EVALUATING SKILL-, RULE-, AND KNOWLEDGE-BASED DESIGN FOR CONTROL ROOM OPERATOR LEVEL 1 OVERVIEW DISPLAYS

A Thesis

Presented in Partial Fulfillment of the Requirements for the

Degree of Master of Science

with a

Major in Psychology

in the

College of Graduate Studies

University of Idaho

by

Joseph Carlos Vargas

January 2015

Major Professor: Steffen Werner, Ph.D.

Authorization to Submit Thesis

This thesis of Joseph Carlos Vargas, submitted for the degree of Master of Science with a major in Psychology and titled "Evaluating Skill-, Rule-, and Knowledge-based Design for Control Room Operator Level 1 Overview Displays," has been reviewed in final form. Permission, as indicated by the signatures and dates given below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor ______________________________Date__________
Steffen Werner, Ph.D.

Committee Members ______________________________Date__________
Brian P. Dyre, Ph.D.

______________________________Date__________
Jason Laberge, M.S.

Department Administrator ______________________________Date__________
Todd Thorsteinson, Ph.D.

Discipline's College Dean ______________________________Date__________
Andrew Kersten, Ph.D.

Final Approval and Acceptance by the College of Graduate Studies

______________________________Date__________
Jie Chen, Ph.D.

Abstract

Control room operators in industries like oil refining, chemical production, or nuclear power plants are required to continuously monitor a plant's status and act on process deviations. Level 1 overview displays provide the operators with support for plant situation awareness to detect process deviations that are indicative of overall system health. This study compared three level 1 overview display prototypes (Calm Water, Heat Map, and Visual Thesaurus) developed by individual design teams of members of the Abnormal Situation Management (ASM®) Consortium.

48 experienced control room operators participated in the study. Each experimental session lasted approximately 2 hours and consisted of a detailed training session followed by data collection. Operators completed process monitoring tasks on the level 1 overview displays while maintaining a healthy system in the secondary task (MAT-B II). Operator performance was assessed through multiple measures of accuracy, workload, and situation awareness. The results show that operators overall performed best on the Calm Water display and experienced the highest subjective workload with the Visual Thesaurus display. However, the visual Thesaurus display showed benefits for plant awareness. The Calm Water display was generally preferred by operators, while the Heat Map display was liked the least. The displays and their respective design elements are discussed in the context of ecological interface design and Rasmussen's skill, rules, and knowledge framework.

Acknowledgements

I would like to express my appreciation and thanks to my advisor Dr. Steffen Werner for providing support in my pursuit of this work. You have always pushed me and helped me better myself, no matter what I am working on. I would like to offer my most sincere gratitude for your guidance during my academic career. It has left a positive impact on my professional and personal life. I would also like to thank my committee members Dr. Brian P. Dyre and Jason Laberge for proving guidance and helpful feedback in support of my thesis work. I would especially like to thank you for your comments and suggestions before and during my thesis defense. My work has improved greatly due to your guidance.

I would like to thank my Parents Sam and Elaine Vargas, Sister Alexandra Vargas, and Grandparents Alex and Corrine Vargas for giving me the tools necessary to realize my dreams. Most importantly, a very special thanks to my beloved fiancée Megan Stewart for being patient in my pursuit for this work. Words cannot express how grateful I am for your support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Evaluating Skill-, Rule-, and Knowledge-based Design for Control Room Operator Level 1 Overview Displays

Control room operators in settings such as oil refining, chemical production, nuclear power plants, and paper manufacturing are required to continuously monitor and act on process deviations to maintain a healthy system. Display design in the control room is important in these applications, as they are typically high-risk environments where unchecked process deviations can lead to costly plant shutdowns or dangerous plant conditions (Vicente, 2007). It is the human's responsibility to monitor these systems to ensure that the plant is operating within safe tolerances (Vicente, 2007).

Since we know that people tend to struggle with tasks that involve monitoring complex systems (Woods, 1995), design teams have implemented overview displays in control rooms to alleviate some of the burden placed on operators by automating aspects of the control room and displaying pertinent plant status information on a single or multiple, related displays. Providing this overview of the plant has been made possible due to the integration of digital systems in the control room, which make it possible to fully integrate displays such that layout is no longer driven by the physical characteristics of the lights, dials, switches, and CRT displays that traditionally populate control rooms. This allows designers to allocate the information and interaction capabilities of the devices to better suit the needs to the human operator. These integrated displays provide operators with a big picture view of plant status and are called level 1 overview displays. The use of an overview display is a method used to mitigate the shortcomings in monitoring performance by reducing the workload on the operator (Jou, Yenn, Lin, & Chiang, 2009). Tharanathan, Bullemer, Laberge, Mclain, and Reising (2010) provided a graphical representation of the operator panel display hierarchy for a typical plant (Figure 1), illustrating the connection between levels of display and the type of information they provide to operators.

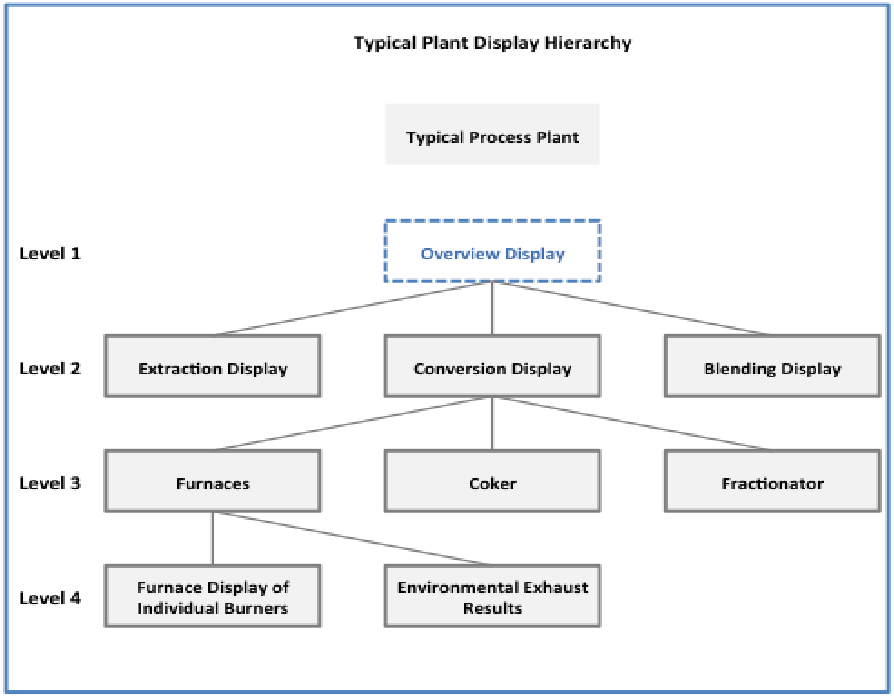

Figure 1. Illustration of a display hierarchy from top (level 1 overview display) to bottom (level 4 component-specific displays). The current study focuses on the level 1 overview display. At each level only one branch of the hierarchical tree is shown.

It is typical to have at least 4 levels of display. Multiple levels of displays are used to provide relevant information to operators, depending on the goals of a particular task. For example, the level of display used to provide overview information of an entire plant will differ from the level of display used to provide information about a particular part of the plant like the individual equipment. In this example, overview information is provided in a level 1 display and information about specific equipment is provided in a level 3 display. Specifically, level 1 differs from subsequent display levels in the operator display hierarchy in that it is primarily an overview display meant to facilitate operator system awareness "at a glance" (Reising, 2006), whereas level 2, level 3, and level 4 displays are intended to provide more detailed information about specific functional areas (units) of a plant or system. The current study focuses on level 1 overview displays and how their design can impact operator performance.

Level 1 Overview Displays

Level 1 displays are intended to provide support for plant situation awareness by providing an “overview” to detect process deviations that are indicative of overall system health (Tharanathan et al., 2010). A key goal for level 1 overview displays is to draw the attention of the control room operator when the automation has detected a process deviation (Xiao & Seagull, 1999). Furthermore, the displays must allow operators to select an appropriate action alternative, as determined by the type and severity of deviation. To effectively alert operators of any potential process deviation, designers are tasked with designing a system that is able to detect rare (and seemingly minor) occurrences of deviations (Parasuraman, Hancock, & Olofinboba, 1997), so the system must be very sensitive to even slight deviations in process conditions to ensure that no important deviations are missed. System designers are forced to engineer with a “fail safe” approach in mind, which drives the threshold of the system low enough to capture even the slightest possibility of a failure (Swets, 1992). In turn, this approach causes decision aids to perform such that they are sensitive to even minor changes in parameters, alerting operators of even the slightest deviations. This causes breakdowns in human-automation interaction, which can lead to poor decision-making performance due to increased false alarms emitted by the system that the operator must address. This places tremendous importance of the level 1 overview displays to effectively communicate process deviations as well as afford quick and efficient action towards resolution of the deviations.

It is clear that level 1 overview displays are a critical part of control room operators' day-to-day responsibilities, providing a "big picture" view of current plant status. This big picture view is intended to allow operators to detect and take appropriate corrective action on process deviations in the plant. This is especially important when dealing with abnormal situations, or plant conditions that have not necessarily been foreseen by designers and operators must evaluate information and take corrective action. Contextual information about process deviations is typically provided in the form of a level 1 overview display, which contains information about the process variables that provide insight about the health of the system, including the major equipment in the plant and provides alerts or notifications when the key variables have deviated from acceptable levels (Tharanathan et al., 2010).

The ability to provide operators with this plant overview is also important because it allows operators to view overall plant status at a glance, without having to spend time searching for information in sub menus or screens while also providing the opportunity for operators to investigate any process deviation further by using the level 1 overview as a top-level window into critical information about specific plant systems. These attributes (providing process deviation alerts and affording overall plant awareness) make level 1 overview displays a critical component in operator workstations.

Previous work has been done to identify the characteristics that a level 1 overview display must possess to assist operators' in detecting and taking appropriate action in regards to process deviations. Bullemer, Hajdukiewicz, and Reising (2002) provide a summary of these characteristics. The display itself should contain key status information such as alarms, key process parameters, process conditions, and an indication of the potential plant effects due to process deviations (e.g., what systems, components will be affected due to a current process deviation in a particular area). With this information readily available, level 1 overview displays are used to provide insight into the existence, severity, location, and direction of process deviations.

In the modern digital control system (DCS), this is accomplished with the integration of an automated aid (or decision support tool), which monitors plant conditions and provides relevant information to operators, in the current context, about process deviations. The decision support tool proves to be valuable in the context of highly complex systems such as the ones used in the process industry.

Situation Awareness

Situation awareness is defined as “the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” (Endsley, 1988).

Parasuraman, Sheridan, and Wickens (2000) noted that when a system provides decision-making functions automatically, it might reduce the operator’s awareness of the plant, including aspects of the work environment that consistently change. This leads to operators not being fully aware of changes in plant status, potentially leading to dangerous conditions due to lack of, or inappropriate, action in response to process deviations. Also noted by Parasuraman et al. (2000), when people are working with an automated system that relegates them to passive decision makers, monitoring the process and providing input only when prompted by the automated aid, they tend to have less awareness of system changes. This occurs both when people interact with an automated system and when they work with another human. Since level 1 overview displays serve the function of prompting the operator when there are process deviations and other important notifications, i.e., turning them into passive decision makers, it is important to note the effect of automation on situation awareness due to its potential to lead to adverse plant conditions if operators lose awareness of system status and make uninformed or misinformed decisions.

Tharanathan et al. (2010) provided a definition more applicable to console operations in process control when they defined it as the understanding of what is happening in the plant, process, and equipment with respect to expected behaviors and performance both at the current point in time as well as in the future relative to the need for operations team action. From these definitions, we can derive 3 aspects of situation awareness that are of interest when evaluating operator performance with level 1 overview displays, and plant operations in general: (1) perception of relevant plant status – detection of process deviations, (2) comprehension of implications of that perceived status – relating process deviations to work goals and objectives, and (3) projection of status into the future – understand how the process deviations may influence the system in the near future. These map to Endsley’s (1988) definitions of the levels of situation awareness (perception, comprehension, projection).

It is important to take the concept of situation awareness into account while evaluating the effectiveness of various level 1 overview displays; this is because a strong ability to detect process deviations, tie those deviations into current objectives, and effectively predict their impact in the near future are imperative for maintaining a healthy system. In the context of the current study, perception and comprehension measures were collected, evaluating the relation with level 1 overview displays in a process-monitoring task. Measures will be described in detail in the method section.

Ecological Interface Design

The critical role of level 1 overview displays place great importance on the design of these displays and poses unique design challenges for display designers. Issues caused by imperfect automation, critical work environments, and situation awareness demands are complicated in nature and require a careful and planned design approach to help prevent or otherwise help mitigate the occurrence of such issues.

One such design method proposed is ecological interface design (EID). The EID framework was discussed in 1989 by Rasmussen and Vicente and is characterized as a theoretical framework for designing interfaces for complex sociotechnical systems (Rasmussen & Vicente, 1989). The main goal of EID is to support workers in adapting to change and novelty such as abnormal plant conditions, where designers may not have the ability to foresee a particular event and therefore are not able to provide support, such as standard procedures, to operators. This is accomplished by first understanding the work domain and its constraints and affordances through a detailed abstraction hierarchy (Rasmussen,1985) and providing users (operators) the necessary design features to encourage three types of behavior: (1) Skill-based behavior, which is automated and guided by sensed information about the environment, (2) Rule-based behavior, which is goal-oriented and guided by states and status of the environment that activate or modify predetermined actions or manipulations, and (3) Knowledge-based behavior, which is goal-controlled and guided by internal conceptual representation of information which is the basis for reasoning and planning (Rasmussen, 1983).

Further research on the topic provides support for improved performance through effective EID, especially when problem solving is required (Vicente, 2002). Assisting in situations where problem solving is required is important within the current context (process control) due to the frequent occurrence of such situations. The EID framework has been applied in other industries such as aviation and medicine and has led to the identification of new information requirements, which enhance user performance (Vicente, 2002). Vicente (2002) stated EID's main objectives as encouraging the use of skill- and rule-based behavior (this helps reduce cognitive resources necessary for a given task) and assisting in knowledge-based behavior (to help operators deal with unfamiliar situations, which typically requires more cognitive resources and is prone to errors).

A neonatal intensive care unit display was used to evaluate the impact of EID on display design in a medical setting. Sharp and Helmicki (1998) applied EID principles to the design of an interface by mapping informational requirements uncovered during a work domain analysis to their display. They compared the existing display with the EID display and found that all but one of their sixteen participants performed better with the existing display in a diagnosis task where they were instructed to perform diagnoses by choosing out of a set of recommended diagnosis options. This finding was in spite of the fact that all sixteen of their participants were already familiar with the existing display and had no prior experience with the EID display. This not only demonstrates EID's ability to yield better performance, but also its ability to provide a display that can quickly be learned and used efficiently.

In a process setting, Reising and Sanderson (2002) demonstrated successful implementation of EID on a display used to illustrate the pasteurization process. A work domain analysis was conducted to identify the properties and relations that should be represented in the interface. They noted significant improvement in fault diagnosis performance with the EID interface over the conventional interface. Of particular interest is their finding that some participants found certain displays within the interface difficult to use upon first use, but were able to become proficient and actually develop a preference with these displays as the sessions progressed. This is interesting because the EID display led to superior initial performance on displays that participants were familiar with and allowed participants to quickly acclimate to unfamiliar displays. This behavior was not noted for the non-EID interface.

Abstraction Hierarchy. Especially important to an EID approach is the implementation of a work domain analysis to assist in designers' ability to provide an interface, which appropriately represents the work domain. The abstraction hierarchy can be used to design and implement interfaces that accomplish this goal of appropriately representing a particular work domain (Vicente, 2002).

The abstraction hierarchy can be viewed as a hierarchical representation of a work domain, highlighting various levels and their relationships with one another. Rasmussen (1985) outlined 5 levels of representation relevant to process control: (1) Functional purpose (why was the plant designed?), (2) Abstract function (what is the mass and energy topology of the plant?), (3) Generalized function (what generic functions implement that topology?), (4) Physical function (what plant equipment realize those functions?), and (5) Physical form (what is the spatial location and appearance of the equipment?). Figure 2 provides a graphical representation of Rasmussen's 5 levels of representation relevant to process control.

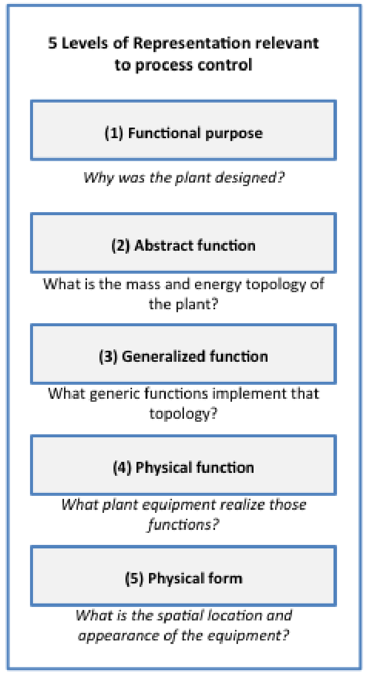

Figure 2. Five levels of representation relevant to process control as a function of the abstraction hierarchy.

In an abstraction hierarchy, higher levels typically represent functional information (e.g., functions of pumps, heaters, valves, etc) and lower levels typically represent physical information (e.g., state of pumps, heaters, valves, etc.). Figure 3 illustrates the relationship between Vicente's 5 levels of representation and the display hierarchy provided by Tharanathan et al. (2010) for a typical process plant.

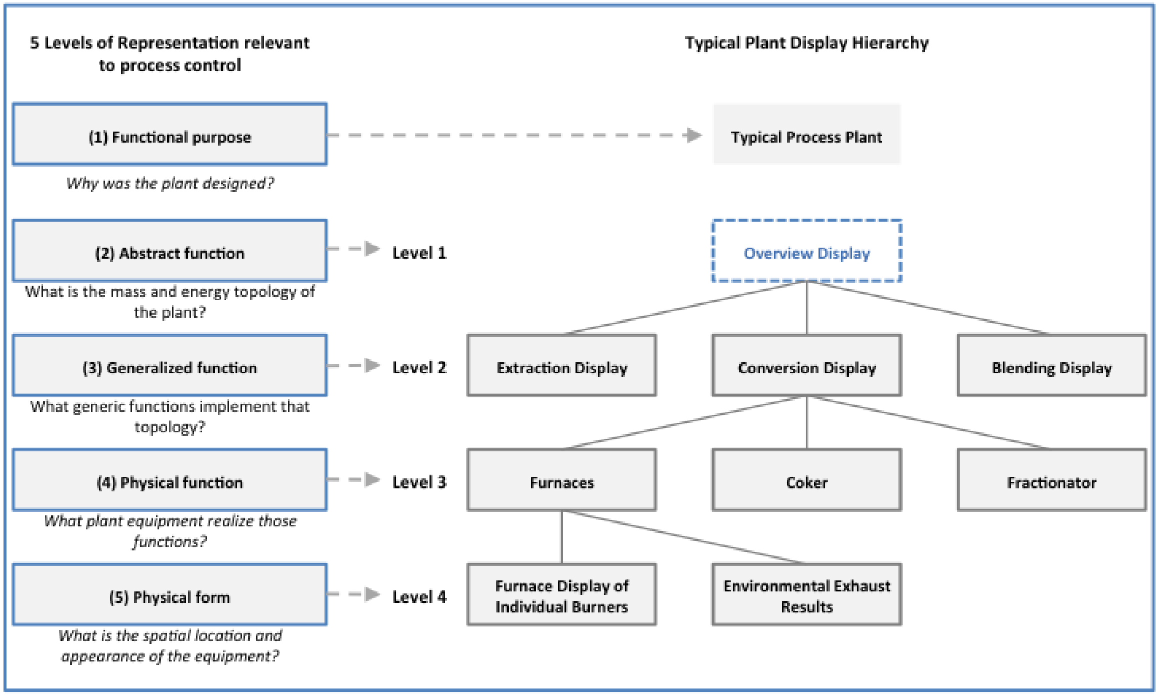

Figure 3. Relationship between the 5 levels of representation and the display hierarchy for a typical process plant.

It is important to note the difference between an abstraction hierarchy (used to effectively support EID) and task analysis. Vicente (1999) defines a task as a set of actions to be performed by one or more actors in pursuit of a particular goal and a work domain as the system that is being controlled by the actors. The key distinction is that a task is something that is done by the operator and a work domain is something that is acted on by the operator (Vicente, 2002). The abstraction hierarchy is intended to model and represent the work domain, not necessarily the tasks to be completed within the work domain. This is accomplished by generating information requirements that inform the design of a particular interface (Jamieson, Hajdukiewicz, & Reising, 2002). It should be noted though, that task analysis is an equally important tool in the development of effective displays as a work domain analysis; however, the current review focuses on the work domain analysis and its relationship with EID and the SRK Taxonomy (see below). In fact, when taking an EID approach to display design, the first step is to identify purposes, laws and constraints, and component resources that are governing processes. An abstraction hierarchy provides this information. This information can then be used in the design process and applied within the SRK taxonomy (Jamieson et al., 2002).

Skill-Rule-Knowledge Taxonomy. Rasmussen (1983) defined 3 modes of control (or behavior) that operators use to act on their environments within the SRK taxonomy: (1) Skill-based behavior, which is automated and guided by sensed information about the environment, (2) Rule-based behavior, which is goal-oriented and guided by states and status of the environment that activate or modify predetermined actions or manipulations, and (3) Knowledge-based behavior, which is goal-controlled and guided by internal conceptual representation of information which is the basis for reasoning and planning.

In the context of the SRK Taxonomy, Vicente (2002) defined the 3 modes: (1) Skill-based behavior is assisted by workers ability to act directly on the interface (direct manipulation), (2) Rule-based behavior is assisted by providing workers with a consistent one-to-one mapping between the work domain constraints and perceptual information in the interface (map perceptual forms), and (3) Knowledge-based behavior is assisted by the interface representing the work domain in the form of an abstraction hierarchy to serve as externalized mental model for problem solving.

Jamieson et al. (2002) commented on the focus of EID and the role it plays in assisting skilled operators in dealing with abnormal plant conditions. Three contributing factors of EID that help operators are ‘showing’ operators the affordances and constraints for action in a process environment, allowing skilled operators to become familiar with plant operations and are able to effectively act on process deviations by selecting action alternatives in unique situations, and allowing operators to manage opportunities in both familiar and unfamiliar plant situations. These three factors map onto the SRK taxonomy developed by Rasmussen (1983) and later reiterated by Vicente (2002).

Skill-based behavior is highly automated and guided by sensed information about the environment and is encouraged by ‘showing’ operators the affordances and constraints for action in a process environment. Providing operators with the affordances and constraints of the plant is achieved by allowing them to act directly on the interface (e.g., show affordances by providing direct manipulation through interface controls). In the context of process control, this allows operators to quickly and accurately react to process deviations.

Rule-based behavior is goal-oriented and guided by states and status of the environment that activate or modify predetermined actions or manipulations and is encouraged by allowing skilled operators to become familiar with plant operations and be able to effectively act on process deviations by selecting action alternatives in unique situations. This is achieved by providing operators with a consistent one-to-one mapping between the work domain constraints and perceptual information in the interface (e.g., help operators become familiar with the plant and effectively select action alternatives by mapping perceptual forms between the plant and the interface). This allows operators to spend less cognitive resources processing visual information about the plant.

Knowledge-based behavior is goal-controlled and guided by internal conceptual representation of information which is the basis for reasoning and planning and is encouraged by allowing operators to manage opportunities in both familiar and unfamiliar plant situations. Allowing operators to manage both familiar and unfamiliar situations is achieved by representing the work domain in the form of an abstraction hierarchy to serve as externalized mental model for problem solving (e.g., help operators manage familiar and unfamiliar opportunities by providing an appropriate functional plant layout) (Woods, 2009; Vicente, 2002; Jamieson et al., 2002). This allows operators to maintain a clear mental model of the plant and its current state. It is important to note that while a particular measure might be closely related to a particular aspect of the SRK Taxonomy (e.g, operator detection accuracy and skill-based behavior), they might also be linked to other aspects. For example, rule-based and knowledge-based behavior could support operator workload by allowing operators to spend less resources processing visual information about the plant and allowing operators to maintain a clear mental model of the plant. Additionally, operator detection accuracy can be aided by knowledge-based behavior in instances when it helps operators form appropriate expectations about the plant conditions. For the purposes of the current analysis, the current review focuses on a general mapping framework to indicate these relationships. Additional work should be done to further investigate these relationships. Figure 4 provides a graphical illustration of the SRK Taxonomy and how each behavior affects operator performance.

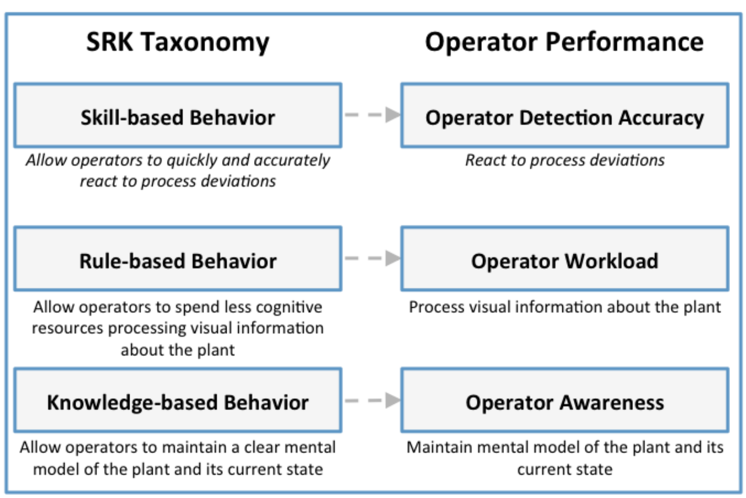

Figure 4. Mapping of SRK Taxonomy to various types of operator performance.

To summarize, interfaces that support skill-based behavior show affordances by providing direct manipulation through interface controls, leading to enhanced operator detection performance. Those that support rule-based behavior help operators become familiar with the plant and effectively select action alternatives by mapping perceptual forms between the plant and the interface, reducing operator workload. And those that support knowledge-based behavior help operators manage familiar and unfamiliar opportunities by providing an appropriate functional plant layout, increasing operator plant awareness.

In process control, responding quickly and accurately to plant deviations, understanding the state of the system, and operator workload levels are important indicators of human-automation performance and therefore system performance. Success in these areas allows operators to efficiently and effectively detect and respond to plant deviations (both anticipated and unanticipated).

In the context of the current line of research, the Abstraction Hierarchy and SRK Taxonomy have been used to evaluate the work domain of a console operator and provide design recommendations for providing effective level 1 overview displays (Bullemer, Reising, Burns, Hajdukiewicz, & Andrzejewski, 2008; Effken, Kim, & Shaw, 1997; Jamieson et al., 2002). Using these resources, display guidelines have been developed and followed, along with industry expertise to develop the displays used in the current research. Findings and recommendations based on this work were applied during the development of the level 1 overview displays evaluated in the current study. Specifically, work domain analysis (abstraction hierarchy) and the SRK taxonomy were used as a basis for design. These are the two key pieces for successful EID implementation (Vicente, 2002). These displays were developed by industry leaders with the intention of being used in real control rooms. Given this, the implementation of the EID varied, usually due to time, cost, and development constraints.

There were 3 level 1 overview displays evaluated in the current study: Calm Water, Heat Map, and Visual Thesaurus. Work domain analysis informed the design of the 3 displays equally. The workplace representation and its constraints were kept consistent across the 3 displays. That is, each display was informed by the same work domain analysis and therefore generally contained the same information. The Heat Map display, however, contained historical process information but the other two displays did not contain this information (more on this below). This information includes the 5 levels of representation discussed by Rasmussen (1985) (functional purpose, abstract function, generalized function, physical function, and physical form) and ensured appropriate representation of everything from plant layout to component states to specific information about the components contained in the plant. The approach of representing the workplace equally (i.e., applying workplace analysis equally) for each display was done so that other factors could be evaluated; specifically, the implementation of the SRK taxonomy and the display's ability to assist in operators being able to leverage skill-, rule-, and knowledge based behavior.

Experimental Level 1 Overview Display Designs

In the current study, the research team worked with 3 separate design teams. These design teams worked to develop 3 distinct level 1 overview displays: Calm Water, Heat Map, and Visual Thesaurus. The design goal of these displays was to provide control room operators with the tools necessary to support in detecting, diagnosing, and acting on process deviations as prescribed in the EID literature above. All 3 design teams developed displays using the outcome of the same abstraction hierarchy, which was previously conducted for another project (Jamieson et al., 2002).

While the design teams based their designs from a common abstraction hierarchy and plant representation, the actual implementation of the displays varied based on multiple factors. In all 3 cases, there were previously established displays that were purposed for the current study by taking existing code and style guides and applying them to the abstraction hierarchy to standardize around a single process. The Calm Water and Heat Map displays had previously been developed to fill specific needs for their respective team's facilities. They did not necessarily serve as level 1 overview displays. The Visual Thesaurus display had been developed specifically for the purpose of serving as a level 1 overview display and was able to be applied to the current context with minimal effort. Another, important, consideration is that each team had independent designers who developed their displays. These factors led to less-than-ideal continuity in the final design of the three displays and three very distinct approaches to level 1 overview displays.

A downside to this is that there was no standardization around some of the key graphical elements such as buttons, colors, or shapes. Each team used conventions that they had previously used as part of their company's user interface style guides. However, terminology and representation of process elements were made consistent between the displays. A more serious downside is that the lack of design continuity is that it makes it difficult to draw the lines on precisely how well EID has been implemented in each display, making direct comparisons difficult. An upside to having distinct approaches to design is that the designs are all fairly different in their appearance and theoretical functionality, proving a range of display options for consideration in future experimentation and implementation.

It is important to note that the research team did not have any influence on the design of the displays. Rather, researchers were engaged to compare the 3 displays in a real world process control environment. An experimental study was developed and conducted to compare the displays.

The main process areas in the level 1 overview displays used in the current study consisted of: crude feed, crude heater, crude tower, vacuum heater, and vacuum tower units. These functional areas (units) were determined based on a combination of previous research (Tharanathan et al., 2010) and the various displays' capabilities to maintain functionally equivalent plant overviews for all experimental conditions. The experimental displays that were used in the study are described below. All 5 functional areas (crude feed, crude heater, crude tower, vacuum heater, and vacuum tower) were represented in the experimental displays. These functional areas were grouped in the displays in reference to their relative position in the overall process. The display used a series of gauges to represent components in each of the functional areas contained on the display. Gauges in the experimental level 1 overview displays were either process values (PV) or operator output percentages (OP).

Each gauge had a button for input associated with it. These buttons were used for operators to acknowledge changes in gauge state (by clicking on the button). Additionally, all of the buttons used in the experimental overview displays followed the same naming convention. There were 3 parts to each button label. The first part indicated what type of gauge it is. Each gauge can indicate flow, temperature, level, or pressure levels and is represented by using the first letter of each type (e.g., "F", "T", "L", "P"). The second part indicated the gauge name (e.g., "Total Pass Flow"), which lets the user know what component the gauge represents. The third part indicated whether the gauge is a process value or operator output percentage gauge (e.g., "PV", "OP").

Additionally, gauge states were defined by the amount of deviation a particular gauge was from normal operating range. A gauge could be in 1 of 3 states: normal, abnormal, and alarm. Normal range was considered to be within normal operating range for a particular gauge, abnormal range was considered to be outside of normal range (high or low), but not yet in the alarm range, and alarm range was considered to be outside of both normal and abnormal ranges (high or low). In the current study, the terms "high" and "low" referred to "above" or "below" normal operating range, respectively. This clarification was made to remove confusion about the terminology (e.g., "high" and "low") being used to indicate the severity of an alarm. Refer to figure 5 for a graphical representation of the possible gauge state definitions. Alarm and abnormal states in the simulations were developed using process expert feedback in preparation for previous research (Tharanathan et al., 2010) and were reused in the current study. Note that this representation is only notional, and not an actual representation of any design element used in the study.

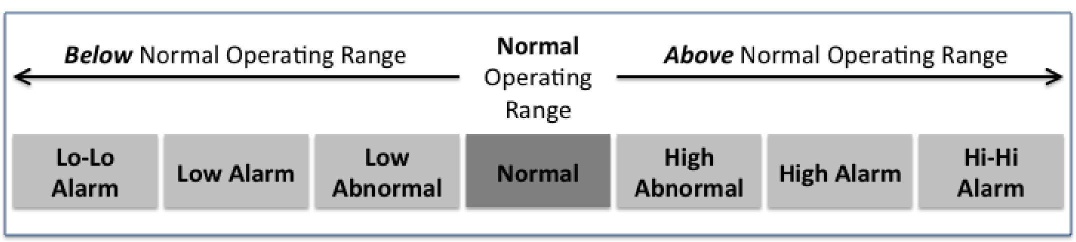

Figure 5. Illustration of possible gauge state definitions used in the current study. The states shown here indicate the possible states of each of the level 1 overview displays used in the current study.

The Calm Water display. The Calm Water display uses a matrix technique to display gauges and process deviations in the plant. Each gauge is represented as a node on this matrix. The matrix is then turned on its side in 3-dimensional space and presented to operators from a side view (figure 6).

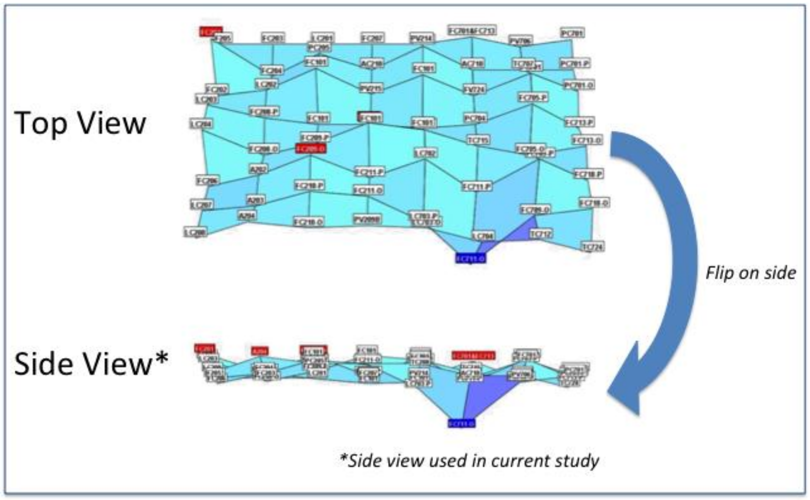

Figure 6. Illustration of Calm Water display matrix perspectives. The side view was used in the current study.

In this side view, operators view plant status as a line across the screen. Due to being flipped, gauges are occluded in the display. Therefore, process deviations are represented as peaks and valleys in the straight line and the size of the peak or valley correlates with the degree of deviation from normal (straight line). Only deviations are visible due to their deviation from the side-view straight line. Generally, product flows from left to right in this display. The term “Calm Water” is used to represent this matrix visually, which will appear like a calm body of water when in a normal state. When the display enters abnormal or alarm states, the matrix will appear like water that has been disturbed (e.g., containing waves). Refer to figure 7 for a graphical representation of the Calm Water overview display.

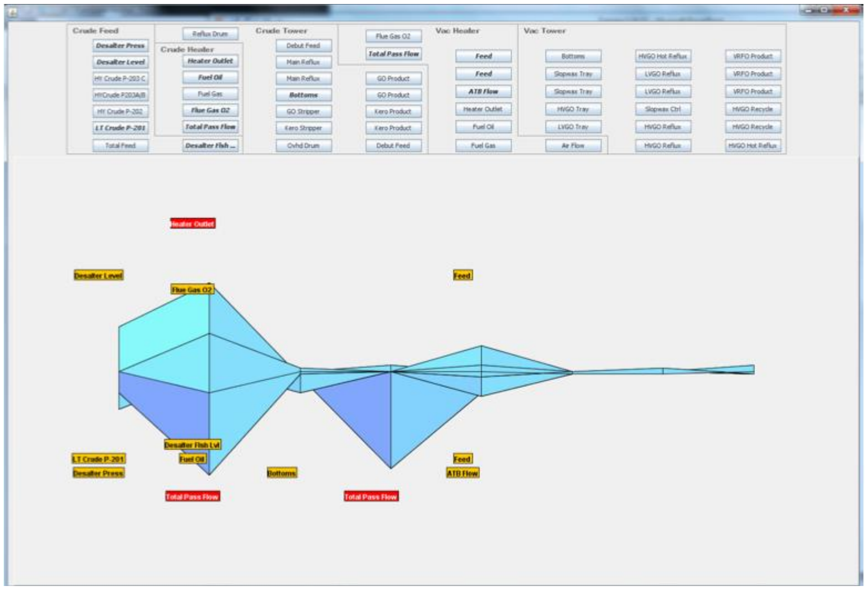

Figure 7. Illustration of the Calm Water overview display

Each gauge is represented as a node in the matrix and the value is provided in the form of a peak or valley in the matrix line. A peak indicates a deviation above normal operating range and a valley indicates a deviation below normal operating range. In addition to the direction of the deviation, the severity is also provided graphically via the size of the peak or valley. In other words, the distance each peak and valley is away from the straight line is a graphical indication of the amount of deviation that a particular gauge has experienced. In addition to the direction and graphical indication of deviation, colored tags are used to indicate whether deviations are in the abnormal or alarm range. Each tag displays a color and gauge name. Yellow tags are used to indicate deviations in the abnormal range and red tags are used to indicate deviations in the alarm range. It is important to note that these tags can indicate either high or low process deviations, depending on the direction of the deviations (e.g., peaks vs. valleys in the line). Additionally, areas with multiple deviations are stacked on top of one another so that individual gauges do not become occluded by other gauges in the same area. This was done so that all abnormal and alarm states are possible to detect by operators. The tag color only indicates the severity of each deviation, not necessarily the direction. There are no tags to indicate normal range, meaning that the inverted matrix does not display tags for gauges in the normal range. This was a design decision made to reduce clutter and only present tags for gauges that have deviated from normal range and require attention. Refer to figure 8 for a graphical depiction of these tags.

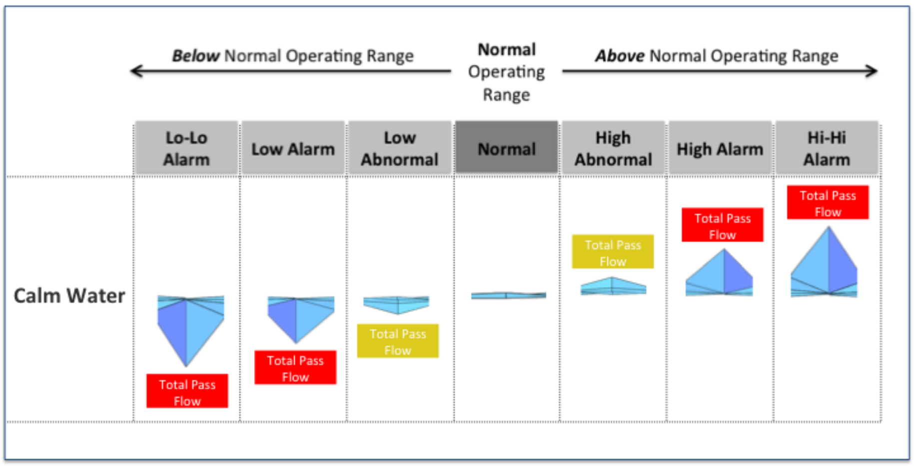

Figure 8. Illustration of tags used in the Calm Water overview display. Color of the tag and vertical distance from the center indicates the amount of deviation from normal operation. There is no tag present to represent normal operating conditions.

In this display, the button for each gauge is always located directly above the associated gauge node (i.e. tag). All buttons are presented above the inverted (side view) matrix display and organized in relation to gauges and components in the physical plant. As per recommendation from the design team, the text in the buttons for input become boldface font when the associated tag enters either the abnormal or alarm state (either high or low) to assist in identification of the relevant button to acknowledge changes in gauge states.

The Heat Map display. The Heat Map display uses gauges with a historical heat map of process values. Colors are used to indicate whether deviations are high (above normal operating range) or low (below normal operating range) and shades to indicate the amount of deviation for each gauge (e.g., lighter shades of color indicate lesser deviations from normal and darker shades of color indicate greater deviations from normal). Shades of red indicate deviations above normal operating range and blue colors indicate deviations below normal operating range. Refer to figure 9 for a graphical representation of the Heat Map overview display.

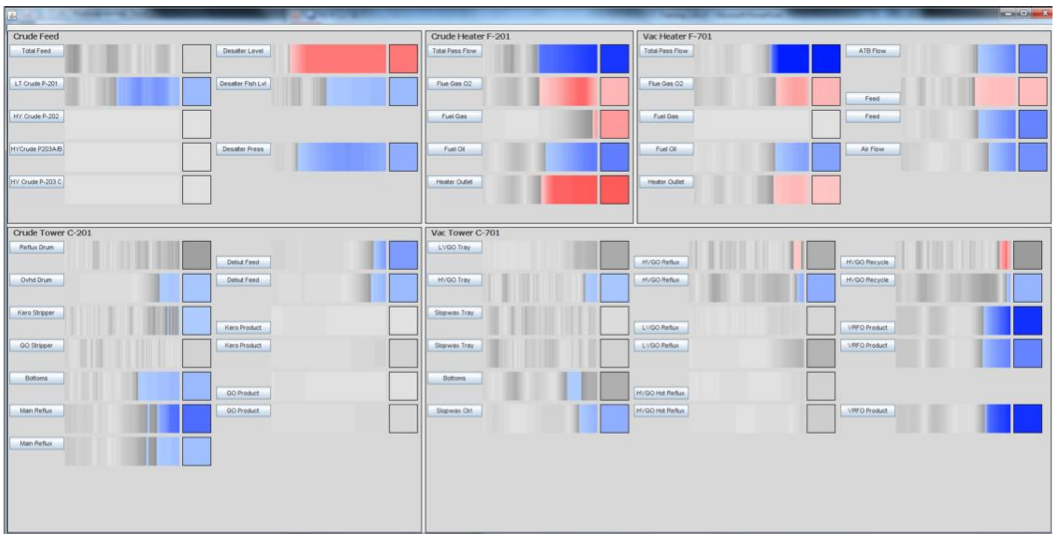

Figure 9. Illustration of Heat Map overview display

As mentioned above, each gauge contains a historical heat map of process values for that particular gauge. This historical heat map represents the last 6 minutes of gauge state data in the form of colors and shades. The historical heat map updates, from right to left, once every 2 seconds. Upon each update, the color of the current state will move to the left into the historical portion of the gauge. This update takes place linearly, meaning that the size of the historical state, for any given time, remains the same throughout the entire 6-minute time period. Each gauge also contains a box indicating the current state of the gauge (color and shade) and a button/ label for that particular component (e.g., “Total Pass Flow”). Refer to figure 10 for a graphical representation of present and historical indicators in the Heat Map display.

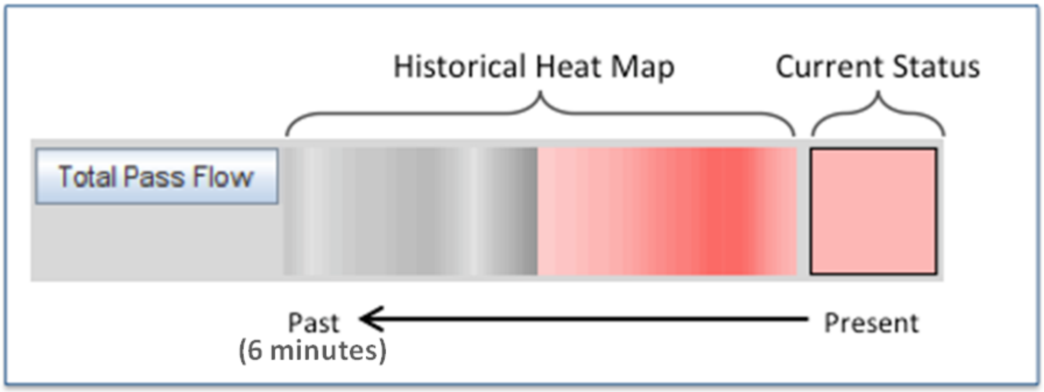

Figure 10. Illustration of Heat Map present and historical indicators.

Shades of color are used to indicate the severity of each deviation. For example, light shades of red or blue indicate lesser deviation from normal operating range and are used to convey abnormal range. Conversely, dark shades of red or blue indicate greater deviations from normal operating range and are used to convey alarm range. Shades of gray are used to indicate normal operating range, with darker shades indicating gauge states approaching the abnormal range (note that the shade of gray does not indicate whether the deviation is high or low, only the severity of the deviation from normal operating range). Refer to figure 11 for a graphical depiction of gauge colors and definitions.

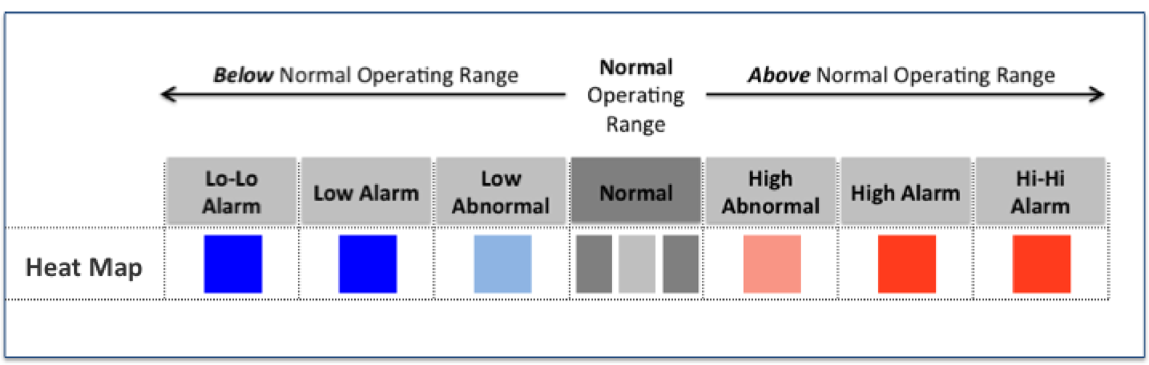

Figure 11. Illustration of Heat Map gauge colors and definitions.

In this display, the button for each gauge is always with the associated gauge; this design decision was made to provide collocation between the gauge and its associated button. As noted above, associating the buttons with the gauge was a different strategy compared to the Calm Water display, where designers placed buttons at the top of the display and relied on visual distinction to allow operators to locate them.

The Visual Thesaurus Display. The Visual Thesaurus display is a digital recreation of traditional analog-style gauges. As such, it uses shapes to represent gauges (e.g., level, temperature, flow, pressure), graphical indicators to represent values, and colors to indicate deviations from normal operating range (e.g., low alarm, high alarm, hi-hi/ lo-lo alarm, abnormal). The icons are selected from a larger set of visual elements (a “visual thesaurus”) that was created to visualize process-related elements for use in overview displays. Refer to figure 12 for a graphical representation of the Visual Thesaurus overview display.

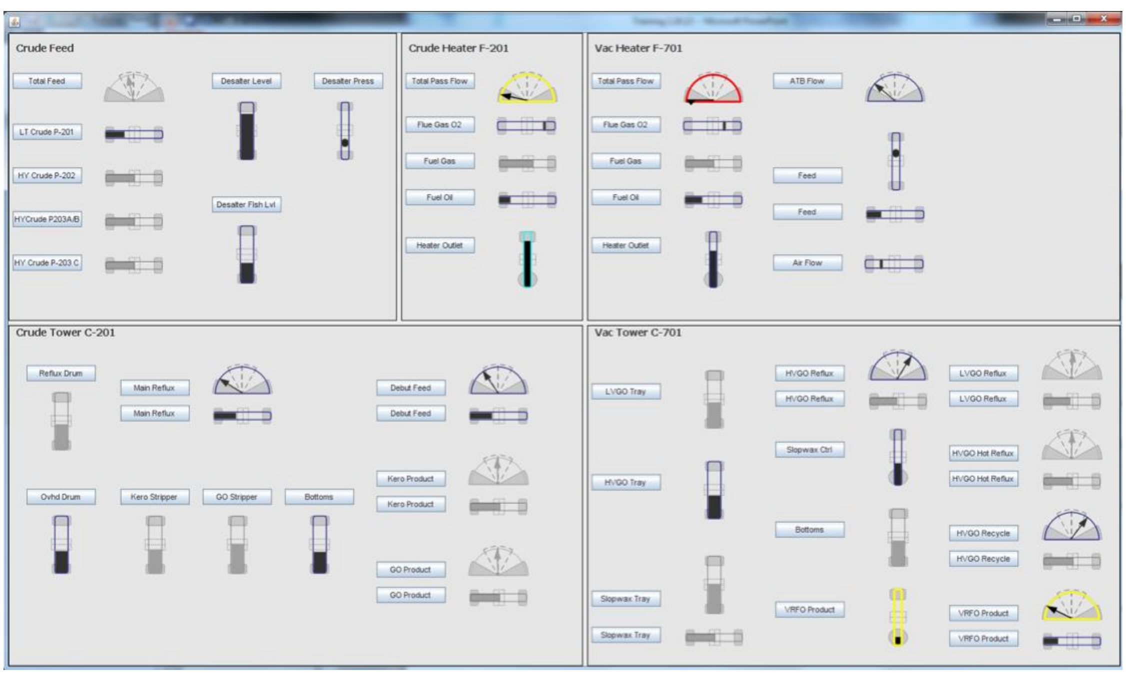

Figure 12. Illustration of Visual Thesaurus overview display

Each gauge uses a colored outline to alert operators of process deviations above or below normal operating conditions in the abnormal and alarm ranges. Refer to figure 13 for a graphical representation of the gauge shapes used in the Visual Thesaurus display.

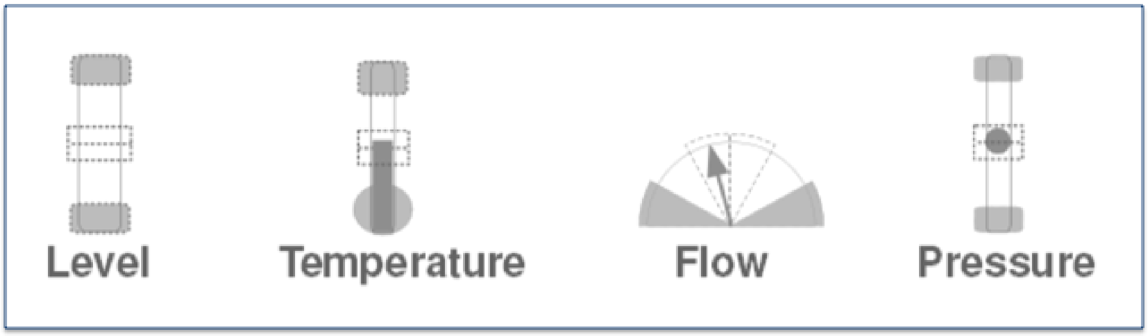

Figure 13. Illustration of gauge shapes used in the Visual Thesaurus overview display.

Yellow outlines are used to indicate deviations in the low alarm range, blue outlines are used to indicate deviations in the high alarm range, red outlines are used to indicate both extreme high and extreme low alarms (Hi-Hi and Lo-Lo), and purple outlines are used to indicate abnormal range. Refer to figure 14 for a graphical depiction of outline colors used in the Visual Thesaurus display.

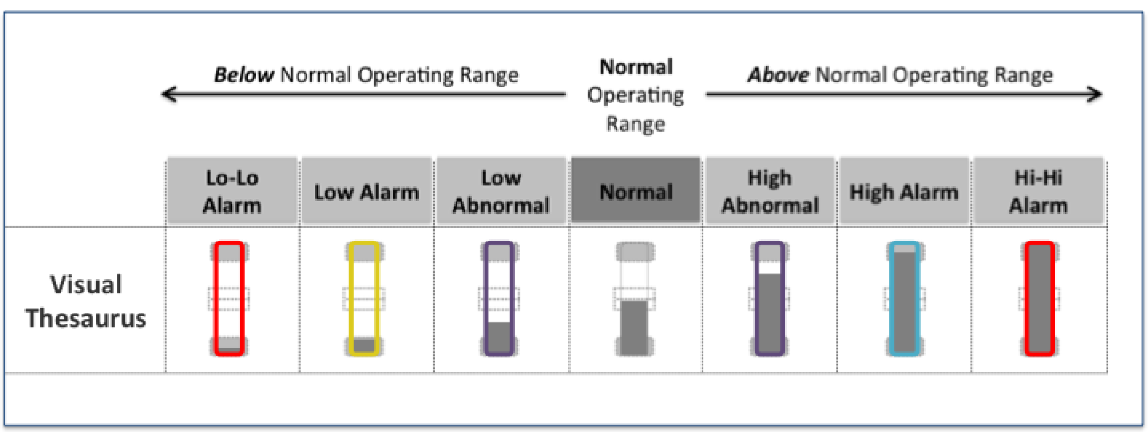

Figure 14. Illustration of outline colors used in Visual Thesaurus overview display. Note that the gauge shape used in this example is the level gauge.

In addition to the colored outlines, gauge indicators are used to indicate current operating conditions for each display. The indicators provide information about the gauge's current status in relation to normal operating range. That is, the indicator provides a graphical depiction of the direction of the deviation (e.g., high, low, etc.). In this display, the button for each gauge is always with the associated gauge; this design decision was made to provide collocation between the gauge and its associated button.

Refer to figure 15 for a graphical summary of the gauge state indicators used in all experimental level 1 overview displays in the current study.

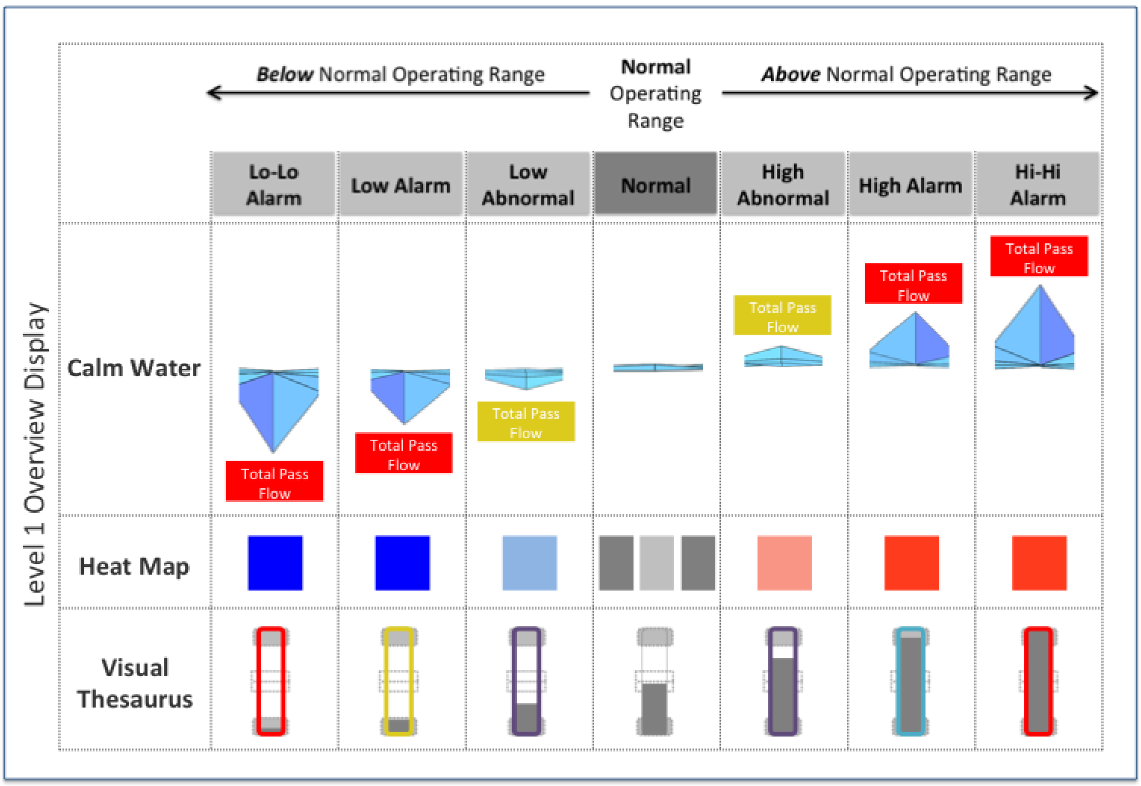

Figure 15. Summary of gauge state indicators for the level 1 overview displays used in the current study.

SRK Taxonomy Implementation

The level at which the SRK Taxonomy was applied in the experimental displays varied between the 3 displays used in the current study due to different constraints on each individual design process; including, time, budget, designer expertise, and intended application. The displays were designed to assist operators in detecting process deviations by alerting operators of deviations and providing contextual information about these deviations. This was the general task that operators conducted during the study. Researchers reviewed the displays in light of the three behaviors (skill, rule, knowledge) below and evaluated how each display is expected to affect the use of each type of behavior.

Skill-based behavior – Show affordances by providing direct manipulation through interface controls. The Calm Water display promotes direct manipulation of the interface controls by consistently presenting buttons directly above related gauges on the display (instead of next to gauges), providing consistent labeling between gauges and buttons, and changing the style of button text to bold when a gauge has deviated into the abnormal or alarm level.

The Heat map and Visual thesaurus displays accomplish this goal by always collocating gauge buttons with the gauges themselves, the only difference between the 2 are that buttons are always directly to the left of the gauge in the Heat Map display and buttons are either to the left or directly above the gauge in the Visual Thesaurus display. Unlike the Calm Water display, the Heat Map and Visual thesaurus display button text does not change to bold style when the associated gauge has deviated from normal. The reasoning for this difference is because the Calm Water display does not collocate buttons with the gauge and therefore takes advantage of a text style change to make the appropriate button more salient and presumably easier to locate. All three displays take advantage of a blue highlight on the buttons to indicate that buttons are “clickable” objects in the interface when they pointed to using the computer mouse.

Rule-based behavior – Help operators become familiar with the plant and effectively select action alternatives by mapping perceptual forms between the plant and the interface. The Calm Water display maps perceptual forms between the plant and interface by using a straight line to indicate current process status, with peaks and valleys in the line indicating process deviations. Labels and button text are used to indicate gauge types. Graphical peaks and valleys on the display (above or below the straight line) are used to indicate gauge levels (peaks = high, valleys = low) and the amount of deviation is indicated by the size of the peaks and valleys. Each gauge is labeled with a tag when it deviates into the abnormal or alarm states. Color tags are used to indicate the current state of a gauge (yellow = abnormal, red = alarm).

The Heat Map display uses both current and historical gauge status with the use of colors (gray, blue, red) and varying shades of these colors (light-to-dark). Button text indicates gauge types. Color (gray, red, and blue) is used to indicate the direction of a process deviation (normal, high, and low). The particular shade of each of these 3 colors indicates the size of the deviation and the current state (darker shades = more deviation and closer to the alarm range, lighter shades = less deviation and closer to the abnormal range). It should be noted that, due to the continuous nature of the shaded indicators, there is no discrete point at which a gauge indicates a particular state. For example, it is not graphically possible to precisely determine when a gauge has moved between abnormal and alarm states. Rather, operators are given the freedom to indicate when the shade of the gauge color has deviated enough so that they indicate a state change.

The Visual Thesaurus display uses graphical objects to represent gauges, with unique gauge shapes to indicate different gauge types. Button text is also used to indicate gauge type. Level indicators on the gauge shapes are used to indicate current gauge levels. Color outlines are used to indicate the direction and state of process deviations (blue = high alarm, yellow = low alarm, red = high or low – hi-hi/ lo-lo alarm). Each gauge shape contains ranges (graphically indicated) to indicate whether the gauge is in normal, abnormal, or alarm state. The level indicators are used to provide this indication.

Knowledge-based behavior – Help operators manage familiar and unfamiliar opportunities by providing an appropriate functional plant layout.

The Calm Water display represents functional plant layout by providing a left-to-right representation of product flow (propagation of product through plant) straight across the screen. Flow starts at the left of the display and moves to the right as product flows through the plant. Gauges are grouped in a matrix-like layout on the display, grouped into functional areas (5 functional areas). The buttons are grouped in a similar manner. Both help reinforce the left-to-right nature of the functional plant layout and allow the interface to show flow of product straight across the display.

The Heat Map and Visual Thesaurus displays provide functional grouping of plant layout, but do not follow the left-to-right convention as strictly as the Calm Water display. The 5 functional areas represented by the display are graphically separated by outlines in the display, which provides a clear indication of functional areas in the plant. Product in these functional areas generally flows from the left of the screen to the right of the screen. However, the graphical representation of product flow is dictated by these 5 areas' location on the display (product flows from one functional area to another). For example, the flow direction of the Heat Map display begins at the upper-left portion of the display and moves one functional area to the right, then down and to the left, then up and to the right, then down. Similarly, the Visual Thesaurus display indicates flow of product beginning at the upper-left functional area and moves to the functional area to the right, then down one, then up and to the right, then down. This representation is dependent on the design of the interface and ability to provide grouped functional areas.

Study Context

Researchers examined the effects of level 1 overview display design – specifically; EID and the use SRK taxonomy on operator performance in a simulated process control-monitoring task. While the abstraction hierarchy was held constant for the 3 experimental overview displays, the overall level of EID that was implemented in each display varied based on design constraints that caused the elements of SRK implementation to vary. These design differences have been enumerated in the text above for each experimental level 1 overview display (Calm Water, Heat Map, and Visual Thesaurus). The level 1 overview displays made up the primary task for operators in the study. A secondary task was used to simulate the varying levels of workload faced by control room operators. The type of secondary task used was a monitoring and control task, where operators were instructed to monitor and make adjustments, as necessary, to maintain a healthy system.

Operator detection performance was measured by collecting operator detection accuracy and response time metrics. Operator workload was measured by collecting subjective workload and secondary task performance metrics. Operator awareness was measured by administering situation awareness probes. Figure 16 illustrates the relationship between the SRK Taxonomy, operator performance, and the measures used in the current study.

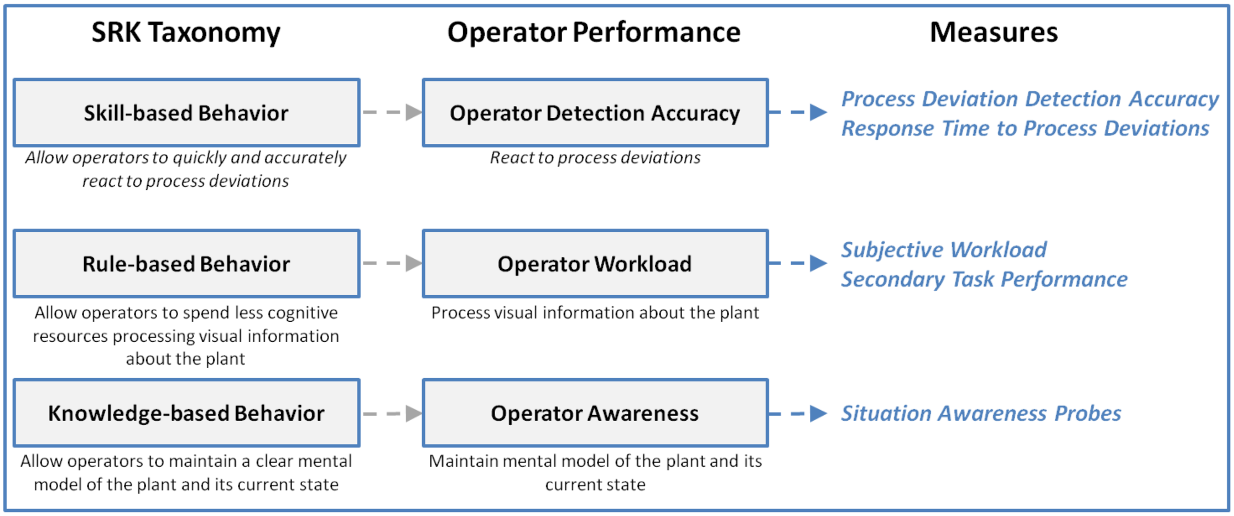

Figure 16. Mapping of SRK Taxonomy, operator performance, and the measures used in the current study.

Given what we know about the SRK Taxonomy and the measures discussed above, we can conclude that a design supporting skill-based behavior will also support operator detection performance by showing interface affordances through direct manipulation – allowing operators to quickly and accurately react to process deviations. Additionally, a design supporting rule-based behavior will also support operator workload by helping operators become familiar with the interface by clearly mapping the interface with the plant layout – allowing operators to spend less cognitive resources processing visual information about the plant. Finally, a design supporting knowledge-based behavior will also support operator awareness by providing a clear functional layout of the plant to the operator – allowing operators to maintain a clear mental model of the plant and its current state.

Hypotheses

Overall, it should be noted that the 3 displays evaluated in the study were very complex in nature and did not share common design themes like color usage, button placement, or gauge representation. This was due to the complex nature of their functional purpose and the fact that, aside from sharing a common work domain analysis, they were developed independently. Limited budget and timelines reduced the already minimal scope to make enhancements that would sync the displays up in terms of design direction, which would have provided more comparable experimental displays. This means that it is nearly impossible to identify precisely what factors lead to performance differences between the designs. Given this, the current study focuses on evaluating each display as a whole in terms of operator detection performance, operator workload, and operator awareness. Future research can be conducted to compare more specific attributes and provide a more precise comparison between the displays.

Operator detection accuracy. We expect operators in the Calm Water experimental display condition to have better response time and response accuracy scores, leading to better operator detection performance. This hypothesis is driven by this display’s qualitatively different method of highlighting buttons once a process deviation has occurred in the simulation – they become bold in addition to the blue highlight when the user hovers over them with the mouse. This creates visual distinction and it is expected to aid in operators’ ability to locate the appropriate button to indicate process deviations. Additionally, while the buttons in the Calm Water display are not directly collocated with the relevant gauges like the Heat Map and Visual Thesaurus displays, they are always directly above the gauges, collocated with the remaining display buttons. Again, this is expected to aid operators’ performance due to creating visual distinction between the actionable button and the rest of the display elements, while also providing a consistent location for buttons.

Operator workload. We expect operators in the Calm Water experimental display condition to have lower subjective workload scores, leading to better secondary task performance scores. While the Visual Thesaurus display qualitatively does a better job at mapping perceptual forms of display gauges (by using literal representations of gauges) compared to the other displays, the Calm Water display provides more salient indications of gauge level and gauge state indicators. This is achieved by using highly salient visual peaks and valleys (deviating from the straight line of the display) to indicate the gauge states and clearly using color-coded tags to indicate gauge state. While the state indication methods of the other two displays are expected to be effective, the location of the tags in the Calm Water display is expected to help provide further distinction from the rest of the display and aid in operators' ability to easily notice, and evaluate the state of the indicator.

Operator awareness. We expect operators in the Calm Water condition to have better situation awareness level 2 probe scores, leading to better operator awareness. While neither display does a particularly effective job of grouping related gauges (all three displays provide relevant but somewhat arbitrary groupings of gauges due to space constraints on the display), the Calm Water's method of laying out gauges in a matrix-like structure helps to provide operators with a visual representation of the plant layout. This is possible on the Calm Water display because the gauges and buttons are not collocated like in the other two displays, providing means for a different layout scheme. Additionally, this layout allows the Calm Water display to clearly indicate direction of flow of product through the display (left to right) by representing gauges along a straight line (via the matrix structure). This is preferable compared to the other two displays because they require operators to maintain a mental model of how product is flowing through the display. Neither display represents flow of product straight across the display due to the dependencies on the gauge grouping constraints mentioned above.

Chapter 2: Method

Research Design

A 3 x 2 x 2 mixed design was used to evaluate the type of display (Calm Water vs. Heat Map vs. Visual Thesaurus), scenario complexity (low complexity vs. high complexity), and experimental trial (1st trial vs 2nd trial). The experimental trial was taken into consideration to identify any potential learning effects over the course of the study. Operator workload was implicitly manipulated by varying the complexity of the scenarios. Each display was developed by an Abnormal Situation Management Consortium member's design team and submitted for evaluation in the current study. The displays each use a unique method to orient the operator to the existence, severity, location, and direction of plant deviations. Operators were assigned to one of the 3 display types, making display type a between-subjects variable. All operators were exposed to both low- and high-complexity scenarios, making this a within-subjects variable. Additionally, all operators completed 2 experimental trials, making this a within-subjects variable. Operator performance measures were collected using multiple measurement methods.

Situation Awareness Level 1 performance data was measured by recording operators' responses to process deviations throughout each scenario – perception of process deviations (Tharanathan et al., 2010). These measures included operator detection accuracy and response time to process deviations and provided an indication of operator detection performance.

Subjective workload measures were collected using the NASA TLX subjective workload probe (Hart & Staveland, 1989) following each experimental session. These measures provided a broad view of operator performance while interacting with the 3 displays and 2 complexity levels. Additionally, secondary task performance data was collected while participants performed resource management and system monitoring and controls on the Multi-Attribute Task Battery II (MAT-B II) (Comstock & Arnegard, 1992). These measures provided an indication of operator workload.

Situation Awareness Level 2 performance was collected by administering Situation Awareness probes at random intervals during each experimental session, measuring operators' ability to relate process deviations to work goals and objectives (comprehension of implications of perceived status, Tharanathan et al. 2010). The probe was presented to operators 2 times per scenario. These measures provided an indication of operator awareness.

Participants

Forty-three male and five female operators participated in the study. These operators worked for Sasol energy and chemical company (Abnormal Situation Management Member Company) and interacted with a display similar to the schematic display (explained in detail in the text below) used in the current study; however, the version they used was a different version of this display. They were recruited based on their availability at the discretion of the shift supervisor. This Sasol facility provides refinement of crude oil and coal across multiple sub-plants on the site location in South Africa. All operators were working their normal control room shift schedule and another operator covered their console while they participated in the simulator evaluation. The average process domain experience for the operators that participated in the study was 13.6 years. Each participant was randomly assigned to one of the three experimental conditions. Participants reported having normal or corrected-to-normal vision. Note that 2 operators reported having color blindness. These operators underwent additional pre-screening to determine whether they were able to detect the colors that were used in the displays used in the study. During screening, participants were asked to indicate the colors of each element of the experimental display they had been assigned, as well as colors from each of the other two displays they would be interacting with. They both passed the pre-screening and were deemed able to participate by researchers. Refer to table 1 for a detailed summary of operator demographics.

Demographic Variable	Total (N=48)		Calm Water (n=16)		Heat Map (n=16)		Thesaurus (n=16)	
Age (years)	M	SD	M	SD	M	SD	M	SD
Operator age	37.7	9.9	38.3	9.1	37.7	12.0	37.2	9.0
Experience (years)	M	SD	M	SD	M	SD	M	SD
DCS experience	7.5	8.0	6.6	6.6	9.3	11.0	6.3	5.4
Field experience	10.3	7.5	10.5	7.7	11.8	9.7	9.3	6.0
Total industry experience	13.6	9.3	12.2	7.8	15.2	11.7	13.5	8.4
Normal vision	N		n		n		n	
Yes	46		15		16		15	
No	2		1		0		1	
Gender	N		n		n		n	
Male	43		15		15		13	
Female	5		1		1		3	

Table 1. Operator demographics broken down by age, experience, vision, and gender across 3 level 1 overview displays.

Operator participation in the study was voluntary and operators were given the opportunity to leave the study at any time, without penalty or notification to their supervisors. One of the conditions of the research team being present and working with console operators was that the sessions had to be carried out during the night shift. However, this should not be of concern because all console operators switch between day and night shifts on a weekly basis. Given this, the research team was able to capture a wide range of operators during a 2-week stay at the plant facilities.

Materials and Apparatus

Two computer workstations were used to conduct the primary and secondary tasks. The primary task consisted of two displays and the secondary task consisted of only one display. The two displays used for the level 1 overview displays were 17 inches and 1280x1024 resolution. The display used for the secondary task was 15 inches and 1280x800 resolution. The primary task was a crude oil process-monitoring task presented on two displays on the left side of the operator's field of view and consisted of a schematic overview display on the upper display (regardless of the experimental condition – see "Level 1 Overview Displays" section below for more information) and the experimental overview display (depending on the condition) on the lower display. The MAT-B II (Comstock & Arnegard, 1992) was used as the secondary task and was located on the right side of the operator workstation. This task battery consisted of resource management and system monitoring tasks. Each workstation had its own set of controls (one standard computer mouse per workstation). It should be noted that operators were instructed to place equal importance on the two tasks and were not told which task was "primary" or "secondary" – these names are simply used in the current report to provide distinction between the two tasks. Figure 17 shows a graphical depiction of the operator workstation.

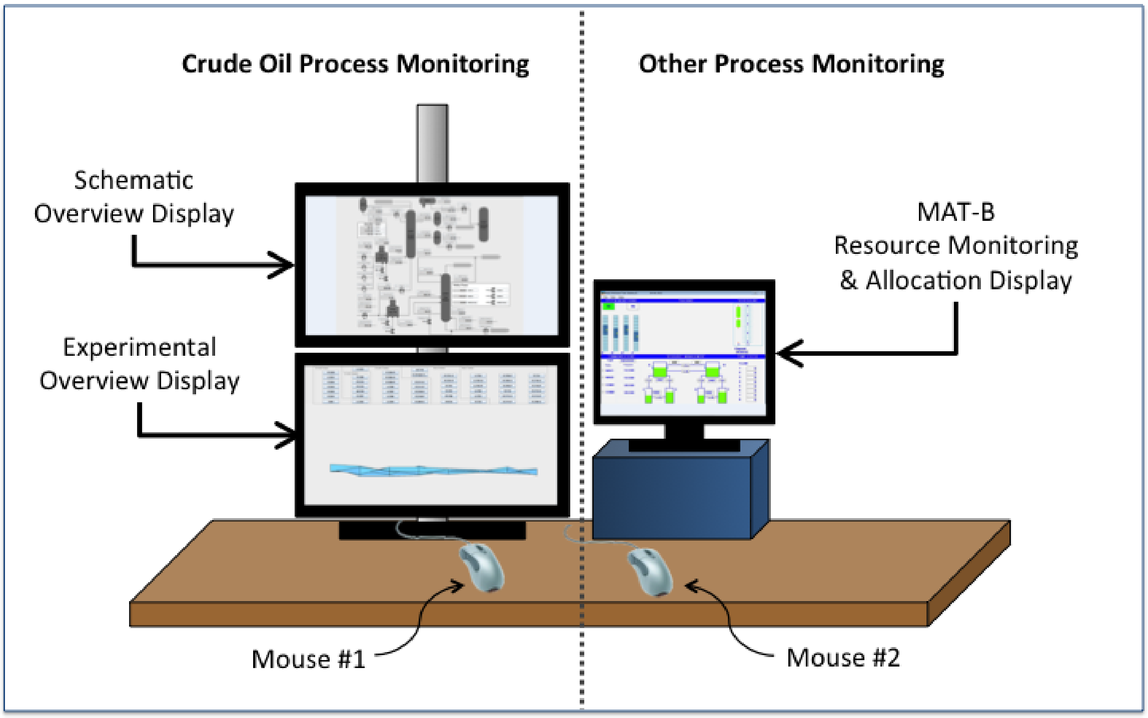

Figure 17. Experimental operator workstation configuration. On the left, top: Schematic overview display of the plant that was the same for all participants. Left, bottom: Experimental Crude Oil Process Monitoring (primary task, calm water display shown). On the right: Other Process Monitoring (secondary task, the same for all participants).

Level 1 Overview Displays. Level 1 overview displays are always in view of the operator and typically placed in the same location in operator workstations (upper-left of the workstation). The workstation used in the current study is comprised of two level 1 overview displays. Given this, the research team decided to place the two overview displays on the left side of the workstation (refer to figure 17 above). In all conditions, there were two level 1 overview displays used – the Schematic overview display (see description below) plus one of the 3 experimental overview displays (Calm Water, Heat Map, Visual Thesaurus), depending on the condition. This was done to compare each display equally; since one of the displays used in the study (Calm Water) was paired with the Schematic display in actual implementation and maintaining that pair (Calm Water plus Schematic) was a requirement for the original designers of this display.

The schematic overview display. This display was present, on the top-left portion of the workstation, at all times during the experiment regardless of experimental condition – it was always paired with the experimental level 1 overview display. The Schematic display is a piping and instrumentation diagram and uses graphic symbols or shapes to represent major equipment units in the plant, lines connecting the equipment units indicate direction of product flow through the plant, and process values are indicated with numeric values for each component. All 5 functional areas (crude feed, crude heater, crude tower, vacuum heater, and vacuum tower) are represented in this display. Refer to figure 18 for a graphical representation of the Schematic overview display.

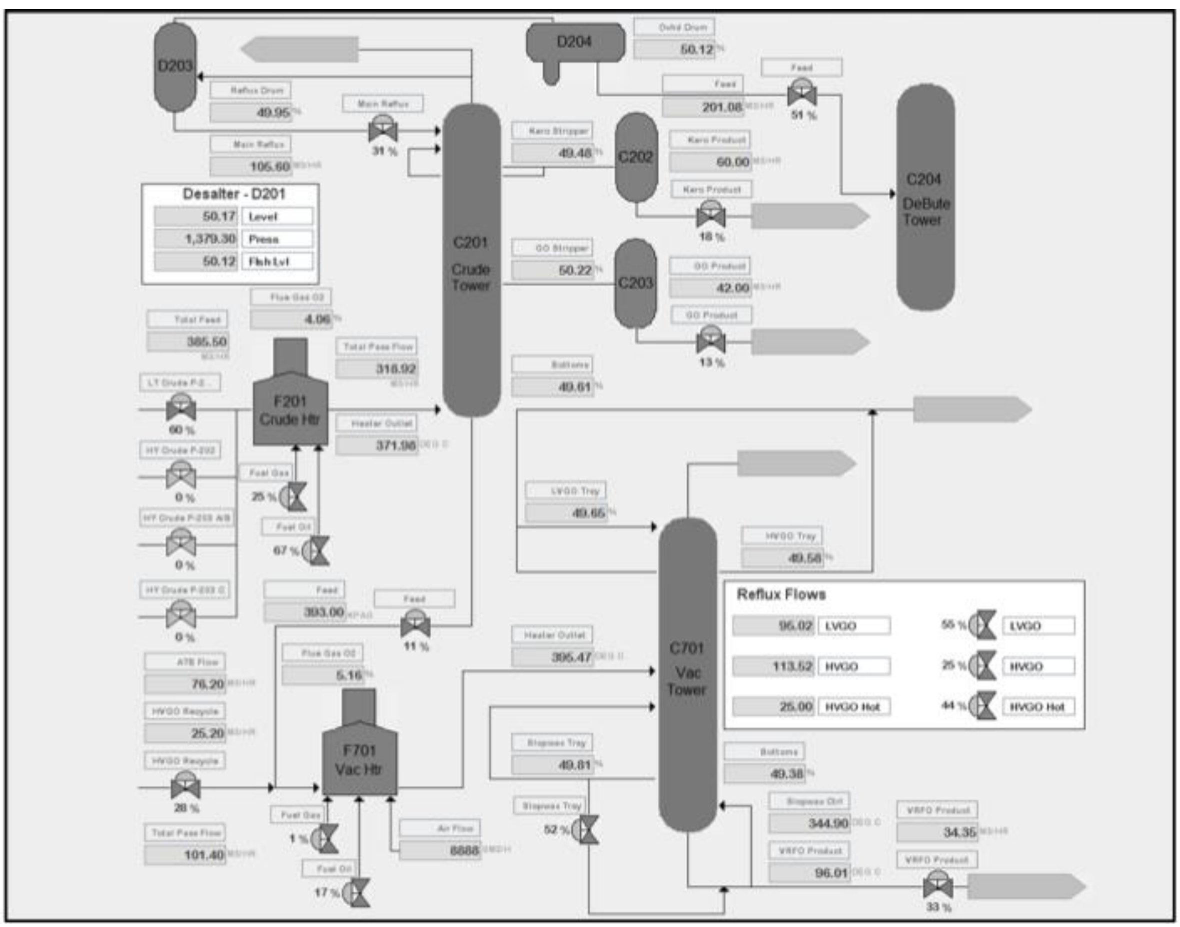

Figure18. Illustration of Schematic overview display. In this representation, there are no process deviations, hence the complete gray scale image.

The display uses a series of gauges to represent components in each of the functional areas contained on the display. These gauges are either process values (PV – indicating the actual value of the component represented by the gauge) or operator output percentages (OP – indicating the percentage of the valve opening, dictated by the set point of the component represented by the gauge). Process values are indicated by a current value and associated units. Operator output percentages are indicated by a valve icon and a numerical percentage to represent the percentage of the valve opening. Each gauge uses colored alarm indicators to alert operators of process deviations above or below normal operating conditions in the alarm range. However, this display does not have any indicators for deviations in the abnormal range. Yellow triangles are used to indicate deviations in the "low alarm" range, blue inverted triangles are used to indicate deviations in the "high alarm" range, and red squares are used to indicate both extreme high and extreme low alarms ("Hi-Hi" and "Lo-Lo"). These are alarms that are extremely high above or extremely low below normal operating conditions. Refer to figure 19 for a graphical depiction of alarms used in the Schematic display.

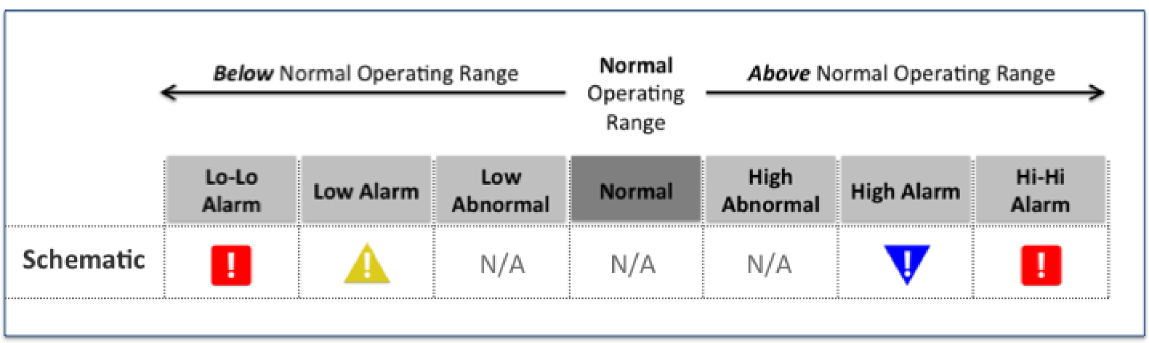

Figure 19. Illustration of alarms used in the Schematic overview display.

Each gauge has a button for input associated with it. These buttons are used for operators to acknowledge changes in gauge state (by clicking on the button). The button also serves as the label for each gauge. Even though both the schematic and experimental level 1 overview displays display the same values and alarms, operators were only required to click one of the two displays to acknowledge an alarm. Refer to figure 17 for a graphical depiction of a gauge with and without an alarm present.

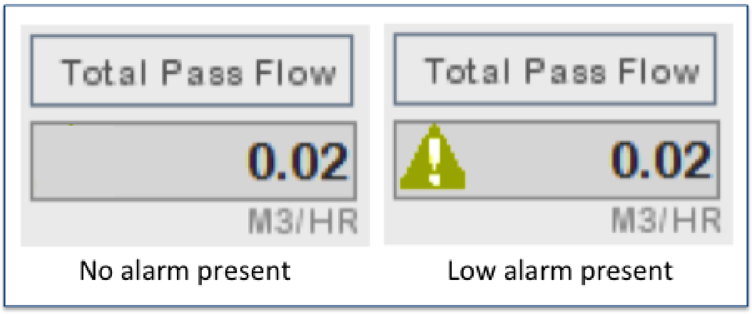

Figure 20. Illustration of gauge and button ("Total Pass Flow" rectangle) used in the Schematic overview display.

Experimental Level 1 overview display. Depending on the condition, one of the 3 experimental overview displays (Calm Water, Heat Map, or Visual Thesaurus) was present, on the bottom-left portion of the workstation, at all times during the experiment. It was always paired with the Schematic display located on the top-left portion of the workstation.

Primary Task. The primary task was a crude oil process-monitoring task that consisted of two types of level 1 overview displays; a Schematic overview display (this display was present during all experimental sessions, regardless of the experimental condition) and a second overview display (this display varied between participants, depending on the experimental condition). Operators were trained on and instructed to monitor a simulated crude oil process where they were responsible for acknowledging changes in gauge states.

The experimental platform followed predetermined simulated process scenarios produced for use in the study and presented the changing process values in the crude oil process. This ensured that the timings of the process changes were consistent for all participants (providing the researchers with the ability to make comparisons between experimental sessions). The simulated crude oil process was designed to imitate typical process control activities to provide realistic scenarios to the operator participants. The scenarios used in the current study were modeled after a previous study done by Tharanathan et al. (2010) and consisted of a crude oil process plant monitoring situation where gauge states would fluctuate between “normal”, “abnormal”, and “alarm” conditions. It is important to note that the representation of “normal”, “abnormal” and “alarm” conditions in each of these displays differs slightly. This is due to differences in implementation and specific design requirements for each display. While there are “real world” drivers for these differences, the limitation is important to note and is expanded upon during the discussion portion of this report. Each scenario lasted approximately 9 minutes and participants were tasked with acknowledging these gauge state fluctuations (gauge state changes) by interacting with the overview displays and clicking buttons for each gauge that changed state.

Secondary Task. The secondary task was a generic process monitoring and control and was comprised of two tasks from the MAT-B II developed by Comstock and Arnegard (1992) as a means for simulating operator conditions. One aspect was a resource-monitoring task. This task simulates a key function of a control room operator: managing dynamic resources to maintain a healthy system. Participants were instructed to monitor and adjust to adequate plant levels. The other aspect was a system monitoring task. This task simulates another key function of a control room operator: monitoring a system to ensure that it is functioning properly and taking corrective action to return to a steady state. Participants were instructed to monitor for levels outside of a predetermined range and return the system to normal by clicking on the appropriate button to reset the system. It is important to note that the MAT-B II task simulated a different process compared to the primary task process. These two processes were not related. The MAT-B II is also capable of simulating Tracking, Communications, and Scheduling tasks; however, these tasks were disabled for this study and participants were instructed not to use these areas of the display. Refer to figure 21 for a graphical depiction of the secondary task.

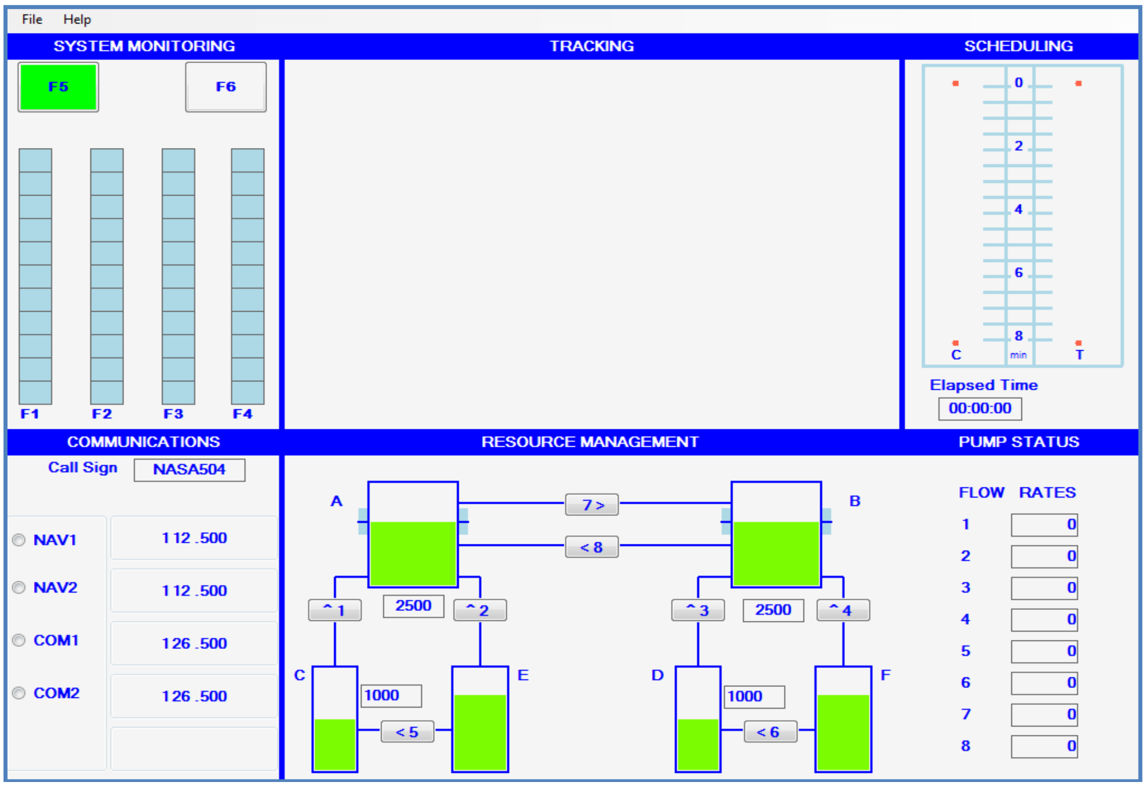

Figure 21. Operator's secondary task display – Multi-Attribute Task Battery.

Subjective Workload. In addition to assessing the potential effects of scenario complexity on subjective (perceived) workload, the impact of the various visualization techniques used across the level 1 displays on subjective workload were also of interest in the current study. NASA-TLX (Hart & Staveland, 1989) was used to collect subjective mental workload measures from participants following each experimental session. The weighted score served as measures of subjective mental workload.

Situation Awareness. In the current study, Two of the 4 situation awareness measures discussed by Pew (2000) were used to collect situation awareness data from operators: direct performance and direct experimental measures. Direct performance measures data was collected in the form of responses to deviations in gauge status during the simulation and was captured in the form of hits, misses, and false alarms to the gauge changes. Direct performance measures were chosen due to the dynamic nature of identifying and reacting to abnormal situations or alarms in the simulated environment. Operator performance was measured using data collected in the form of situation awareness probes, administered twice randomly, during each experimental session. These probes were aimed at assessing operators' current awareness of the plant status (e.g., *"Since the beginning of the scenario, three AREAS (not just gauges) have deviated from normal state. A) True B) False"*). See Appendix D for a complete set of questions and correct responses. Direct experimental measures were chosen due to the ability for experimenters to administer these probes at random intervals during the dynamic simulation, and resume the scenario exactly where it left off upon completion of the probe questions, providing a relatively non-disruptive way to measure operators' current awareness of plant conditions. When these probes were administered, the screens went blank and prompted operators and experimenters to begin the probe. Both direct performance and direct experimental measures were created and administered in accordance with previous research (Tharanathan et al., 2010).

Eye Tracker. Operators were fitted with a head mounted eye tracker and completed the tasks while wearing the eye tracker. This was intended to provide region of interest data across two tasks: (1) the top-left (overview) display, (2) the bottom-left (overview) display, and (3) the right (MAT-B II) display. However, during analysis it was discovered that the data collected did not yield this information. There was an unknown error with the tracking software and the eye tracking data did not provide reliable data so it was omitted from the analysis.

Procedure

Participants entered the testing room individually and at different times. They read and signed an informed consent form and completed a demographic questionnaire. They were assigned to an experimental group, based on a predetermined Latin Square table to ensure that the presentation order of scenarios was equal across participants. Every experimental group completed one high-complexity scenario and one low-complexity scenario. There were a total of four possible scenario orders (1H, 2H, 1L, 2L). Four orders were used (1H + 1L = Order A, 1H + 2L = Order B, 2L + 1H = Order C, 2L + 2H = Order D). Only two orders were chosen per experimental session, one for trial 1 and one for trial 2. A researcher briefly explained the nature of the experiment and provided detailed and scripted instructions about how to perform the required tasks. Participants were fitted with the eye tracker and calibration was conducted. Participants completed the primary and secondary tasks individually and then concurrently to familiarize themselves with the nature of each task. The goal of this was to establish familiarity to the tasks and interface used during the study. This was an opportunity for participants to ask questions and receive help and feedback prior to beginning the experimental trials. Once familiarity was established, participants completed four 9-minute sessions, separated by the opportunity for a break. The experiment lasted approximately two and a half hours, this included greeting participants, providing an overview of the study, calibration of the eye tracker, training, and conducting the study itself. Calibration of the eye tracker was the most time consuming portion of the study – nearly 30 minutes per participant was allocated to calibration. Refer to figure 22 for a graphical depiction of the experimental procedure followed for this study.

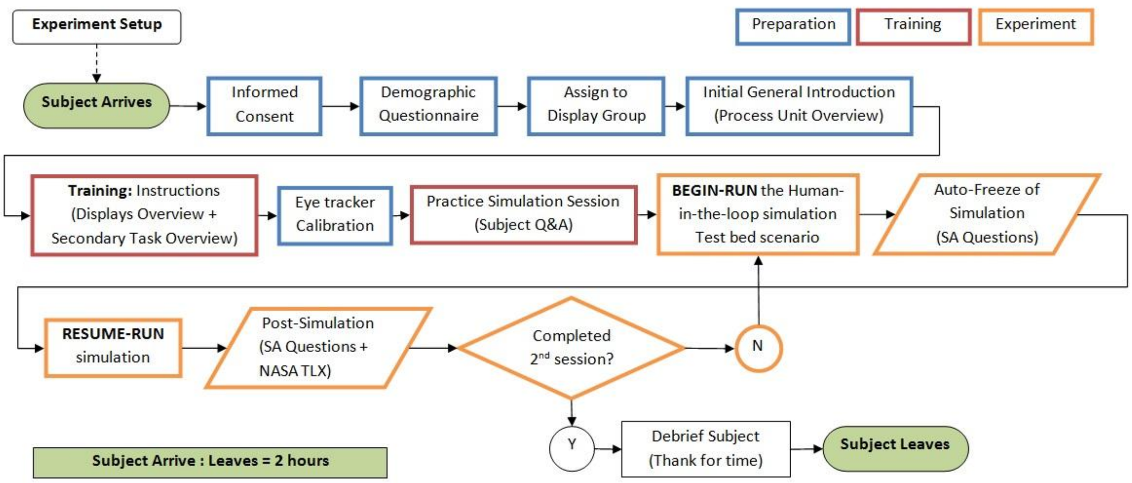

Figure 22. Illustration of the experimental procedure followed for the current study. Overall study duration lasted approximately 2 hours.

Experimental Sessions. An experimental session consisted of two monitoring tasks: (1) Monitor and acknowledge gauge state changes (process deviations) in the 2 level 1 overview displays by clicking the appropriate button on the displays (participants had access to both level 1 displays, not just the experimental one) and (2) monitor and interact with the secondary task (MAT-B II), where they were required to make adjustments in the MAT-B II interface to maintain a healthy system on the MAT-B II portion of the task. The experimental session lasted approximately 9 minutes. During this time, there were two random situation awareness probe freezes. The first freeze occurred near the middle of the 9-minute scenario (upon completion of the first situation awareness freeze probe, the session resumed where it left off) and the second freeze marked the end of the scenario (however, participants were not aware when these freezes would occur). During these freezes, operators were presented with a series of 8 questions about the crude oil process that they were monitoring. Following the conclusion of each experimental session (i.e., the conclusion of the second situation awareness freeze), operators were administered NASA-TLX subjective workload probe.

Scenario Complexity. Scenarios were created using input from process experts and a Crude Distillation Unit. This unit was used to represent a typical Crude Unit console operator unit. As previously mentioned, the main process areas (functional areas) represented in the scenarios consisted of: crude feed, crude heater, crude tower, vacuum heater, and vacuum tower. Scenario complexity was varied in accordance with Tharanathan et al. (2010) as a within-subject variable. Following a practice scenario, each operator completed one low-complexity and one high-complexity scenario. Presentation order of low- and high-complexity scenarios was systematically counterbalanced across all operators in the study. Scenarios were broken down such that low-complexity scenarios consisted of 24 or 32 total process deviations and high-complexity scenarios consisted of 47 process deviations, depending on the scenario assigned to that particular condition. The amounts of deviations were intended to simulate actual operating conditions over a similar period, while monitoring an actual process. The initiating event for each scenario (the first gauge state change) occurred in different main functional areas of the overview displays to minimize the chance for operator bias to look for deviations in particular parts of the workstation. It is important to note that operators were not instructed which display was the display of interest and therefore not given any guidance as to which display they should interact with more. This decision was made by the research team and project stakeholders because the Calm Water display was paired with the Schematic display in actual use.

Chapter 3: Results

The study design leads to a three-factorial design with the factors display type (between subjects), scenario complexity (within subjects), and first- vs. second trial (within subjects). Due to the specific nature of the study, scenario complexity and trial number were not fully crossed but dependent on each other. If the high-complex scenario was seen first by the operators, the low-complexity scenario was experienced second. For an initial analysis, the data was recoded with the within-subject variable display complexity and the two between subject variables display type and trial order (high-low vs. low-high) to see whether the order of trials had any effect. Trial order or interactions involving trial order were only significant for one dependent variable (NASA-TLX) and will be reported there. In all other cases we simplified the ANOVA to increase the statistical power of the analysis and the interpretability of the results to a two-way mixed ANOVA, assessing the impact of level 1 overview display type and scenario complexity on operator detection performance, operator workload, and operator awareness. Additionally, a Two-Way Mixed ANOVA was conducted to assess the impact of level 1 overview display type and practice effects on performance between the first and second experimental trials. For all analyses, the level of significance was set at alpha = .05; if results were non-significant, they are generally not reported below. See descriptions of each operator performance measure in Table 2; including how the data was collected and what it represents for each measure. Table 3 provides descriptive statistics for scenario complexity and Table 4 provides descriptive statistics for experimental trial.

Operator Performance Measures	Data Collection Apparatus	How The Data Was Collected	Data Range	What the Values Represent
Operator Detection Accuracy				
Process Deviation Detection Accuracy	Experimental Display (primary task)	By recording operator responses to process deviations and determining whether they were accurate	0-100	% Accuracy
Response Time to Process Deviations	Experimental Display (primary task)	By recording how long it took operators to respond to process deviations	0-60	Seconds
Operator Workload				
Subjective Workload	NASA TLX Probes	By recording operator responses to the NASA TLX subjective workload probe	0-100	% Scale (higher means more workload)
Secondary Task Performance	Multi-Attribute Task Battery (secondary task)	By recording operator responses on resource management and system monitoring tasks and determining whether they were accurate	0-100	% Accuracy
Operator Awareness				
Situation Awareness Probes	Situation Awareness Level 2 Probes	By recording operator responses to situation awareness probes about plant status and determining whether they were accurate	0-100	% Accuracy

Table 2. Reference table for the data collected and used to analyze the operator performance measures in the study.

Scenario Complexity

Operator Detection Accuracy

Process Deviation Detection Accuracy	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Scenario Complexity	M	SD	M	SD	M	SD	M	SD
Low	41.35	9.79	36.06	16.14	33.95	12.52	37.12	13.18
High	28.40	9.06	19.66	8.24	19.58	7.22	22.54	9.06
Overall	34.87	11.70	27.86	17.16	26.76	13.52		

Response Time to Process Deviations	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Scenario Complexity	M	SD	M	SD	M	SD	M	SD
Low	27.90	18.51	31.84	15.48	41.53	23.57	33.76	19.91
High	26.35	11.44	26.70	17.66	35.16	10.85	29.40	14.00
Overall	27.13	20.20	29.27	19.89	38.34	24.78		

Operator Workload

Subjective Workload	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Scenario Complexity	M	SD	M	SD	M	SD	M	SD
Low	62.63	24.19	56.44	23.99	67.13	18.79	62.06	22.42
High	67.13	21.79	52.94	22.00	63.19	17.83	61.08	21.07
Overall	64.88	28.68	54.69	28.59	65.16	22.63		

Secondary Task Performance: Mat-B System Monitoring	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Scenario Complexity	M	SD	M	SD	M	SD	M	SD
Low	58.38	18.51	45.06	17.55	51.88	16.44	51.77	18.00
High	58.31	13.78	42.38	18.17	51.75	10.26	50.81	15.60
Overall	58.34	20.92	43.72	21.75	51.81	17.97		

Secondary Task Performance: Mat-B Resource Management	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Scenario Complexity	M	SD	M	SD	M	SD	M	SD
Low	78.00	21.52	79.13	23.31	66.75	28.85	74.63	24.87
High	72.56	26.06	74.94	27.87	69.81	25.43	72.44	25.99
Overall	75.28	28.33	77.03	30.52	68.28	34.00		

Operator Awareness

Situation Awareness Probe	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Scenario Complexity	M	SD	M	SD	M	SD	M	SD
Low	31.64	21.22	37.11	11.96	41.80	11.11	36.85	15.70
High	35.16	19.88	42.97	12.26	41.80	10.13	39.97	14.80
Overall	33.40	25.45	40.04	14.77	41.80	13.22		

Table 3. Descriptive statistics for scenario complexity and level 1 overview display type on operator detection accuracy, workload, and awareness measures.

Experimental Trial

Operator Detection Accuracy

Process Deviation Detection Accuracy	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Experimental Trial	M	SD	M	SD	M	SD	M	SD
Trial 1	33.04	10.00	28.39	10.12	27.45	14.61	29.63	11.78
Trial 2	36.71	12.66	27.33	19.21	26.08	10.24	30.04	15.01
Overall	34.87	13.42	27.86	16.94	26.76	16.30		

Response Time to Process Deviations	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Experimental Trial	M	SD	M	SD	M	SD	M	SD
Trial 1	23.84	16.48	28.11	14.29	35.76	10.68	29.24	14.60
Trial 2	30.41	13.42	30.42	18.94	40.93	23.80	33.92	19.44
Overall	27.13	19.02	29.27	19.58	38.34	19.93		

Operator Workload

Subjective Workload	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Experimental Trial	M	SD	M	SD	M	SD	M	SD
Trial 1	67.81	21.35	54.94	22.65	70.50	15.06	64.42	20.70
Trial 2	61.94	24.43	54.44	23.51	59.81	19.79	58.73	22.41
Overall	64.88	27.46	54.69	28.10	65.16	20.56		

Secondary Task Performance: Mat-B System Monitoring	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Experimental Trial	M	SD	M	SD	M	SD	M	SD
Trial 1	58.25	12.33	41.50	17.40	50.69	8.52	50.15	14.70
Trial 2	58.44	19.51	45.94	18.13	52.94	17.33	52.44	18.68
Overall	58.34	18.50	43.72	21.61	51.81	14.92		

Secondary Task Performance: Mat-B Resource Management	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Experimental Trial	M	SD	M	SD	M	SD	M	SD
Trial 1	72.50	24.63	79.44	21.98	67.44	25.36	73.13	24.04
Trial 2	78.06	23.12	74.63	28.87	69.13	28.98	73.94	26.81
Overall	75.28	29.56	77.03	30.00	68.28	32.60		

Operator Awareness

Situation Awareness Probe	Display							
	Calm Water (n=16)		Heat Map (n=16)		Visual Thesaurus (n=16)		Overall	
Experimental Trial	M	SD	M	SD	M	SD	M	SD
Trial 1	23.05	21.74	40.63	10.46	37.11	8.68	33.59	16.39
Trial 2	43.75	12.29	39.45	14.20	46.48	10.20	43.23	12.42
Overall	33.40	23.41	40.04	14.50	41.80	11.29		

Table 4. Descriptive statistics for experimental trial and level 1 overview display type on operator detection accuracy, workload, and awareness measures.

Operator Detection Accuracy

***Process deviation detection accuracy*.** Data was collected by recording operator responses to process deviations on the primary task and determining whether they were accurate. Accuracy was determined by operators' ability to detect a process deviation and click on the appropriate gauge before that deviation automatically returned to normal. Results represent the accuracy scores on a scale of 0-100 for detection performance. A score of 100 indicates perfect performance.

The analysis showed a significant main effect for display on process deviation detection accuracy $F(2, 45) = 3.41$, $p = .042$, $\eta_p^2 = .131$, observed power = .611. Fisher's LSD post-hoc comparisons indicated a significant difference between the Calm Water display ($M = 34.87$, $SD = 11.70$) and both the Heat Map display ($M = 27.86$, $SD = 17.16$) and Visual Thesaurus display ($M = 26.76$, $SD = 13.52$). The Calm Water display yielded better process deviation detection accuracy. There was no significant difference between the Heat Map display and the Visual Thesaurus display ($p = .747$). See figure 23 below.

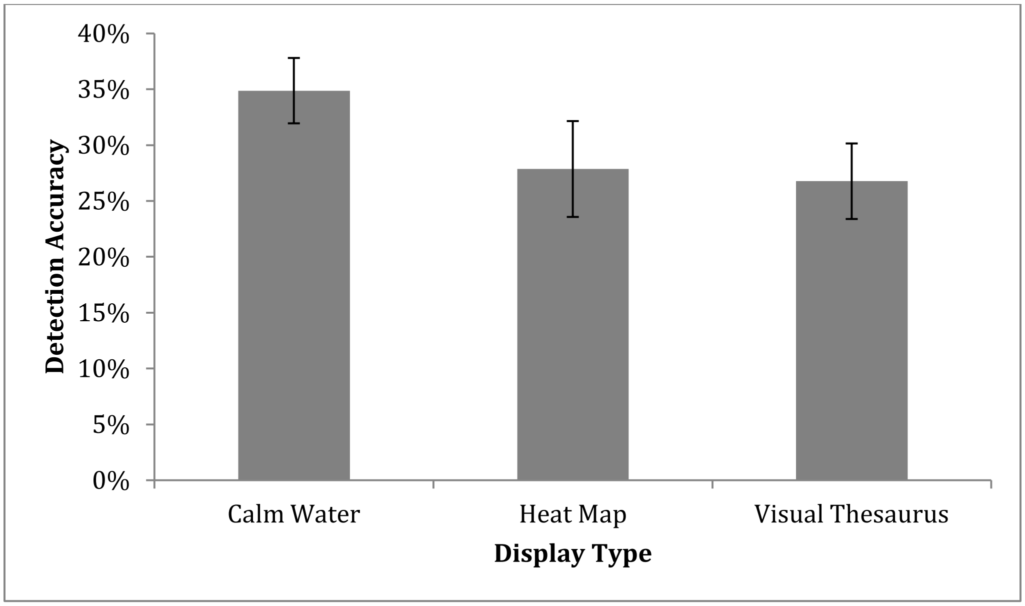

Figure 23. Average detection accuracy for different display types. The Calm Water display yielded better performance click accuracy. Error bars represent the standard error of the mean.

As expected there was a main effect for scenario complexity, $F(1, 45) = 90.19$, $p < .001$, $\eta_p^2 = .667$, observed power = 1.00. Overall, performance was better during the low complexity scenarios ($M = 37.12$, $SD = 13.18$) compared to the high complexity scenarios ($M = 22.54$, $SD = 9.06$). These results are graphically depicted in figure 24.

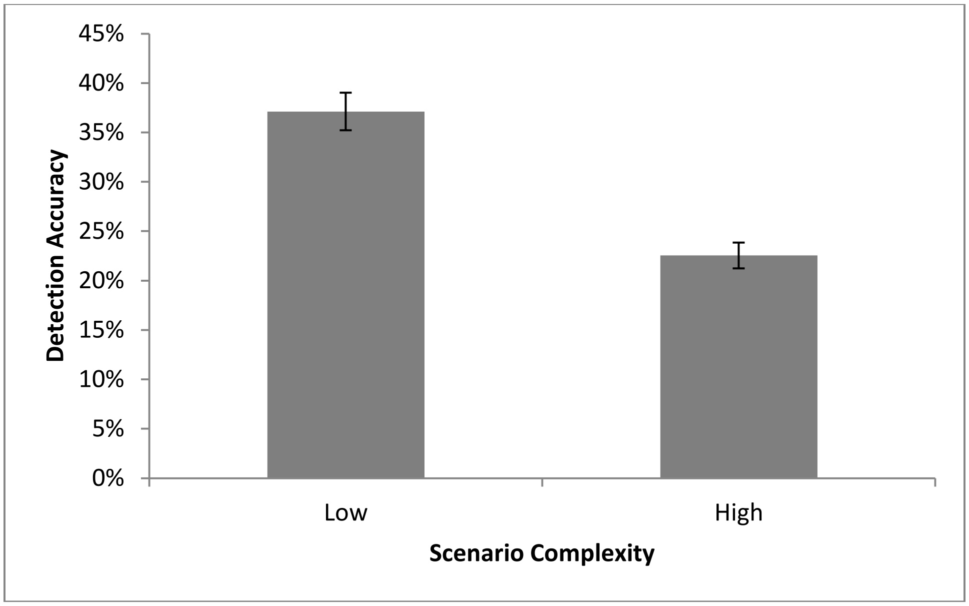

Figure 24. Average detection accuracy high and low scenario complexity levels. Performance was better during low complexity scenarios. Error bars represent the standard error of the mean.

There were no significant interaction effects found for scenario complexity and display type (*p* = .655). Additionally, there were no significant main effects found for experimental trial (p = .878), indicating that there were no learning effects for process deviation detection accuracy.

Response time to process deviations. Data was collected by recording how long it took operators to respond to process deviations on the primary task. Therefore, results represent the amount of time (in seconds) it took for operators to respond to process deviations. Participants were allowed to respond to process deviations at any time, including after a minute has passed; however, there were not any mean response times greater than 60 seconds for any of the experimental groups.

The analysis showed a significant difference in response time among the three level 1 overview displays $F(2, 45) = 4.44$, $p = .017$, $\eta_p^2 = .165$, observed power = .735. Fisher's LSD post-hoc comparisons indicated a significant difference (p =.007) between the Calm Water display (M = 27.13, SD = 20.20) and the Visual Thesaurus display (M = 38.34, SD = 24.78). The Calm Water display yielded better response time performance than the Visual Thesaurus display. Post-hoc comparisons also indicated a significant difference (p = .028) between the Heat Map display (M = 29.27, SD = 19.89) and the Visual Thesaurus display (M = 38.34, SD = 24.78). The Heat Map display performed better than the Visual Thesaurus display. In both cases, the Visual Thesaurus display yielded slower reaction time performance. There was no significant difference found between the Calm Water and the Heat Map display (p = .595). See figure 25 for a graphical depiction of the results.

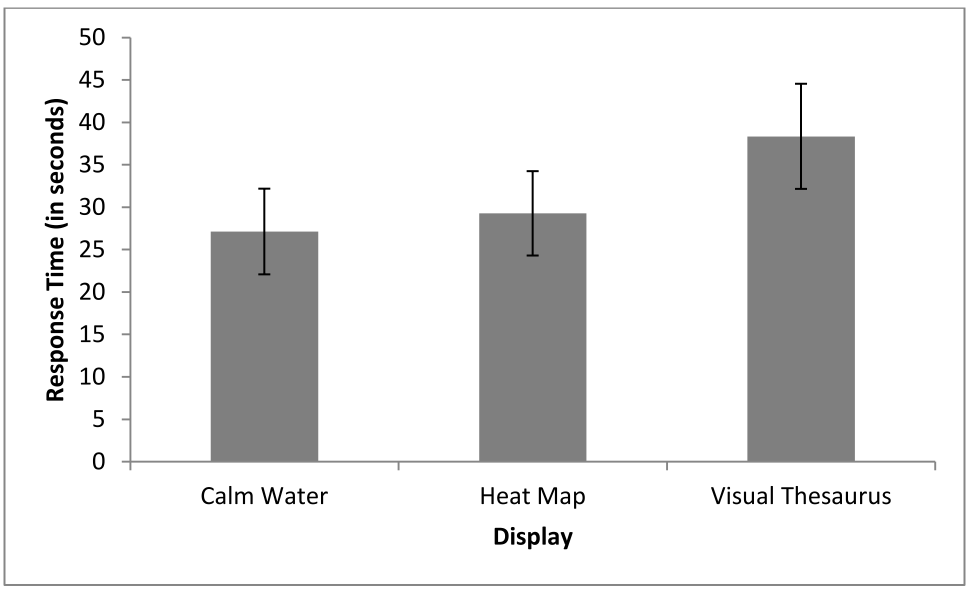

Figure 25. Average response time for different display types. Performance was worst with the Visual Thesaurus display. Error bars represent the standard error of the mean.

There were no significant main effects found for scenario complexity and display type (p = .852). Additionally, there were no effects found for experimental trial (p = .199), indicating that there were no learning effects for response time.

Operator Workload

Subjective workload: NASA TLX. Data was collected by recording operator responses to the NASA TLX subjective workload probe. Results represent the amount of workload reported by operators on a scale of 0-100. A score of 0 indicates no reported subjective workload. See Appendix E for an example of the NASA TLX Subjective workload probe administered to operators for the current research.

The analysis showed a statistically significant difference in workload between the first and second experimental trials across all displays $F(1,45) = 7.42$, $p = .009$, $\eta_p^2 = .141$, observed power = .760. Workload was higher in the first experimental trial ($M = 64.42$, $SD = 20.70$) compared to the second experimental trial ($M = 58.73$, $SD = 22.41$) overall. Figure 26 provides a graphical depiction of these results.

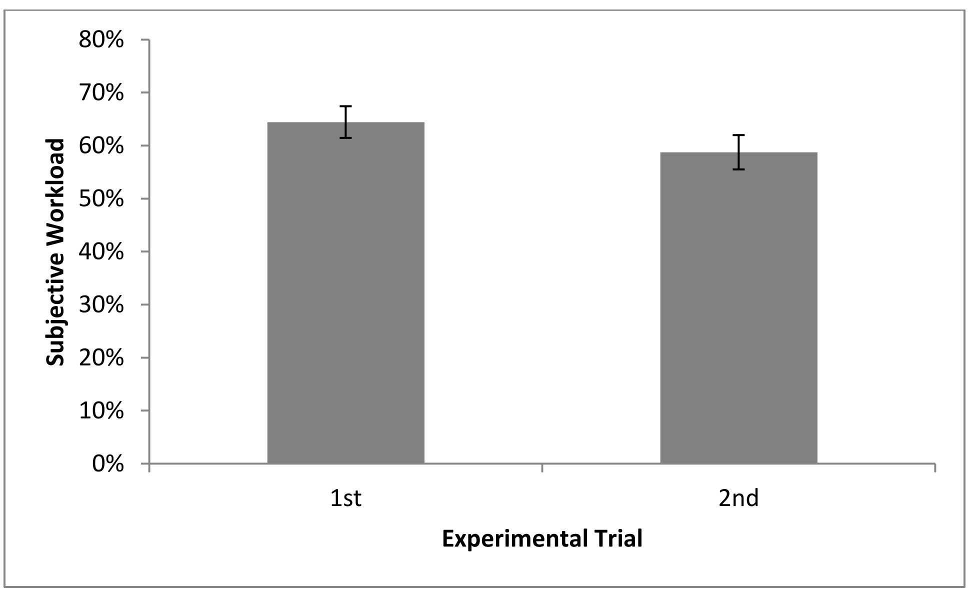

Figure 26. Subjective workload scores for first and second experimental trials. Workload was lower during the second experimental trial. Error bars represent the standard error of the mean.

For the subjective workload measure, there were no significant main effects found for scenario complexity and display type (p = .240). Additionally, there were no interaction effects found for experimental trial and display (p = .149).

For trial order, the analysis revealed a significant interaction effect of trial order and scenario complexity $F(1, 45) = 7.115$, $p = .011$, $\eta_p^2 = .377$, observed power = .926. When presented with order two (low then high complexity), operators reported lower workload during high complexity scenarios (M = 56.00, SD = 16.22) compared to low complexity scenarios (M = 62.52, SD = 18.61). These results are graphically depicted in figure 29.

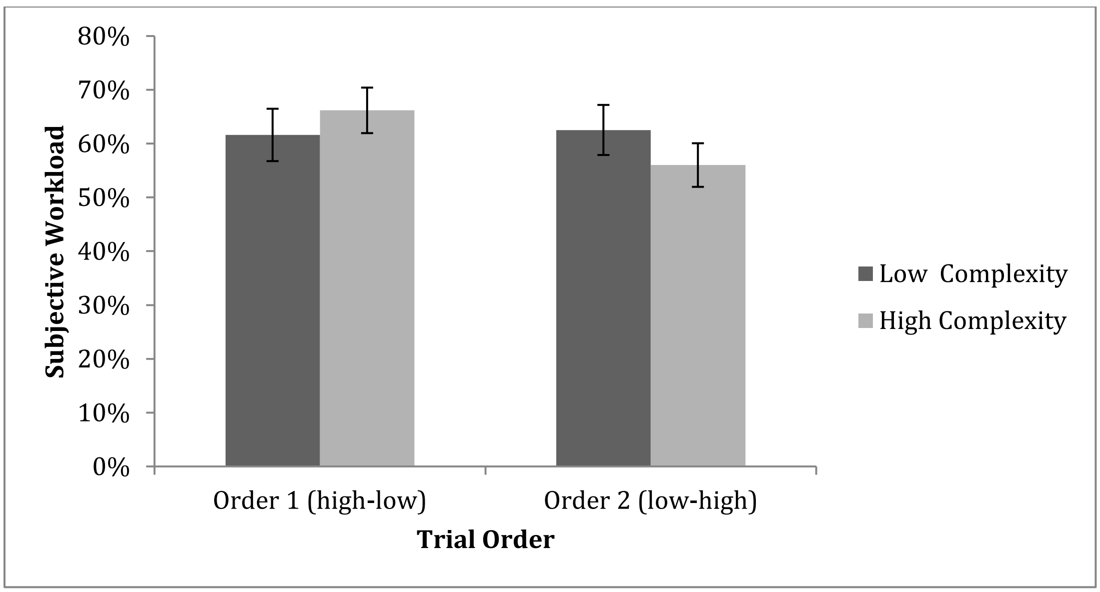

Figure 27. Subjective workload scores for trial order and scenario complexity. Workload was lower in order 2 during high complexity scenarios.

Secondary task performance: MAT-B II. Data was collected by recording operator responses on resource management and system monitoring tasks and determining whether they were accurate. Results represent the accuracy scores on a scale of 0-100 for detection performance. A score of 100 indicates perfect performance.

The analysis showed a significant difference in secondary task (MAT-B II) System Monitoring performance accuracy among the three level 1 overview displays $F(2, 45) = 4.53$, $p = .016$, $\eta_p^2 = .168$, observed power = .744. Fisher's LSD post-hoc comparisons indicated a significant difference between the Calm Water display ($M = 58.34$, $SD = 20.92$) and the Heat Map display ($M = 43.72$, $SD = 21.75$). Operators in the Calm Water condition had better performance compared to operators in the Heat Map condition. These results are graphically depicted in figure 28. There were no significant differences found between the Calm water and Visual Thesaurus display ($p = .186$) or the Heat Map and Visual Thesaurus display ($p = .103$).

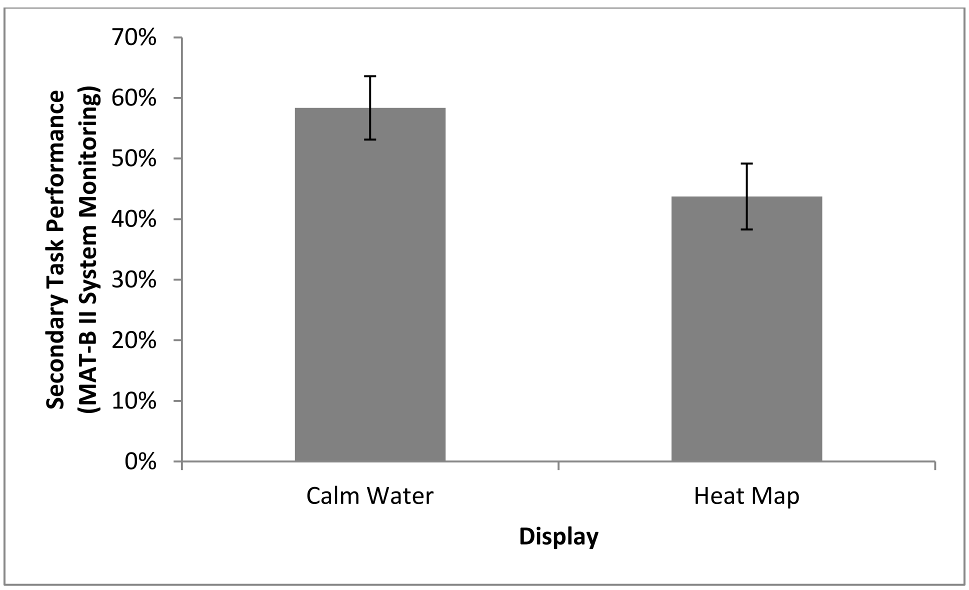

Figure 28. Secondary task performance scores for the MAT-B II system workload task for Calm Water and Heat Map displays. Performance was higher with the Calm Water Display. Error bars represent the standard error of the mean.

There were no significant interaction effects found for scenario complexity and display type ($p = .877$) on the System Monitoring portion of the secondary task. Additionally, there were no effects found for experimental trial ($p = .336$), indicating that there were no learning effects for secondary task performance. For the Resource Management portion of the secondary task, there were no effects found for scenario complexity (p = .321) or experimental trial (p = .229).

Operator Awareness

Situation awareness level 2. Data was collected by recording operator responses to situation awareness probes about plant status and determining whether they were accurate. Results represent the accuracy scores on a scale of 0-100 for detection performance. A score of 100 indicates perfect performance.

Regarding experimental trial, the analysis showed a significant difference in situation awareness level 2 performance $F(1, 45) = 7.42$, $p < .001$, $\eta_p^2 = .252$, observed power = .968. Operators performed better during the second experimental trial (M = 43.23, SD = 12.42) compared to the first (M = 33.59, SD = 16.39). This finding is depicted in figure 29 below.

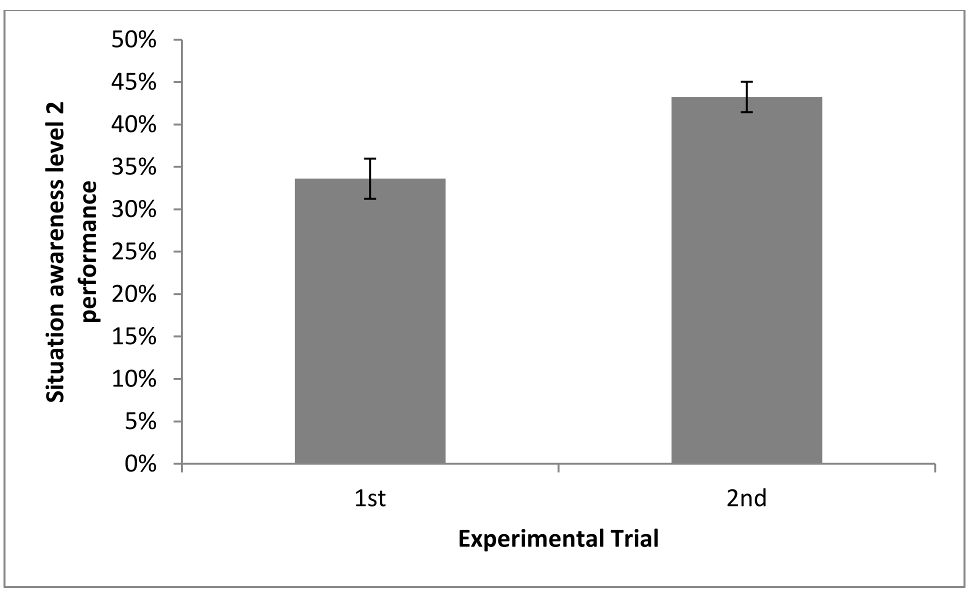

Figure 29. Situation awareness level 2 performance for experimental trial. Performance was better during the second experimental trial. Error bars represent the standard error of the mean.

The analysis also showed a statistically significant interaction effect of experimental trial and display $F(2, 45) = 6.53$, $p = .003$, $\eta_p^2 = .225$, observed power = .889. When using the Visual Thesaurus display, operators showed better situation awareness level 2 performance during the second trial of the experimental session ($M = 46.48$, $SD = 10.20$) compared to the first ($M = 37.11$, $SD = 8.68$). These results are graphically depicted in figure 30. There were no other significant effects for experimental trial.

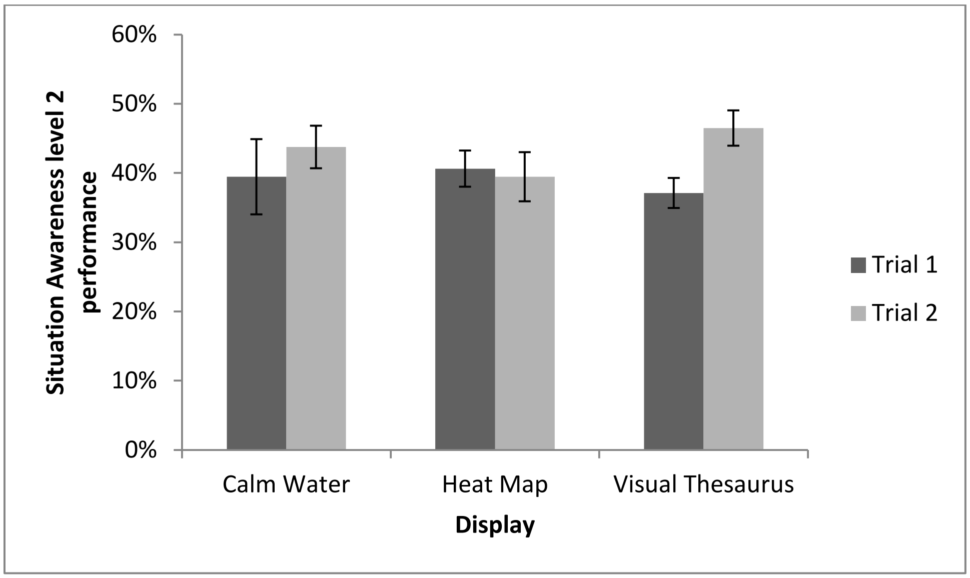

Figure 30. Situation awareness level 2 performance for experimental trial and display type. Performance while using the Visual Thesaurus display was better during the second experimental trial compared to the first. Error bars represent the standard error of the mean.

There were no significant effects found for scenario complexity and display type ($p = .319$), indicating that scenario complexity did not influence operator awareness.

Chapter 4: Discussion

Operator Performance

Operator Detection Accuracy. The finding that the Calm Water experimental display performed better at process deviation detection accuracy than Heat Map and Visual Thesaurus displays provides support for the prediction that this display would aid operators in correctly identifying process deviations by creating visual distinction on the display inputs when deviations occurred. Visual distinction was achieved by making the button responsible for indicating a process deviation bold, providing the operator an indication of precisely what part of the interface to intact with to acknowledge an alarm. Operators were able to notice and act on deviations more easily than with the other displays. One possible explanation for this advantage could be a strategy whereby operators simply watched the portion of the display that included the buttons and waited for them to turn bold so that they could respond to them without even having to reference the graphical portion of the display. While this is possible, it is not likely. The same display also provides very salient gauge state indications, so it is likely that operators were relying on this salient information in addition to watching for the changing buttons. Another reason to believe that operators did not use this strategy stems from the awareness measures. Focusing solely on the changes in the appearance of the buttons would cause operators to have less awareness of the system and reduce the overall information relevant for the status of the plant. However, measures of operator awareness do not show a significant difference between the calm water display and the other displays, making it unlikely that operators relied on this simple strategy.

Process deviation detection accuracy was best for all displays during the low complexity scenarios. Interestingly, this effect was not present with any of the other measures. This indicates that the increased scenario complexity leads to worse detection performance, but does not affect other operator performance measures. This is a finding that should be investigated further to determine whether there was a particular reason for this effect only on a single measure.

Response time performance was slower with the Visual Thesaurus display compared to the other two displays. This is not unexpected because of the nature of the layout on the Visual Thesaurus display. The buttons on the Calm Water display are located across the top of the screen and become visually distinct when a process deviation occurs. The Heat Map display, on the other hand, maintains a strict grid-style layout where the button is always directly to the left of the gauge. The Visual Thesaurus display also groups the buttons next to the gauges, but the graphical nature of the Visual Thesaurus may break up the perception of the 'natural' grid in the display and therefore lead to worse performance compared to the Heat Map display. There are also instances where buttons on the Visual Thesaurus display are located *above* the gauges, instead of to the left. The Heat Map display layout is more consistent and therefore likely allows operators to be quicker with responses.

Operator Workload. Overall, there appeared to be a learning effect across all 3 displays for the NASA-TLX workload measure. Operators indicated that they experienced less subjective workload during the second experimental trial compared to the first. In fact, operators experienced less subjective workload during the second scenario when they were presented with the low complexity scenario first. Some potential reasons for this could be that the tasks were too complicated or that the training session did not appropriately prepare operators for the tasks. Since workload did drop during the second scenario, it is more likely that lack of training factored into this result and caused the learning effect to occur.

We found that the secondary task performance for the System Monitoring in the MAT-B II task was better with the Calm Water display compared to the Heat Map display. We expected that the Calm Water would effectively aid in workload by providing indications of gauge level and gauge states through highly salient visual peaks and valleys on the display. This might have allowed operators to take advantage of the strong visual cues to capture their attention, assisting in their ability to multi task on the secondary task (Mat-B II). The comparison between the Calm Water display and the Visual Thesaurus display were non-significant. It should be noted that this was only for the system-monitoring task in the MAT-B II task. The Resource Monitoring task did not yield any significant results. So, while the Calm Water display led to better workload performance, it was not exceedingly better than the other displays across all operator workload measures.

Operator Awareness. Operators using the Visual Thesaurus display showed a learning effect between experimental trials 1 and 2 for the operator awareness measure. In fact, operators with the Visual Thesaurus display scored better during the second experimental trial (46%) compared to the Calm Water display (44%) and the Heat Map display (39%); however, these results were not significant. Responses to the situation awareness level 2 probe indicated that operators had better awareness of the plant during the second trial. This learning effect for operator awareness was not present in either of the other two experimental displays. The Visual Thesaurus display led to equal awareness performance during the first experimental trial, but allowed operators to improve their awareness during the second trial. While contrary to our predictions, this finding may have been due to the graphical indicators used to display gauges, which could have aided operators in identifying which gauges belonged to the various parts of the plant and therefore more easily map these to understand the status of the plant. The potential advantage of the graphical gauge indications likely benefits operators new to the display. It would be expected that operators would be able to perform equally well with this display as with the Calm Water or Heat Map displays after a few weeks of use due to becoming familiar with the plant and the processes they are monitoring. Operators' would develop a mental model of the plant as they interact with either of the displays, aiding their ability to maintain plant awareness. Another explanation for this finding is that operator performance during the first experimental trial might have been limited based on cognitive resources necessary to complete the process-monitoring task with a new display. Once operators became familiar with the displays and the task, the Visual Thesaurus display may have provided additional information compared to the Calm Water and Heat map displays. This would indicate that performance during the second trial might have been limited due to data available to operators. For instance, the Visual Thesaurus displays quantitative information with more precision than the other display types, since that information is represented as levels of gauges and dials. The Calm Water and Heat Map displays, on the other hand, use relative size and shades of color to indicate the same information. Processing the information on the Visual Thesaurus display is arguably easier, since quantitative information is harder to gain from relative size or shade of color.

Level 1 Overview Displays

Tharanathan et al. (2010) provided a description for how level 1 overview displays should support plant operators. A level 1 overview display should aid in operators' ability to learn about plant situation awareness by providing an overview of the plant. This helps operators detect process deviations and maintain an overview of the current status of the plant. They must aid in capturing the attention of operators when process deviations are detected and allow operators to act on those events (Xiao & Seagull, 1999). By these definitions, an effective level 1 overview display should: help maximize operators' detection and minimize response time to process deviations, minimize operators' workload, and supply effective awareness of the plant. The tenets of EID (Rasmussen & Vicente, 1989) and the SRK framework (Jamieson et al., 2002; Rasmussen, 1983; Vicente, 2002) can be applied to interface design to develop effective level 1 overview displays for process monitoring (Jamieson et al., 2002). In the current study, we have evaluated various implementations of EID and the SRK framework in three level 1 overview displays (Calm Water, Heat Map, Visual Thesaurus) within the context of process control monitoring with expert operators to identify the effective aspects of each display.

Generally, the Calm Water display led to better operator detection accuracy, response time, and workload scores; however, the results were not definitive. The Heat Map display showed strong performance on the response time measure but was not different from the Calm Water display. The Visual Thesaurus display had the worst accuracy and response time performance but showed promise with a learning effect from trial 1 to trial 2 on the operator awareness measure, where the other displays did not show any improvement.

Operator detection accuracy was consistently better with the Calm Water display and response time was the worst with the Visual Thesaurus display. The implementation of buttons on this display likely led to the better performance measures – something that could easily be implemented into the other display types. Locating the buttons at the top gave a consistent and dedicated place for operators to look and making them bold when there was a process deviation provided simple visual indication of which button they need to press to complete their task – acknowledge the deviation. These same attributes were also likely accountable for the improved response time performance over the Visual Thesaurus display. Again, operators were likely easily able to locate the appropriate buttons whenever there were process deviations. Additionally, they were likely accountable for the Calm Water display's better workload performance on the System Monitoring portion of the MAT-B II – task switching was aided by consistent button placement at the top and the visual indication when deviations occurred, allowing operators to spend more resources on the secondary task. Button placement regularity also likely led to the Heat Map display's better response time performance over Visual Thesaurus as well. The buttons on the Heat Map display were located near the gauges and did not turn bold which is very similar to how they were implemented on the Visual Thesaurus display; but unlike the Visual Thesaurus display, the buttons on the Heat Map were always directly to the right of the relevant gauges. This provided consistency in where to find the button for a process deviation and likely led to decreased response time.

These findings align with the qualitative feedback gathered following the sessions. Operators indicated that they liked the Calm Water display and noted that they would use it if it were available to them. They felt confident that they would be able to master that display with minimal training. Additionally, they commented on how the features of the Calm Water display worked well for them. These comments reflect some of the performance indicated by the analysis – even though there was not a clearly better performing display, the Calm Water display yielded better performance in operator detection accuracy compared to any of the other displays on any of the other measures. Interestingly, the operators also indicated that the Visual Thesaurus display appeared to be easier to use than the other displays. This display did not perform well from an operator performance perspective, but this preference might have been due to the graphical nature of the display (graphical gauges), making it appear easier to use. A note here is that the Visual Thesaurus display did yield improved plant awareness during the second experiential trial. They indicated the Heat Map as their least preferred display overall.

Limitations, Future Research, and Conclusions

One of the main limitations of the current study lies in the coupling of the individual displays with an additional overview display. The current design does not allow to differentiate whether the experimental displays on their own would have led to the same findings in the form of the operator performance measures as in conjunction with the schematic display. The reason for incorporating the additional schematic overview display was that one of the 3 experimental displays (Calm Water) had been paired with the schematic display in actual use. Due to this, project stakeholders requested that all sessions be conducted using both overview level 1 displays. During training, operators were instructed to place equal importance to all displays during the sessions (schematic level 1 overview display, experimental level 1 overview display, and MAT-B II secondary task), but we were unable to validate whether they treated the displays as equally important. It is likely that the schematic display had an impact on the results. An eye tracker was implemented to account for this, but data for that measure was unable to be analyzed due to issues with the data logs generated by the eye tracking software. If we had this data available, we would be able to isolate performance data to instances when the operator was watching the experimental display. For example, we would be able to know whether the operator was responding to alarms on the experimental display or the schematic display. Additionally, the data would reveal potential display preferences between the displays in the form of ratio of time spent watching each display. Future research could ensure successful implementation of an eye tracker to track operators' visual attention or ensure that the experimental level 1 overview display could be used on its own so that the use of a schematic display would be unnecessary. In addition, future tests could be run evaluating the display types without an accompanying additional overview display.

The scenarios and measures used in the study did not necessarily examine all parts of the designs. For example, the scenarios that were used did not require operators to use the historical trend information contained in the Heat Map display – which is only easily accessible in this display. This aspect of the Heat Map display could potentially provide operators with very useful information about current and future states, supporting skill and rule-based behavior. Researchers could have used additional situation awareness level 3 (projection) probes to evaluate whether the displays help operators make projections into the future, relevant to the process scenario. Using this type of situation awareness probe would have provided some indication whether the historical information was useful for operators, as they would have been able to use that historical information to understand the trends in the process and make predictions using that information. Another aspect of the Heat Map display that might be valuable to operators, but was not evaluated, is the use of distinct colors to indicate alarms above normal operation (high alarms – red) and alarms below normal operation (low alarms – blue). Further investigation can be done to understand whether using distinct colors to indicate high vs. low conditions would lead to better operator performance.

Another area where the scenarios used in the current study might have been limiting to the current research questions is the implementation of the event rate used in the scenarios.

Each experimental trial consisted of 24 to 47 process deviations (events) over the span of 9 minutes. The amount of events depended on the scenario complexity (low vs. high). At 160 to 313 events per hour, these event rates are much higher compared to what an operator would typically be exposed to in the real world. Mattiasson (1999) found that abnormal situations, or situations that are unforeseen and therefore not supported by current procedures, could occur 3 times per hour during a normal week. This number increases to 53 events per hour during upset conditions. This increased event rate was useful to help researchers collect response rates over a short period of time (9 minute experimental trials), but may not be ecologically valid and therefore might have let to vigilance levels more conservative than would be typical for control room operators. Parasuraman and Giambra (1991) noted that vigilance can decline when event rates are increased during an experimental session designed with increased event rates compared to the real world. Specifically, higher event rates lead to decreased responses by participants. More conservative levels of criterion due to increased error rates may lead to decreases in responses (Sarter, Givens, & Bruno, 2001). This means that operators in the current study may have had artificially low response rates, leading to the low accuracy rates reported in the analysis (roughly 20-40% detection accuracy). In fact, it has been noted that the hit rate can decrease by two-thirds (from about 90% to about 30%) when event rates increase from 5 events per minute to 30 events per minute (Jerison & Pickett, 1964). This increase in events generally corresponds to the event rate used in the current study (up to 313 per hour) compared to the event rate I the real world (up to 53 per hour). Based on this, it would be expected that, under normal conditions, the 20-40% accuracy rates reported would rise to levels closer to 60-80%. Additional research should be done to understand whether vigilance (and therefore response rates) would shift with a more realistic event rate with each of the displays. To do this, scenario event rates should be adjusted to reflect real world event rates of 3 times per hour during normal operations and 53 times per hour during upset conditions.

Mouse click data was collected in the form of accuracy and response time such that events were recorded when a response was correct (hits), whether it was missed (misses), and what the response time was for that event. This could have been expanded to data for all mouse clicks on the display to reveal measures such as false detections or correct rejections in addition to hits and misses. This would have allowed us to use signal detection theory to analyze the data. This more rigorous statistical analysis method, teamed with a better experimental design would likely lead to stronger statistical evidence for operator performance due to SRK design implementation.

One of the greatest challenges in the current research was managing expectations across the three independent design teams in terms of what they want out of the research and what design their display will contain. Regarding display color, each display followed a unique color scheme. This becomes particularly troublesome when the experimental display is paired with the schematic display, which has yet another color scheme. For example, the schematic display uses yellow to indicate a high alarm state. The Calm Water and Heap Map displays use red to provide the same indication. Additionally, the Visual Thesaurus uses a light blue for the same. The use of colors in general should be evaluated in future studies. Having greater control over this would have allowed the research team to correct some of the limitations mentioned above as well as ensure consistent design implementation across all three displays. It would have also provided additional freedom in the experimental design so that the researchers could ensure a sound experimental design.

Based on the findings above, we could theorize a display which incorporates various elements of SRK in the design to aid in the three areas of operator performance (accuracy, workload, and awareness). For example, making buttons turn bold when process deviations occur could lead to better detection accuracy by highlighting the correct input for operators. Additionally, the use of icons to highlight process deviations for the experimental displays could also be explored. Ensuring consistent placement of buttons in relation to gauges could lead to better response time by allowing operators to easily locate the button for each gauge, regardless of where in the plant it is located. Highly salient visual peaks and valleys in gauge levels could potentially aid in workload and operators' ability to multitask by providing visual indications that can easily capture attention, even if it is currently placed on a secondary task. This needs to be explored further with the use of an eye tracker or some other mechanism to isolate which part of the display operators are focusing on. Even though it is unlikely, operators might not have paid attention to the visual peaks and valleys in the calm water display and solely relied on the bold buttons for indication. The inclusion of valid eye-tracking data would help in disambiguating this situation. Finally, providing graphical indications of the gauges could help aid in learning the display and understanding overall plant status. Future research could focus on isolating and testing each of these elements to investigate any potential effects on operator performance.

References

Bullemer, P., Hadjukiewicz, J., & Reising, D. V. (2002). *Effective operator display design.* ASM Consortium Guidelines (Version 4.00). Minneapolis, MN: ASM Consortium.

Bullemer, P., Reising, D.V., Burns, C., Hajdukiewicz, J. & Andrzejewski, J. (2008). *Effective operator display design*. ASM Consortium Guidelines (Version 4.00). Minneapolis, MN: ASM Consortium.

Comstock, J. R., & Arnegard, R. J. (1992). *The multi-attribute task battery for human operator workload and strategic behavior research* (NASA Technical Memorandum No. 104174). Hampton, VA: National Aeronautics and Space Administration, Langley Research Center.

Dinadis, N. (2004). *Advanced process graphics for operator effectiveness*. Presentation made at the 2004 Honeywell Asia Pacific Users Group.

Effken, J. A., Kim, N. G., & Shaw, R. E. (1997). Makingthe constraints visible: Testing the ecological approach to interface design. *Ergonomics , 40*, 1-27.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd Annual Meeting.* Santa Monica, CA: Human Factors Society, 97-101.

Hart, S. G. & Staveland, L. E. (1989). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Handcock & N. Meshkati (Eds.), *Human mental workload*. Amsterdam: North Holland.

Jamieson, G., Hajdukiewicz, J., & Reising, D. V. (2002). *EID design rationale case study* (Version 1.03). ASM Consortium Technical Report. Minneapolis, MN: ASM Consortium.

Jerison, H. I, & Pickett, R. M. (1964). Vigilance: The importance of the elicited observing rate. *Science*, *143*,970-971.

Jou, Y.-T., Yenn, T.-C., Lin, C. J., & Chiang, C.-C. (2009). Evaluation of operators' mental workload of human–system interface automation in the advanced nuclear power plants . *Nuclear Engineering and Design , 239*, 2537–2542.

Mattiasson, C. (1999). The alarm system from the operator's perspective. Human Interfaces in Control Rooms, Cockpits and Command Centres, 1999. *International Conference,* 217-221.

Parasuraman, R., & Giambra, L. (1991). Skill development in vigilance: Effects of event rate and age. *Psychology and Aging, 6*, 155-169.

Parasuraman, R., Hancock, P. A., & Olofinboba, O. (1997). Alarm effectiveness in driver-centered collision-warning systems. *Ergonomics, 40* (3), 390-399.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans , 30* (3), 286-297.

Pew, R. W. (2000). The state of situation awareness measurement: Heading toward the next century. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 33-47). Mahwah, NJ: LEA, Inc.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics , SMC-13*, 257-266.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Transactions on Systems, Man, and Cybernetics , SMC-15*, 234-243.

Rasmussen, J., & Vicente, K. J. (1989). Coping with human errors through system design: Implications for ecological interface design. *International Journal of Man-Machine Studies , 31*, 517-534.

Reising, D., & Sanderson, P. M. (2002). Ecological Interface Design for Pasteurizer II: A Process Description of Semantic Mapping. *Human Factors , 44* (2), 222.

Reising, D., Errington, J., Bullemer, P., & McLain, R. (2006). *2006 Visual Thesaurus Project.* Minneapolis, MN: ASM Consortium Technical Report (Version 2.0).

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: Where top–down meets bottom–up. *Brain Research and Brain Research Review*, *35*, 146– 160.

Sharp, T. D., & Helmicki, A. J. (1998). The application of the ecological interface design approach to neonatal intensive care medicine. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *42*, pp. 350-354.

Swets, J. A. (1992). The science of choosing the right decision threshold in high stakes diagnostics. *American Psychologist , 47* (4), 522-532.

Tharanathan, A., Bullemer, P., Laberge, J., Reising, D., & Mclain, R. (2010). Functional versus schematic overview displays: Impact on operator situation awareness in process monitoring. *Proceedings of the 54th Annual Meeting of the Human Factors and Ergonomics Society.* San Francisco.

Vicente, K. J. (1999). *Cognitive work analysis: Towards safe, productive, and healthy computer based work.* Mahwah, NJ: Erlbaum.

Vicente, K. J. (2002). Ecological Interface Design: Progress and Challenges. *Human Factors, 44* (1), 62-78.

Vicente, K. J. (2007). Monitoring a nuclear power plant. *Attention; From Theory to Practice*.

Woods, D. D. (2009). Rasmussen's SRK 30 Years Later: Is Human Factors Best in 3's? *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* , *53*, pp. 217-221.

Woods, D. D. (1995). The alarm problem and direct attention in dynamic fault management. *Ergonomics* , 38(11), 2371-2393.

Xiao, Y., & Seagull, F.J. (1999). An analysis of problems with auditory alarms: Defining the roles of alarms in process monitoring tasks. *Proceedings of the Human Factors and Ergonomic Society 43rd Annual Meeting* (Santa Monica, CA: Human Factors and Ergonomics Society), 256-260.

Appendix A: Protocol Approval: Research Approval Letter

ASM

20 March 2013

Department of Home Affairs
Republic of South Africa

To Whom It May Concern:

I hereby confirm that Mr. Joseph Carlos Vargas (Date of Birth – 22nd February, 1987; Passport no - [redacted] Expiration Date of Passport – 23 September 2022) will need to travel to South Africa, with the port of entry as O.R. Tambo International Airport. He needs to travel to South Africa to conduct a human machine interface study comparing various forms of Overview Displays at Sasol Synfuels in Secunda as a part of the Abnormal Situation Management Consortium work. The study will start on May 6 and will be completed by May 18, 2013.

Mr. Joseph Carlos Vargas is currently employed as a Research Engineer with Honeywell International, and Honeywell International leads the ASM Consortium. Sasol Synfuels is a member of the Abnormal Situation Management Consortium.

Please do not hesitate to call me if you have any questions.

Thank you for your support and willingness to help with this matter.

Dr. Anand Tharanathan
Principal Investigator
Abnormal Situation Management Consortium
Honeywell International

Direct telephone	001 763-954-4181
Direct facsimile	001 763-954-5436
Email	anand.tharanathan@honeywell.com

Appendix B: Consent Form

Informed Consent Form

Sasol is a member of the Abnormal Situation Management (ASM®) Consortium. Your plant management has agreed to conduct a study to evaluate the effectiveness of different designs for span-of-control overview displays aimed at helping you identify and respond to abnormal situations. By participating in the study, you will help us identify what aspects of the designs work for you and potential improvements for the displays.

After you review and sign this form you will be asked to complete a demographic questionnaire, some training modules, and answer some questions about the training material. Next, you will be given a chance to review the overview displays during a practice scenario. Afterwards, you will review several scenarios and your task will be to maintain awareness of the situation while identifying abnormal situations and alarm states. At random times during each scenario, we will pause the scenario and ask you some questions about the current status of the plant. Finally, we will thank you for your time and answer any questions that you have about the displays or the usability study.

To collect the necessary information, we will be recording eye-tracker data. In addition, we will record all of your mouse clicks automatically as well as the responses to the questions which will be presented on the computer screens. All the data will be kept confidential.

By signing this form, you agree that you will allow us to use the data we will collect from you, and any verbal responses, for the purposes of evaluating the three alternative overview displays. The data we collect will be kept confidential, and your recorded answers will not be identified by your name in anyway. Please keep in mind that we are evaluating the overview displays and not you. We will provide Sasol with a report of the results of this display evaluation, and that report will be shared with the other ASM member companies.

Your participation is voluntary. You may withdraw from this study at any time, for any reason. However, the data collected to that point may still be used for evaluation purposes.

If you have any questions, please ask now or at any time during the study.

If you agree with these terms, please indicate your acceptance by signing below.

Signature: ______________________________

Printed Name: ______________________________

Date: ______________________________

Appendix C: Experiment Instructions (Protocol)

NOTE 1: START all computers and ensure data collection folders are empty (displays and MATB).

NOTE 2: Note the operator number in the operator assignment matrix. Record the operator number on all demographic questionnaires, consent forms, response forms, etc.

NOTE 3: Open the corresponding training pdf on the right monitor. Copy/Paste the correct training scenario for the displays on the left monitors and open NetBeans.

[Describe who the research sponsor is]

Sasol is a member of Honeywell's ASM® Consortium. ASM stands for Abnormal Situation Management. The Consortium does research on better solutions that help operators prevent and respond to abnormal plant conditions. The Consortium has funded a project to investigate the impact of different design approaches for span-of-control overview displays. We have obtained permission from your management to conduct this preliminary study.

[Describe study objective]

In this study, we will be evaluating different overview display designs in terms of their ability to support operator awareness of the current situation. One display is a P&ID-based schematic display and the other is a new display that provides the same information as the P&ID display, but in a different format.

[Describe the overview of the project scope]

This study will last approximately 2.5 hours. This time will consist of the following:

First, you will review and sign an informed consent form that provides a summary of the study. The most important part of the consent form is that it explains that your participation in this study is voluntary and you can stop at any time, for any reason, and not be penalized.

Second, you will complete a short demographic questionnaire so we can get a sense of your background and experience, which may be useful when we analyze the evaluation data later.

Third, you will review some training material to get introduced to the displays and tasks that you will be completing.

Fifth, we will ask you some questions about what you learned from the training material on the overview displays.

Sixth, you will complete a practice scenario so that you can become familiar with the study protocol.

Seventh, you will spend the remainder of the session completing four scenarios that show real process data in the two alternative overview displays.

Finally, you will complete a post-session questionnaire and we will ask you a few debriefing questions. Before you leave, we will answer any questions that you have about the study or displays.

[Distribute informed consent form and have operator sign it]

As we mentioned, participation in this study is voluntary and you can stop at any time, for any reason, and not be penalized.

Note: Ask the operators whether their management asked them to participate in this study. Even if they were asked to participate in the study, tell them that they can stop at any time, for any reason, and we will not disclose that to your management.

The data that we collect will be kept confidential, and your recorded answers will not be identified by your name in any way. Please keep in mind that we are evaluating the overview display designs and not you. We will provide Sasol with a report of the results of this display evaluation, and that report will be shared with the other ASM member companies.

If you have any questions, please ask now or at any time during the study.

Please indicate your willingness to participate in this study by reviewing and signing the informed consent form.

[Distribute demographic questionnaire]

Please complete this demographic questionnaire so we can get a better sense of your experience and background. This information will not be used to evaluate you personally but could help us understand how using the different overview display designs could relate to operator experience or familiarity with the unit.

[Set expectation for experimental session]

During this study, you will be monitoring two Level 1 Overview displays and indicating gauge state changes by interacting with the displays. Level 1 overview displays are intended to enhance operator awareness by providing a 'big picture' view for detecting process deviations in the key process variables that characterize the health of each major equipment area. In addition to monitoring the Level 1 overview displays, you will also be monitoring a generic process task where you will be required make adjustments to maintain a healthy system.

Each scenario lasts approximately 10 minutes. During each scenario, we will pause (freeze) the scenario twice at random times and present you with questions on the computer screen. You will be provided with several multiple-choice questions to gauge your awareness of the situation. Please try to respond as accurately as you can. We will discuss these questions in detail later in the training. Also, please do not make any notes while doing the tasks.

At the end of each scenario, we will ask you about the workload you experienced during the scenario. We will review workload questions during the practice session.

Keep in mind that this study is focused on identifying the effectiveness of the Level 1 overview display designs – **the study does not focus on an individual operator's performance**. Instead, we want to evaluate how effective the displays are for the different scenarios you completed. All of the results will be kept anonymous and will be averaged across all operators, to protect each individual's identity.

[Start training module]

Because you are not familiar with either of the overview display designs, we want you to complete some training prior to participating in the rest of the study. The training material consists of a number of PowerPoint slides to illustrate some of the key concepts.

**// [CONDUCT APPROPRIATE TRAINING SCRIPT] **

Master Training Script

<training with active simulation, using PowerPoint as reference only>

[ ] (prepare: open Netbeans on the left monitor)

[ ] (prepare: copy and paste the appropriate wasp.dat file into program folder - PRACTICE/Visual Thesaurus/Heatmap/Calmwater)

[ ] (prepare: open appropriate training slides on the right AND left monitors)

Experiment Setup (slide 3)

The experiment contains two simulations which require continuous monitoring by you.

On the left two screens, you will be monitoring a crude oil process. The top-left screen will contain a specific type of display and the bottom-left screen will contain another type. However, both of the screens on the left will essentially show the same underlying process. The two displays on the left screens may present the process in different manners, and both are designed as high-level overview displays.

You will use the left mouse to interact with the two left screens.

On the right screen, you will be monitoring a completely different process. This process is unrelated to the process presented on the left screens but is equally important and should be given the same priority as the process on the left.

You will use the right mouse to interact with the right screen.

First, we will review the process and task on the left (point), and then discuss the one on the right (point)

Do you have any questions?

Process Model Overview (slide 4)

The crude oil process model used in the simulation contains two processes, the atmospheric unit and the vacuum unit. The atmospheric unit contains three main functional areas: the crude feed, the crude heater, and the crude tower. The vacuum unit contains two main functional areas: the vacuum heater and the vacuum tower.

The diagram on the bottom of the slide indicates the direction of the process.

Do you have any questions?

Defining Gauge States (slide 5)

One of the tasks that you will be doing is a crude oil process task where you will be required to attend to the situation and click gauge buttons when you detect a SIGNIFICANT change in the gauge state. A significant change is when a gauge moves between normal, abnormal, and alarm states. For example, a gauge moves from normal to abnormal, normal to alarm, abnormal to normal, abnormal to alarm, alarm to normal , or alarm to abnormal.

In addition to this task of attending to the situation, you will have to do the task on the other computer. We will discuss this task in more detail later.

Please make sure that you give EQUAL IMPORTANCE to both the tasks.

Schematic Display (slide 6)

This is a piping-and-instrumentation diagram (P&ID). Shapes represent major equipment, process flow lines show direction, and process values are numeric indicators.

This Level 1 overview display will be present at all times during the experiment

This display will be paired with another Level 1 overview display and both displays will follow the same scenario.

You are tasked with monitoring these displays simultaneously and click the buttons to acknowledge changes in gauge states. However, you do not need to click the same gauge button for both displays.

The schematic display will have colored icons to indicate alarm states. There is no visual cue for abnormal states on the schematic display.

Normal Operation (slide 7)

This is how steady state (normal) operation will look. There are currently no alarms present.

Abnormal Operation (slide 8)

These are examples of alarms. They are indicated by the colored shapes. We will get back to these shortly.

Main Areas (slide 9)

These are the main functional areas and the process flow: the crude feed (point), the crude heater (point), the crude tower (point), the vacuum heater (point), and the vacuum tower (point).

Alarms (slide 10)

These are buttons which you are asked to click if you notice an alarm (point).

Buttons for Input (slide 11)

These are the three types of alarms you will see on the schematic display.
(explain the three different types)

Buttons for Input (slide 12)

These are Process Variable (PV)/ Operation Output Percentage (OP) Gauges.
PV = Indication of current value
OP = Controlled by operator, influences PV

Do you have any questions?

Visual Thesaurus Display (slide 13)

This is a display that uses functional areas to represent major equipment, with shapes for analog gauges in each area.

This Level 1 overview display will be present at all times during the experiment.

The Visual Thesaurus display will be paired with the schematic display and both displays will follow the same scenario.

You are tasked with monitoring these displays simultaneously and click the buttons to acknowledge changes in gauge states. However, you do not need to click the same gauge button for both displays.

The Visual Thesaurus display will have colored outlines for the gauge shapes to indicate either abnormal OR alarm states.

Normal Operation (slide 14)

This is how steady state (normal) operation will look. There are currently no alarms present.

Abnormal Operation (slide 15)

These are examples of abnormal states and alarm states. They are indicated by the colored outlines. We will get back to these shortly.

Main Areas (slide 16)

These are the main functional areas and the process flow: the crude feed (point), the crude heater (point), the crude tower (point), the vacuum heater (point), and the vacuum tower (point).

Alarms (slide 17)

These are the three types of alarms you will see on the lower display.
(explain the three different types)

Buttons for Input (slide 18)

These are buttons which you are asked to click if you notice an abnormal or alarm state (point).

Buttons Review (slide 19)

The first letter indicates the type of gauge: F=Flow, P=Pressure, L=Level, T=Temp.

The text after that is the name of the gauge, example Main Reflux.

The letters at the end (immediately to the right of each button) indicates whether it is a process value or an operator output % gauge. PV=Process Value, OP=Operator Output %.

Gauge Shapes (slide 20)

These are what the shapes mean.
(explain the different types)

Gauge Shapes (slide 21)

This is how the shapes are designed.
(explain the normal, abnormal, and alarm ranges)

Heatmap Display (slide 13)

This is a display that uses functional areas to represent major equipment, with a color heat-map of historical values for each gauge.

This Level 1 overview display will be present at all times during the experiment.

The Heatmap display will be paired with the schematic display and both displays will follow the same scenario.

You are tasked with monitoring these displays simultaneously and click the buttons to acknowledge changes in gauge states. However, you do not need to click the same gauge button for both displays.

The Heatmap display will have shades of BLUE and RED to indicate either abnormal OR alarm state.

Normal Operation (14)
This is how steady state (normal) operation will look. There are currently no alarms present.

Abnormal Operation (15)
These are examples of abnormal states and alarm states. They are indicated the shades of BLUE and RED. We will get back to these shortly.

Main Areas (16)
These are the main functional areas and the process flow: the crude feed (point), the crude heater (point), the crude tower (point), the vacuum heater (point), and the vacuum tower (point).

Gauge States (17)
This slide explains the way in which the current and historical states are represented for each gauge.
(go through the slide in detail)

Alarms (18)
The abnormal states AND alarms are indicated by the colors BLUE and RED
(explain the slide in detail)

Buttons for Input (19)
These are buttons which you are asked to click if you notice an abnormal or alarm state (point).

Buttons Review (20)
The first letter indicates the type of gauge: F=Flow, P=Pressure, L=Level, T=Temp.

The text after that is the name of the gauge, example Main Reflux.

The letters at the end (immediately to the right of each button) indicates whether it is a process value or an operator output % gauge. PV=Process Value, OP=Operator Output %.

Calmwater Display (slide 13)
This is a display that uses a matrix-like structure to link all gauges together into a single-object surface graph. The idea is that a steady-normal process will be presented as a calm-matrix ('calm-water'); abnormal events disturb this with peaks.

This Level 1 overview display will be present at all times during the experiment.

The Calmwater display will be paired with the schematic display and both displays will follow the same scenario.

You are tasked with monitoring these displays simultaneously and click the buttons to acknowledge changes in gauge states. However, you do not need to click the same gauge button for both displays.

The Calmwater display will have peaks/spikes to demonstrate abnormal and alarm states. Gauge name labels will appear in several colors to represent the state.

Basic Concept (14)
(explain matrix structure and point-of-view during simulation)

Normal Operation (15)
This is how steady state (normal) operation will look. There are currently no alarms present.

Abnormal Operation (16)
These are examples of abnormal states and alarm states. We will get back to these shortly.

Main Areas (17)
The buttons are grouped into the functional areas: Crude Feed, Crude Heater, Crude Tower, Vacuum Heater, and Vacuum Tower.

Layout (18)
The buttons on top represent the gauges immediately below on the graph (vertically).
(show slide and explain)

Layout (19)
(an example of previous slide when a tag turns ON)
When a tag is present, the text in the button on top which represents that tag becomes bold.

Alarms (20)
The abnormal states are indicated by a yellow/orange tag, the alarms are indicated by red tags.
The horizontal plane is an indicator if the status is a high or low abnormal/alarm.
(explain the slide in detail)

Buttons for Input (21)
These are buttons which you are asked to click if you notice an abnormal or alarm state (point).

Buttons Review (22)
The first letter indicates the type of gauge: F=Flow, P=Pressure, L=Level, T=Temp.

The text after that is the name of the gauge, example Main Reflux.

The letters at the end (immediately to the right of each button) indicates whether it is a process value or an operator output % gauge. PV=Process Value, OP=Operator Output %.

//[CONDUCT MAT-B TRAINING]

MATB Display (page 1)
[ ] (action: minimize Netbeans on left screens)
[ ] (action: close PowerPoint on right screen)
[ ] (action: open MATB PowerPoint on left screen)
[ ] (action: open MATB and choose Practice Scenario, and start the program)

This is a independent display which requires monitoring of a separate process, different than the crude oil process.

This display will be present at all times during the experiment.

The other two (crude oil) displays will be on the left screens throughout the experiment. You are tasked with monitoring all of these three displays simultaneously.

The MATB display will have several features which require your attention to maintain a healthy system.

Layout (page 2)
The areas that are not in use are these (point them out)

Areas (page 3)
There are two areas of interest to you. We will review them one at a time.

System Monitoring (page 4)
The first area is a System Monitoring Task
(read off the slide and explain the features within the task)

System Monitoring (page 5)
(show example of red light ON and how to resolve)

System Monitoring (page 6)
(show example of green light OFF and how to resolve)

System Monitoring (page 7)
(show example of abnormal scale and how to resolve)

[ ] (instruction: ask the subject to click the buttons and answer any questions)
1 minute max

Resource Management (page 8)
The second area is a Resource Management Task
(read off the slide and explain the features within the task - continued on next slide)

Resource Management (page 9)
(read off the slide and explain the features within the task)

Resource Management (page 10)
(example 1 abnormal scenario - use if needed)

Resource Management (page 11)
(example 2 abnormal scenario - use if needed)

[ ] (instruction: ask the subject to click the buttons and answer any questions)
1 minute max

//[CALIBRATE EYETRACKER]

[Start practice session – be sure to choose the appropriate display type]

NOTE 1: Use the assignment matrix to determine the correct scenario #

NOTE 2: Remember to continue calibrating the eyetracker at the start of each scenario

NOTE 3: Remember to adjust eyetracker ROIs during each scenario

NOTE 4: Remember to freeze MATB at the freeze time for each scenario

NOTE 5: Remember to copy/paste the output data file at end of each scenario

Before you complete the four scenarios, we want you to get some practice doing the tasks simultaneously.

During this practice training, if you do not understand the questions, please ask for clarification. If you don't understand something you see in the displays, please ask. Note however that we may not be able to answer all of your questions about the displays because we might bias your subsequent responses during the scenarios if we do so.

Please remember to keep your head still, as best as you can, during the scenario. You will be allowed to reposition in between each scenario, if necessary.

[Open NASA TLX on the right monitor]

Before we end this scenario, we would like you to fill another questionnaire that asks some questions about the total workload you experienced.

NOTE 1: Remember to copy/paste output files

[Start first experimental trial - be sure to choose the appropriate display type]

NOTE 1: Use the assignment matrix to determine the correct scenario #

NOTE 2: Remember to continue calibrating the eyetracker at the start of each scenario

NOTE 3: Remember to adjust eyetracker ROIs during each scenario

NOTE 4: Remember to freeze MATB at the freeze time for each scenario

NOTE 5: Remember to copy/paste the output data file at end of each scenario

[Open NASA TLX on the right monitor]

Before we end this scenario, we would like you to fill another questionnaire that asks some questions about the total workload you experienced.

NOTE 1: Remember to copy/paste output files

[Start second experimental trial - be sure to choose the appropriate display type]

NOTE 1: Use the assignment matrix to determine the correct scenario #

NOTE 2: Remember to continue calibrating the eyetracker at the start of each scenario

NOTE 3: Remember to adjust eyetracker ROIs during each scenario

NOTE 4: Remember to freeze MATB at the freeze time for each scenario

NOTE 5: Remember to copy/paste the output data file at end of each scenario

[Open NASA TLX on the right monitor]

Before we end this scenario, we would like you to fill another questionnaire that asks some questions about the total workload you experienced.

NOTE 1: Remember to copy/paste output files

[Start third experimental trial - be sure to choose the appropriate display type]

NOTE 1: Use the assignment matrix to determine the correct scenario #

NOTE 2: Remember to continue calibrating the eyetracker at the start of each scenario

NOTE 3: Remember to adjust eyetracker ROIs during each scenario

NOTE 4: Remember to freeze MATB at the freeze time for each scenario

NOTE 5: Remember to copy/paste the output data file at end of each scenario

[Open NASA TLX on the right monitor]

Before we end this scenario, we would like you to fill another questionnaire that asks some questions about the total workload you experienced.

NOTE 1: Remember to copy/paste output files

[Start fourth experimental trial - be sure to choose the appropriate display type]

NOTE 1: Use the assignment matrix to determine the correct scenario #

NOTE 2: Remember to continue calibrating the eyetracker at the start of each scenario

NOTE 3: Remember to adjust eyetracker ROIs during each scenario

NOTE 4: Remember to freeze MATB at the freeze time for each scenario

NOTE 5: Remember to copy/paste the output data file at end of each scenario

[Open NASA TLX on the right monitor]

Before we end this scenario, we would like you to fill another questionnaire that asks some questions about the total workload you experienced.

NOTE 1: Remember to copy/paste output files

[Debrief and thank operator]

DEBRIEF QUESTIONS

For us to get accurate information about the effectiveness of the displays, it is important that you do not discuss the displays or scenarios with other operators who may also be participating in the study. If they have advanced knowledge of what we are doing it may reduce the usefulness of the study results.

Thank you for taking the time to help us in this study.

Appendix D: Experiment Debriefing Form

OPERATOR#__________ DATE: ____________ CONDITION: _____________

For each item in the table below, please place a check mark (✓) in the appropriate box to indicate the degree of agreement.

		Strongly Disagree			**Neither**			**Strongly Agree**
	Question							
1	I think I would use this overview display frequently							
2	I found this overview display to be simple							
3	I thought the overview display was easy to use							
4	I think that I would not need the support of a technical person to use this overview display							
5	I found that the various features in this overview display worked well							
6	I thought this overview display behaved predictably							
7	I imagine that most operators would learn to use this overview display very quickly							
8	I found these graphical objects very easy to use							
9	I felt very confident using this overview display							
10	I needed to learn minimal things before I could get started with this overview display							
11	This overview display helped me with my situation awareness							
12	This overview display captured the critical variables							
13	I had trouble with the overall layout of the equipment areas in this display							
14	I found the use of colours in this display to be useful							

For each numbered item below, please indicate your response by placing a check mark (✓) in the appropriate box.

1. Indicate how realistic you thought the scenarios were today.

   ☐ Very Unrealistic
   ☐ Slightly Unrealistic
   ☐ Neutral
   ☐ Slightly Realistic
   ☐ Very Realistic

2. What could have been improved for the scenarios (please circle your response)?

   a. The rate by which disturbances moved to other areas should have been

      ☐ Slower
      ☐ Faster
      ☐ Not applicable

   b. The number of variables affected should have been

      ☐ Less
      ☐ More
      ☐ Not applicable

   c. The magnitude of the deviations was

      ☐ Too little
      ☐ Too much
      ☐Not applicable

3. Have you had any prior experience with the overview display that you used today?

   ☐ Yes (please elaborate below)
   ☐ No

____
____
____
____

4. Do you have any general comments of feedback?

   ☐ Yes (please elaborate below)
   ☐ No

____
____
____
____

Appendix E: Situation Awareness Probe Questions

FIRST SET (FREEZE QUESTIONS)

Correct Answer	#	Question	A	B	C	D
B	1	Since the beginning of the scenario, ***desalter H2O pressure in the crude feed*** has:	Significantly increased	Significantly decreased	Stayed constant with no significant deviation from normal state	
A	2	Since the beginning of the scenario, ***controller output of the light crude flow in the crude feed*** has:	Significantly increased	Significantly decreased	Stayed constant with no significant deviation from normal state	
B	3	When the freeze happened, ***total pass flow in the crude heater*** was:	In abnormal state	In alarm state	In normal state	
D	5	When the freeze happened, which of the following was significantly increasing?	Total pass flow in the crude heater area	Kerosene product flow in the crude tower	GO product in the crude tower	None of the above
A	7	Since the beginning of the scenario, ***more than three AREAS (not just tags)*** have deviated from normal state.	TRUE	FALSE		
A	N1	When the freeze happened ***all tags in the vacuum tower*** were in a normal state.	TRUE	FALSE		
B	N2	When the freeze happened, ***total feed in the crude feed*** was in a normal state.	TRUE	FALSE		
A	N3	When the freeze happened, ***total pass flow in the vacuum heater*** was in an abnormal state.	TRUE	FALSE		

SECOND SET (END QUESTIONS)

Correct Answer	#	Question	A	B	C	D
C	1	Since the first freeze, ***controller output of the fuel gas in the crude heater*** has:	Significantly increased	Significantly decreased	Stayed constant with no significant deviation from normal state	
B	3	When the freeze happened, ***total pass flow in the vacuum heater*** was:	In abnormal state	In alarm state	In normal state	
A	5	When the freeze happened what amongst the following was true for the ***vacuum heater***?	More than two abnormal states and one alarm	Two abnormal states and one alarm	More than three abnormal states and no alarm	More than three abnormal states and two alarms
A	6	When the freeze happened, which of the following was significantly decreasing?	Air flow in the vacuum heater	Controller output of the slopwax tray level in the vacuum tower	Controller output of the fuel gas in the crude heater	None of the above
B	7	When the freeze happened, ***temperature of the VRFO product in the vacuum tower*** was in an abnormal state.	TRUE	FALSE		
C	N1	When the freeze happened, ***temperature of the VRFO product in the vacuum tower*** was:	In normal state	In abnormal state	In alarm state	
B	N2	When the freeze happened ***all tags in the crude feed*** were in a normal state.	TRUE	FALSE		
A	N3	When the freeze happened, ***VRFO product flow in the vacuum tower*** was in an abnormal state.	TRUE	FALSE		

Appendix F: NASA TLX Subjective Workload Probe

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High