurations on the operator's mental workload, situation awareness, and performance in responding to critical alarms. In the three scenarios, the level of nuisance alarms will differ, with the third scenario having a high number of nuisance alarms, creating a flood of alarms. Another configuration is alarm prioritization, represented by priority indications in the alarm list and a visual cue that distinguishes critical, medium, and low alarms. Figures 4 and 5 show the alarm and no-alarm prioritization conditions for scenario 1. Group 1 will test without alarm prioritization, whiles groups 2 and 3 will test with alarm prioritization.

| Ack | Sound | Priority | State | Time | Tag | Section |
|---|---|---|---|---|---|---|
| ☐ | ☐ | 3 | Active | 00:05:03 | LAH01 | Tank |
| ☐ | ☐ | 1 | Active | 00:05:04 | PAL01 | Tank |

Fig. 4.: List of alarms with prioritization

| Ack | Sound | Priority | State | Time | Tag | Section |
|---|---|---|---|---|---|---|
| ☐ | ☐ | 1 | Active | 00:05:03 | LAH01 | Tank |
| ☐ | ☐ | 1 | Active | 00:05:04 | PAL01 | Tank |

Fig. 5.: List of alarms without prioritization

***Procedures:*** Intervention Procedures are key supports for control room operators to manage safety. The intervention procedure in this study is written for each alarm as a hierarchical rule-based task representation format. It is further presented on both paper and screen for the operators. Each intervention procedure is written under three broad task contexts; troubleshooting, control, and evaluation. An example is shown in Figure 6. In this study, two mediums of procedure presentation are studied; paper-based procedures and screen-based procedures. Here, Groups 1 and 2 will test the paper procedures, whiles Group 3 will test with the screen-based procedure.

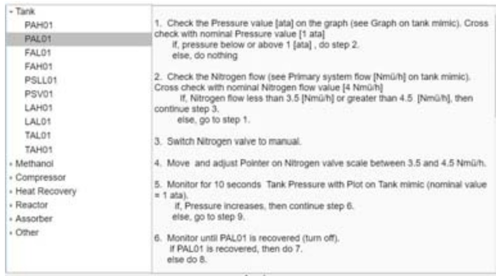



Fig. 6.: Screen-Procedure for PAL01

***Task complexity:*** The task complexity corresponds to the selected process safety events (scenarios). Three scenarios are considered here; each represents one level of complexity: low, medium, and high. Task complexity combines the number of task steps, control actions, and the number of alarms. These increase from scenarios 1 to 3.

#### 2.2.4. *Dependent variables*

***Mental workload:*** The impact of mental workload on cognitive performance is a crucial research area in aviation, chemical, and nuclear power plants, where operational alertness and attention are vital. Particularly in process industries, where operator intervention is often necessary to keep the plant stable. Operators rely on mental process models, prior knowledge, and experience in process control rooms to maintain this stability. When the task demand exceeds the capabilities of a human, it can result in human errors. On the other hand, a discrepancy between an individual's expectations and the available mental resources to perform a task leads to cognitive workload. Three methods are commonly used to evaluate mental workload: subjective methods, performance-based methods, and methods based on psycho-physiological measurements. Subjective methods rely on the user's self-report of their sense of the task conditions. One widely used subjective method is the NASA-TLX load index assessment method Hart and Staveland (1988), which assesses physical and mental demands Fernandes and Braarud (2015).

***Situational awareness:*** Optimal Situational awareness levels (perception, comprehension, projection) are essential for control room operators to perform effectively during normal operations and, even more importantly, given abnormal events. This becomes apparent in current interaction configurations where the operator is kept outside the loop due to automation and increased complexity with lots of information to process. Therefore, understanding operators' situational awareness helps adapt proper supports that enhance decision-making during process operations and in designing training plans. Different methods and metrics have been used to assess situational awareness. Subjective methods, such as the situation present assessment method (SPAM) and situation awareness global assessment technique (SAGAT), are used for situational awareness assessment Endsley (2021). Objective techniques such as eye-tracking are used through experimental setups to assess situational awareness Bhavsar et al. (2016).

There have been concerns about using these metrics as a standalone, and recommendations are to combine them with other objective and subjective methods for better situational awareness classification or predictions Endsley (2021),Gao et al. (2013). In this setup, a few subjective and objective methods have been selected to address these concerns as applicable to this experimental study.

#### 2.2.5. *Control Variables*

The following variables are controlled in this study, the time of day (for which the participants of the different groups are randomized), demographics, and training.

#### 2.2.6. *Control Performance:*

The participant's performance on the task (ToT)/response time will also be collected, with a record of the error on the macro-cognitive levels.

## 3. Data Collection and Selected Metrics

Both subjective and objective data are considered in this study. Objective measures will be collected using monitoring sensors that can help infer the cognitive states of operators to predict their workload or SA ahead of an accident. According to Bhavsar et al. (2016), this has not been explored well in the chemical process industries. To ensure good data quality during collection precisely for the EEG, the participants will be asked to move as little as possible, including head movements.

### 3.1. *Objective measures*

***A. Process Measures - Electroencephalogram:***

Electroencephalography, or EEG, captures the brain's electrical activity caused by neural oscillations. In human factors-related research, primarily in aviation, nuclear power plant, and petrochemical industries, EEG signals have been shown to provide information about the relative amounts of mental effort expended to perform a task Costa (1993), workload Staal (2004), Iqbal et al. (2020) among other cognitive measures. This experiment will collect EEG data using the semi-dry 24-channel SMARTING mbt device (MbrainTrain, Serbia), a head-mounted system with a semi-dry and wireless electrode layout. The data is recorded in real-time and can be monitored through the device-supported SMARTING app either on a smartphone or laptop via Bluetooth. The cap will be mounted on the head of the participant before the start of the experiment.

***B. Process Measures - Eye Tracking:***

Tobii Pro Glasses 3 will be used in the data collection. The sampling rate is 50 Hz. The glasses come with wide-angle scene cameras and can provide comprehensive and accurate gaze data. Tobii Pro 3 enables data collection on gaze and pupil metrics used as workload, situational awareness, and fatigue metrics.

Before the simulation runs, after the participant has worn the glasses, a 1-point calibration is done to ensure the validity of eye-gaze data. It also enables focused observation of areas of interest (AOI) on specific tasks. Based on the Overall task, there are 6 AOIs (1 on Interface 1, 4 on Interface two, and one on Interface 3). However, this further varies depending on the sub-tasks.

The selected eye tracking and EEG metrics for SA and workload are shown in Figure 7. A combination of these monitoring tools and the other selected measures will be helpful in subsequently assessing both temporal and overall situational awareness and mental workloads and better classifying these cognitive states as compared to having just one data collection mode, as also emphasized by Gao et al. (2013). A summary of these measures already used for similar process industry-related studies can be found in a review by Amazu et al. (2022).

Objective data is further collected on the interaction from the system, including data on the process during the simulation runs.

***C. Direct SA measures - SPAM:***

SPAM is an objective SA measure used to assess operators' knowledge of a situation in real-time and verbally at specific points during the task. The questions developed in this study to assess the situational assessment of the participants at three points during each scenario run are asked without a simulation freeze but using a concurrent think-aloud approach. The responses, response accuracy, and time to respond will be recorded.

***D. Direct task load and overall procedure support assessment:***

The participants are asked to rate the task load on a scale of 1 - 5 just after the action task, after assessing the support of the intervention strategy (SA level 3 question). The participants are further asked to rate the overall support of the paper or screen-based procedure and the alarm prioritization using the same scaling at the end of each simulation round.

## 3.2. *Subjective Measures*

The subjective metrics used are the Nasa-TLX for workload and SART for overall situational awareness assessment.

***A. Nasa-TLX***

The assessment technique needs to be sensitive, diagnostic, selective, non-intrusive, reliable, and easy to implement to assess workload. The NASA Task Load Index (TLX) has proven to meet these criteria Rubio et al. (2004). It is the most commonly used self-report index of mental workload and is often used as a reference point to evaluate the effectiveness of other measures, theories, or models. To derive an overall workload score, the NASA-TLX is used. This subjective technique considers six subscales, including mental demand, physical demand, temporal demand, performance, effort, and frustration level. These subscales have been found to reduce variability and provide diagnostic information about workload sources Cao et al. (2009).

***B. Direct SA measures - SART:***

The major limitation of subjective metrics is that they can only access the overall workload or situational awareness from the impact of the factors and cannot reflect changes in workload during the execution of operations Gao et al. (2013). And this overall approach for subsequently performing task analysis on the intervention procedure does not give us insight into a peak or critical steps that can influence system performance. Hence the need to integrate objective measures and, even so, multi-psycho-physiological metrics.

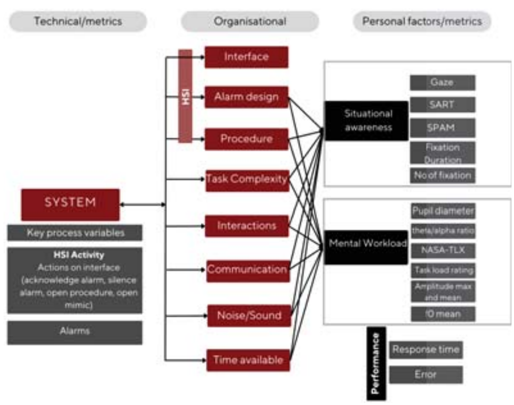



Fig. 7.: PSFs and metrics considered in this setup

## 4. Discussion and Future Work

In this study, situational awareness would be measured using the metrics within the allocated box in Figure 7. Likewise, the mental workload would be assessed using the metrics assigned in Figure 7. Results will show the influence of emergent events and the applicable condition of procedural design principles and alarm management to supporting operators. Using statistical analysis, the subjective and objective measures will be used to understand the influence of the different performance-shaping factors (PSFs) and support tools on each experiment group. In further work, the authors intend to use this data to classify workload and situational awareness, understand and build predictive models of human behavior, and update prior error rate estimations for process safety data modeling.

## 5. Conclusion

In conclusion, the present study proposed a methodology for assessing the impact of human system interfaces and associated performance-shaping factors on cognitive states of situational awareness and mental workload and how this impacts process safety in the context of process control rooms. This experimental study setup is proposed to be adopted by process plants and other safety-critical related domains for more effective and efficient optimization of human system interfaces that provide the necessary support to enable operators to make better decisions, reduce errors, and manage safety. Furthermore, the results of this study will provide insights into the performance-shaping factors and how they impact different HITL configurations, specifically for the study groups in this research. It will also provide data and context for informed human performance and process safety predictive models.

### Acknowledgement

The CISC project funds this work as part of the grant H2020-MSCA-ITN-2020 with grant agreement no. 955901

## References

Amazu, C. W., M. Demichela, and D. Fissore (2022). Human-in-the-loop configurations in process and energy industries: a systematic review. pp. 3234–3241. Research Publishing, Singapore.

Bhavsar, P., B. Srinivasan, and R. Srinivasan (2016). Pupillometry Based Real-Time Monitoring of Operator's Cognitive Workload to Prevent Human Error during Abnormal Situations. *Industrial and Engineering Chemistry Research 55*(12), 3372–3382.

Bhavsar, P., B. Srinivasan, and R. Srinivasan (2017). Quantifying situation awareness of control room operators using eye-gaze behavior.

Cao, A., K. K. Chintamani, A. K. Pandya, and R. D. Ellis (2009). Nasa tlx: Software for assessing subjective mental workload. *Behavior research methods 41*, 113–117.

Chang, J. Y., D. Bley, L. Criscione, B. Kirwan, A. Mosleh, T. Madary, R. Nowell, R. Richards, E. M. Roth, S. Sieben, and A. Zoulis (2014). The SACADA database for human reliability and human performance. *Reliability Engineering and System Safety 125*, 117–133.

Costa, G. (1993). Evaluation of workload in air traffic controllers. *Ergonomics 36*(9), 1111–1120.

Demichela, M., G. Baldissone, and G. Camuncoli (2017). Risk-Based Decision Making for the Management of Change in Process Plants: Benefits of Integrating Probabilistic and Phenomenological Analysis. *Industrial and Engineering Chemistry Research 56*(50), 14873–14887.

Endsley, M. R. (2021). A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM. *Human Factors 63*(1), 124–150.

Fernandes, A. and P. Ø. Braarud (2015). Exploring measures of workload, situation awareness, and task performance in the main control room. *Procedia Manufacturing 3*, 1281–1288.

Gao, Q., Y. Wang, F. Song, Z. Li, and X. Dong (2013). Mental workload measurement for emergency operating procedures in digital nuclear power plants. *Ergonomics 56*(7), 1070–1085.

Hart, S. G. and L. E. Staveland (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, Volume 52, pp. 139–183. Elsevier.

HPOG (2021). HPOG - Human Performance Oil Gas. Best Practice in Procedure Formatting. Retrieved January 30, 2023.

Iqbal, M. U., B. Srinivasan, and R. Srinivasan (2020). Dynamic assessment of control room operator's cognitive workload using electroencephalography (eeg). *Computers & Chemical Engineering 141*, 106726.

ISA and ANSI (2009). Isa-18.2: Management of alarm systems for the process industries. *International Society of Automation. Durham, NC, USA*.

ISO (2022). ISO 11064, BSI Standards Publication Ergonomic design of control centres Part 10 : Introduction to the control room design series of standards. (August).

Rubio, S., E. Díaz, J. Martín, and J. M. Puente (2004). Evaluation of subjective mental workload: A comparison of swat, nasa-tlx, and workload profile methods. *Applied psychology 53*(1), 61–86.

Staal, M. A. (2004). Stress, cognition, and human performance: A literature review and conceptual framework.