2014

Using eye-tracking to investigate strategy and performance of expert and novice control room operators

Christina Marie Koffskey

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses

Part of the [Mechanical Engineering Commons](#)

Recommended Citation

Koffskey, Christina Marie, "Using eye-tracking to investigate strategy and performance of expert and novice control room operators" (2014). *LSU Master's Theses*. 3353.

https://digitalcommons.lsu.edu/gradschool_theses/3353

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

USING EYE-TRACKING TO INVESTIGATE STRATEGY AND PERFORMANCE OF
EXPERT AND NOVICE CONTROL ROOM OPERATORS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment
of the requirements for the degree of
Master of Science

in

The Department of Mechanical and Industrial Engineering

by
Christina M. Koffskey
B.S., Louisiana State University, 2010
May 2014

ACKNOWLEDGMENTS

First and foremost I would like to sincerely thank Dr. Laura Ikuma for her guidance, advice, patience and enthusiastic encouragement. I would also like to thank my other committee members, Dr. Fereydoun Aghazadeh and Dr. Craig Harvey for their valuable time and feedback. This research would not have been possible without funding from the Center for Operator Performance and generous technical assistance and support from Mark Nixon, Emerson Process Management.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF TABLES	iv
LIST OF FIGURES	v
GLOSSARY OF ABBREVIATIONS	vi
ABSTRACT.....	vii
CHAPTER 1. INTRODUCTION	1
CHAPTER 2. REVIEW OF THE LITERATURE	4
2.1 Introduction.....	4
2.2 Insights into Cognitive Processes Using Eye-Tracking	4
2.3 Expertise: What It Means to be an Expert	5
2.4 Eye-Tracking to Improve Training and Interface Design	6
2.5 Using Eye-Tracking to Investigate Expert Strategy.....	7
2.6 Basic Eye-Tracking Measurements	8
2.7 Eye-Tracking in Aviation and Health Care.....	8
2.7.1 Number of Fixations per AOI	9
2.7.2 Gaze Percentage per AOI.....	9
2.7.3 Gaze Duration Mean per AOI.....	10
2.7.4 Scan Path.....	11
2.8 Variability of Eye-Tracking Metrics.....	12
CHAPTER 3. METHODS	14
3.1 Background on Previous Research	14
3.1.1 Simulation and Interface Design.....	14
3.2 Experiment Overview and Design	15
3.3 Participants.....	17
3.4 Eye-Tracking.....	19
3.5 Experimental Procedure (Previous Research).....	19
3.5.1 How to Address Failures Within Scenarios	20
3.6 Experimental Procedure (Current Research)	21
3.7 Obtaining Eye Metrics through BeGaze Software.....	22
3.7.1 Fixation Frequency per AOI	22
3.7.2 Gaze Duration Mean per AOI.....	23
3.7.3 Gaze Percentage per AOI.....	23

3.7.4 Scan Path.....	23
3.8 Areas of Interest.....	24
3.9 AOI Mapping.....	26
3.10 Alarm Activity	27
3.11 Performance	27
3.12 Hypothesis.....	28
3.13 Data Analysis	29
CHAPTER 4. RESULTS	32
4.1 Quantitative Analysis.....	32
4.2 Scan Plot Qualitative Analysis.....	46
CHAPTER 5. DISCUSSION.....	57
5.1 Model 1	57
5.2 Model 2	59
5.3 Correlation	59
5.4 Setpoint Change.....	63
5.5 Future Research and Limitations	67
CHAPTER 6. CONCLUSION.....	69
6.1 Expertise	69
6.2 Interface Design	69
REFERENCES	71
APPENDIX 1. GREY AND BLACK DISPLAYS.....	75
APPENDIX 2. AOIS	77
APPENDIX 3. SEMANTIC GAZE MAPPING.....	78
APPENDIX 4. AOI EDITOR	79
VITA.....	80

LIST OF TABLES

Table 1 - Eye-Tracking Metrics and What They Measure Based on Task Type	13
Table 2 - Independent and Dependent Variables for Previous and Current Research	15
Table 3 - Comparison of the Design Changes for the Grey and Black Interface.....	15
Table 4 - Experiment Design and Predictor Variables.....	16
Table 5 - Demographic Information Collected for Novice Students	18
Table 6 - Demographic Information Collected for Expert Control Room Operators	18
Table 7 - AOI List for Active and Inactive Periods	26
Table 8 - Sample Output for Event Statistics.....	26
Table 9 - Independent T-Test between Thesis Participants and Entire Sample Size	28
Table 10 - Independent and Dependent Variables Analyzed Within Each Model	30
Table 11 - Descriptive Statistics for Expert and Novice AOI Fixation Details	32
Table 12 - Model 1: MANOVA Main Effects Results	33
Table 13 - Model 2: MANOVA Main Effects Results	33
Table 14 - Single ANOVA Results for Model 1: Main Effects Due to Interface	34
Table 15 - Single ANOVA Results for Model 1: Main Effects Due to Expertise	35
Table 16 - Single ANOVA Results for Model 1: Main Effects Due to Alarm Activity	36
Table 17 - Single ANOVA Results for Model 2: Main Effects Due to Interface	38
Table 18 - Single ANOVA Results for Model 2: Main Effects Due to Expertise	39
Table 19 - Non-parametric Results: Interface Type.....	40
Table 20 - Non-parametric Results: Expertise.....	41
Table 21 - Non-parametric Results: Alarm Activity.....	42
Table 22 - Correlation Between Fixation Frequency and Gaze Duration.....	43
Table 23 - Correlation Between Gaze Percentage and Dependent Variables	44
Table 24 - Correlation Between Speed and Accuracy and Dependent Variables	45
Table 25 - Participant Elapsed Time to Open Faceplate and Correct Setpoint.....	64

LIST OF FIGURES

Figure 1 - Sequence and timestamps of failures for the hard workload level	17
Figure 2 - SMI eye-tracking glasses used during the experiment (http://eyetracking.tistory.com/14)	19
Figure 3 - Location of the setpoint failure	24
Figure 4 - Diagram illustrating how AOIs are used to include or exclude data for analysis	25
Figure 5 - Participant 1: 60 seconds after setpoint failure	47
Figure 6 - Participant 1: 30 seconds before and after opening FP	48
Figure 7 - Participant 2: 60 seconds after setpoint failure	48
Figure 8 - Participant 2: 30 seconds before and after opening FP	49
Figure 9 - Participant 3: 60 seconds after setpoint failure	50
Figure 10 - Participant 3: 30 seconds before and after opening FP	50
Figure 11 - Participant 4: 60 seconds after setpoint failure	51
Figure 12 - Participant 4: 30 seconds before and after opening FP	51
Figure 13 - Participant 5: 60 seconds after setpoint failure	52
Figure 14 - Participant 5: 30 seconds before and after opening FP	52
Figure 15 - Participant 6: 60 seconds after setpoint failure	53
Figure 16 - Participant 6: 30 seconds before and after opening FP	53
Figure 17 - Participant 7: 60 seconds after setpoint failure	54
Figure 18 - Participant 7: 30 seconds before and after opening FP	54
Figure 19 - Participant 8: 60 seconds after setpoint failure	55
Figure 20 - Participant 8: 30 seconds before / after opening FP	55
Figure 21 - Participant 9: 60 seconds after setpoint failure	56
Figure 22 - Participant 9: 30 seconds before and after opening FP	56
Figure 23 - Participant 3 scan path: 30 seconds before and after opening C-TIC-13807 faceplate.....	58

GLOSSARY OF ABBREVIATIONS

AOF: Alarm on Faceplate

AOI: Area of Interest

C-FIC-13803: Fuel Oil Controller

C-FIC-13804: Fuel Gas Controller

C-TIC-13807: Crude Temperature Controller

FP: Faceplate

MOF: Mode on Faceplate

MPDT: Mean Percentage Dwell Time

PHV: Process History View

PV: Process Variable

ROF: Reading on Faceplate

SCADA: Supervisory Control and Data Acquisition

SCLTC: South Central Louisiana Technical College

SMI ETG: SensoMotoric Instruments Eye Tracking Glasses

SP: Setpoint

TTC: Time to Complete

TTF: Time to Faceplate

ABSTRACT

There is lacking research within the Petrochemical Industry that uses eye-tracking to explore the differences between the strategies of expert and novice control room operators as they monitor and address process parameters that could be used to improve novice training programs and interface design. Scan paths and three eye-tracking metrics (Fixation Frequency, Gaze Duration Mean, and Gaze Percentage) were used to investigate the differences in eye behavior of three expert control room operators and six novice students as they monitored and corrected a Crude Refinement simulation.

A 2x2x2 mixed factor design was used to explore the effects that expertise (expert and novice), interface type (black and grey), and alarm activity (active and inactive) had on participant eye behavior specifically, fixation frequency, gaze duration mean, and gaze percentage for certain areas of interest. The display was separated into 6 different areas and each area resulted in distinct eye statistics. Scan paths were plotted surrounding a subtle setpoint change within the simulation and were qualitatively analyzed to reveal differences due to expertise, interface type and alarm activity.

The MANOVA revealed no significant differences due to expertise, interface type, and alarm activity. The single ANOVAs revealed that participants had higher fixation frequencies on the Main display during monitoring periods than during active periods revealing that both expert and novice participants' attention was more divided when there were failures and alarms present than when the process was running at normal conditions. Also, experts spent a larger percentage of time monitoring the critical crude temperature and flow controller than novices. Pearson's correlation between dependent variables revealed a positive correlation between fixation frequency and gaze percentage that indicated that participants typically had many, quick fixations rather than few, long fixations. Scan path analysis revealed that active monitoring and interface background color influenced how quickly operators discovered the setpoint change on screen. Overall, eye-tracking successfully detected differences between participants and interface types that can benefit novice training and display design.

CHAPTER 1. INTRODUCTION

The petrochemical industry experiences losses of up to \$10 billion dollars annually due to preventable work-place accidents. If not controlled and regulated properly, petrochemical processing can be volatile and may lead to catastrophic incidents, which is why safety is such a high priority. Good interface design is a crucial safety factor that contributes to accident prevention (Johannsen, 1997; Yung-Sen & Sheue-Ling, 1995). In critical situations, good interface designs allow operators to easily access and interpret necessary real-time information in a timely manner. By providing displays with appropriate process information, operators can make speedy and correct decisions to avoid damages and losses.

In several accident investigations, inadequate interface designs were identified as contributing factors to the accident. In 2005, a series of explosions took place during the start-up of a hydrocarbon isomerization unit at the BP Texas City refinery killing fifteen workers and injuring 180 more. The investigation determined that a poor computerized control board display contributed to the accident. The display did not provide the flow information in and out of the raffinate unit on the same screen. The failure to have both readings on the same screen impeded the operator from knowing how much liquid was in the unit and that he needed to divert some liquid to storage (U.S. Chemical Safety and Hazard Investigation Board, 2007). Another accident attributed in part to poor interface design occurred in 2008, at the Bayer CropScience facility in Institute, West Virginia, where a runaway chemical reaction caused a methomyl vessel to explode, killing 2 and injuring 8 people. After further investigation by the U.S. Chemical Safety and Hazard Investigation Board (2008), the introduction of a new control system and inadequate board operator training were contributing factors. The new system was supposed to decrease human-related error but operators found it difficult to familiarize themselves with the new display screens and different units of measure than the old system.

Efficient visual displays can contribute to safe operations. Poorly designed console displays have led to tragic accidents as was the case at the BP Texas City refinery. One goal to increase safety is aimed at increasing operator performance by improving the usability of the display/user interface. Interface

design strongly influences the ease of access to real-time information allowing operators to make accurate and timely decisions. General interface design research has considered the effects of color, presentation style and location of information (pictorial, numerical, graphical, etc.), as well as navigation on performance (Burns, 2000; D'Souza & Greenstein, 1999; Wu, Chen, Lee, & Chen, 2010). If the interface used to control and monitor the process is distracting and difficult to navigate operators are more likely to commit mistakes. When critical information is hard to find within the display or there are distracting items on-screen, mistakes or time delays are likely to occur. By considering these common features, designers can create displays that support high performance and ease of use.

Recently, eye-tracking has become a popular technique within Human Computer Interaction (HCI) research to develop improved visual display designs. Eye-tracking allows researchers to identify how users scan a display, locate areas that are frequently looked at versus areas that are often neglected, and detect instances when the user is confused by an object on-screen. Such knowledge would be very difficult to gain through other measurement tools such as observations and surveys (Cooke, 2005). By acquiring information using eye-tracking technology, designers can ensure pertinent information is visible and remove or redesign objects that are difficult to interpret. HCI researchers have also studied color appeal in website design (Cyr, Head, & Larios, 2010), user adaptability to website design (Tzanidou, Petre, Minocha, & Grayson, 2005), and the cognitive processing of visual information during reading and searching tasks (Mele & Federici, 2012). Information obtained from collecting and analyzing eye movement data can be applied to improve usability of visual displays.

The data recorded from tracking eye movements can be used to improve the usability of a system by identifying differences between expert and novice cognitive processing and describing an individual's attention allocation. It can also be used to study monitoring and search task strategies and as an indicator of performance (Poole & Ball, 2005). Useful information, such as techniques and strategy, can be attained from studying the eye movements of experts. How experts scan the display or what information experts are focusing on when certain events occur can aid in display design (Poole & Ball, 2005). By

knowing what information experts attend to during certain tasks, display designers can ensure such information is salient. Expert strategy can also be useful when training novice operators on how to interact with the display.

Although there is research that uses eye-tracking technology regarding cognitive processing of visual information for general displays, there is little research that uses eye-tracking to evaluate the cognitive processing of displays in SCADA systems. SCADA, Supervisory Control and Data Acquisition, systems are a type of control system that collects and displays data and allows users to manipulate and control the system from a distant location. The petrochemical industry uses SCADA systems to control the petroleum refining process from a centralized control room. Addressing this gap, the goal of this research is to use eye-tracking technology to investigate the strategy and behavior of expert and novice control room operators while managing a simulated petrochemical SCADA system. Using eye-tracking to explore the strategy of experts and novices may lead to insightful improvements to operator interfaces. By identifying their strategy, displays and presentation of information can be designed in a manner that supports high performance. The scope is limited to investigate the strategies of operators during the crude distillation process within petroleum refining using one SCADA system, DeltaV (Emerson Process Management, Round Rock, TX).

CHAPTER 2. REVIEW OF THE LITERATURE

2.1 Introduction

Research into usability and interface design is relevant to the petrochemical industry, which relies on interface consoles to control highly complex systems. The petrochemical industry uses SCADA displays primarily to monitor and control processes. In order for operators to monitor and control the process effectively, the interface needs to provide real-time information rapidly and in a form consistent with their mental model. Operator interfaces should be designed compatibly with the user's needs (Murphy & Mitchell, 1986). Understanding the cognitive mechanisms behind user behavior is instrumental when designing effective displays; eye-tracking is a tool that can aid in understanding such mechanisms by helping identify the "whys" behind operator behavior.

2.2 Insights into Cognitive Processes Using Eye-Tracking

Eye-tracking provides insight into the cognitive processing behind human behavior that is not available through observation and think-aloud practices (Cooke, 2005). Just and Carpenter (1976) researched eye fixation and simple cognitive tasks and concluded that what the eye is focused on is assumed to be the thought at the forefront of the mind during cognitive processing. This means that eye-tracking can be used to locate and track the user's attention (Poole & Ball, 2005). Several industries including aviation and healthcare have used eye-tracking as a way to know exactly what individuals are focused on and the path that the eye takes when looking from one point to the next. The aviation industry has used it to study visual scanning behavior during vigilance tasks (Lavine, Sibert, Gokturk, & Dickens, 2002), pilot monitoring strategies during automated flight simulation (Sarter, Mumaw, & Wickens, 2007), as well as expert aircraft inspection strategy to use as a standard when evaluating novice inspector's performance (Gramopadhye et al., 2002). The healthcare and medical industries have used eye-tracking to investigate attention allocation of surgeons (Tomizawa, Aoki, Suzuki, Matayoshi, & Yozu, 2012) and surgeon vigilance in the operating room (Zheng et al., 2011). In regard to expertise research, tracking the user's attention when investigating the strategies of experts and novices is

extremely useful since experts' improved decisions are attributable to their effective attention allocation strategies (Schriver, Morrow, Wickens, & Talleur, 2008). By recording where the experts' attention is focused, their valuable attention allocation strategy can be extracted and used to improve display designs and training programs as explained in section 2.4.

2.3 Expertise: What It Means to be an Expert

Webster's Dictionary defines expert as one "having, involving, or displaying special skill or knowledge derived from training or experience" (Merriam-Webster's online dictionary, n.d.), though researchers in expertise research do not agree on a specific definition of expertise. Different research areas define expertise differently. Cognitive science research views expertise in regard to cognitive functioning (i.e. memory, learning to problem solve and reasoning). Based on these cognitive measurements, cognitive science researchers find that expert performance is superior to that displayed by novices (Shanteau, 1992). In decision-making research, experts are judged mostly on accuracy and repeatability. When considering the accuracy and repeatability of expert judgments, studies show that experts are often ineffective decision makers with poor quality judgments (Shanteau, 1992). Studies on clinical psychologists (Oskamp, 1962) and court room judges (Ebbesen & Konecni, 1975), found these experts made inaccurate and unreliable decisions. Considering both cognitive sciences and decision-making sciences, it can be reasoned that just because experts have higher cognitive functioning does not mean they always make the best decisions. In general, the agreement is that experts are more skilled, competent, and think in qualitatively different ways than novices do within their specific domain (Cellier, Eyrolle, & Marine, 1997; Glaser & Chi, 1988; Shanteau, 1992).

Although cognitive science research cannot answer all questions regarding expertise, it does give insight into expert performance. Cognitive science investigates expert performance by looking at the cognitive abilities of experts assumed to be attained through an extensive period of training and development (Ericsson, 2005). Experts are able to perceive meaningful patterns within their domain, which does not imply that experts have a general superior perceptual ability but rather have a greater

organization of their knowledge base; they are able to recall things in meaningful clusters or chunks (Glaser & Chi, 1988). “Chunking” allows experts to perceive and store information in terms of chunks rather than at the lowest-level unit (Anderson, 1996).

In addition to the ability to “chunk” information, experts are overall faster than novices and solve problems with little error (Glaser & Chi, 1988). There is a large amount of literature that supports the claim of performance differences between experts and novices found predominantly in tasks that are knowledge based (Chase & Simon, 1973; Hampton, 1994; Vicente, 1992). Although experts are faster overall, they are slower in the initial stages of problem solving (Glaser & Chi, 1988). It is common for experts to spend more time analyzing a problem quantitatively trying to understand and retain all facets of the problem (Glaser & Chi, 1988). On the other hand, novices tend to hastily start solving problems without analyzing them (Glaser & Chi, 1988).

Shanteau (1992) describes five different characteristics observed in expert decision strategies. First, experts are more disposed to making adjustments to initial decisions whereas novices tend to blindly commit to initial choices. Second, experts rarely work in isolation and instead work in groups or seek help from others to make better decisions. Operation in groups increases accuracy and confidence of decisions (Sniezek & Henry, 1990). Third, experts often use decision aids to avoid biasing decisions. The fourth characteristic of expert strategies is the effort to avoid making large mistakes. Experts initially make an estimate then make minor adjustments after a more careful analysis. Lastly, experts often operate by breaking larger problems into smaller, more manageable pieces.

2.4 Eye-Tracking to Improve Training and Interface Design

Identification of cognitive mechanisms and strategies of experts has strong implications for novice training. One of the motivations behind expertise research is to improve the training of the less skilled to behave and think like experts (Ericsson, 2005; Klein, 2008). Expert strategies are desirable as they correlate with superior performance. Bellenkes, Wickens, and Kramer (1997) observed that expert

and novice aviation pilots utilized different scanning techniques when flying and found higher performance ratings (measured by reduced flight path error and faster reaction times) were correlated with scanning techniques of expert pilots over novices. This research supports training novices to think and act like experts to increase novice performance. Likewise, Gopher et al. (1994) demonstrated this approach to training in a study on flight cadets' video game strategy. Cadets were trained in the game task to use strategies based on analysis of expert behavior. The number of errors reported by flight cadets trained to apply the same type of strategy used by expert gamers decreased compared to cadets not trained. The use of eye-tracking to investigate expert strategy can be highly advantageous for training programs for novice individuals with the possibility of expediting the novice to expert process.

Another advantage to using eye-tracking to investigate expert and novice strategy is its implications for improved usability and design of interfaces. Interface deficiencies can be identified when there is a large gap between expert and novice performance on a particular task (MacDorman, Whalen, Ho, & Patel, 2011). Good interface design allows ease of use when interacting with the system and extracting information for all users. If there is a specific task identified where novice performance is highly inferior to experts, certain eye-tracking measurements can identify specific objects that are difficult to interpret or extract information from (Jacob & Karn, 2003; Just & Carpenter, 1976). By investigating eye movements of users, designers can create displays suited to the user.

2.5 Using Eye-Tracking to Investigate Expert Strategy

Eye-tracking has been used to observe expert strategies as an alternative to the think-aloud technique (users vocalize what they are looking at, thinking about, and doing). Eye-tracking allows researchers insight into behavior with minimal interruption and into periods when the users are not aware of their own behavior. In 2007, Fisher and Pollatsek used eye-tracking to study driver attention allocation and found that when novices exploited expert scanning strategies, their ability to identify risky situations increased and drivers allocated more of their attention toward areas of high-risk. Bellenkes et al. (1997) used eye-tracking to investigate the scanning techniques of expert and novice pilots. They found that

experts tended to scan to vital instruments more frequently and novices tended to dwell on instruments longer. Fisher and Pollatsek (2007) found that experts had more fixations in high-risk areas indicating their ability to identify risk while driving. Bellenkes et al. (1997) found experts had more fixations on important instruments while flying indicating their ability to attend to relevant information. Both studies used fixation frequency to extract information regarding expert strategy.

2.6 Basic Eye-Tracking Measurements

There are basic elemental measurements used to record eye movements:

1. Area of Interest (AOI) – a defined area within a field of view in which statistics and measurements can be calculated for the individual area. The AOI size and location is at the discretion of the researcher depending on the purpose of the task.
2. Gaze – (or “dwell”) the combination of consecutive fixations within a small prescribed area.
3. Fixation -- a quick stop of the eye or when the eyes are relatively stationary and are most commonly measured in regard to duration and frequency.
4. Point-of-regard – the location where the eye is actually focused
5. Saccade – the rapid movement of the eye from one fixation point to another.
6. Scan path -- the combination of fixations and saccades. It identifies the order and location of each fixation; it follows a saccade-fixation-saccade pattern (Poole & Ball, 2005).

2.7 Eye-Tracking in Aviation and Health Care

Extensive research in the aviation and healthcare industries uses eye-tracking to investigate the behavior and strategy of expert and novice pilots and physicians. Since the needs of the petrochemical operator are similar to the needs of aviation pilots and physicians in that they all require access to real-time process information to make decisions in stressful, high time pressure situations, the research should be applicable to petrochemical operators as well. The most common eye metrics used to investigate pilot and physician strategies are the number of fixations per AOI (or Fixation Frequency per AOI), gaze or “dwell” percentage on each AOI, gaze duration mean for AOIs, and scan path.

2.7.1 Number of Fixations per AOI

The number of fixations per AOI is a count of fixations that are located within a defined area. This metric indicates how often participants glance at specific areas; more frequent fixations to an area can indicate its importance (Fitts, Jones, & Milton, 1950). Ottati, Hickox, and Richter (1999) found differences in Fixation Frequency between expert and novice pilots during a navigation flight simulation. Pilots were able to reference a 90° forward field of view map representation of their current location during flight to stay on course which researchers treated as the area of interest. Experts had a higher Fixation Frequency to the reference map than novice pilots. Novices made more frequent fixations outside of the cockpit and also had more difficulty keeping the correct altitude. This may indicate novices' decreased ability to navigate based on the map; they relied more heavily on terrain features outside of the cockpit to navigate. Experts paid more attention to the map in the cockpit and were able to keep the correct altitude, which may demonstrate the importance of being able to navigate based on the map with less reliance on the terrain. Within the health care industry, Zheng et al. (2011) conducted a study on surgeon vigilance in the operating room and found that expert surgeons glanced more frequently at the anesthetic monitor than novices especially during times when the patient was unstable. Novices tended to focus most on the surgical task. In both studies, higher fixations were associated with more relevant information.

2.7.2 Gaze Percentage per AOI

Gaze or “dwell” percentage per AOI refers to the proportion of time the eyes dwell on a specific AOI for the entire session. Schriver et al. (2008) used mean percentage dwell time (MPDT) per AOI (same as “gaze” percentage) to identify attention allocation between expert and novice pilots during a flight simulation. Pilots encountered failures such as low oil pressure, loss of significant power, and electrical failures while in simulated flight. Pilots participated in 16 trial flights; in 8 of the trials failures occurred and in the other 8 trials no failures occurred. MPDT allocated to AOIs containing information relevant cues of a failure was measured from the onset of the failure until the pilot's first action to address

the failure. They found that experts allocated significantly more attention to problem relevant AOIs when a failure occurred than novice pilots. This indicates pilot expertise led them to look at appropriate instruments. Expert pilots also allocated more attention to highly relevant diagnostic cues and less to lower diagnostic cues during failure trials; novice pilots showed the opposite trend. This means that during flights where failures occurred, expert pilots paid more attention to critical areas after failures occurred.

Caution should be taken when using Gaze Percentage per AOI as the frequency of fixations and duration of fixations are not differentiated. Two different participants may be found to have the same Gaze Percentage for each AOI but one participant focused on the AOI many times for short durations and the other focused on the AOI a few times but for long periods of time. These metrics are different and should be analyzed separately. According to Fitts et al. (1950), frequency indicates the importance of the element whereas duration indicates the difficulty of interpreting the information. Although the importance and difficulty of interpretation is difficult to decipher using Gaze Percentage per AOI, this metric can indicate which AOIs are attended to for larger percentages of the time with regard to other AOIs. A post-testing interview can help reveal why the user paid more attention to certain areas.

2.7.3 Gaze Duration Mean per AOI

Gaze duration mean per AOI describes the average length of gazes within an AOI. Longer gazes during encoding tasks may indicate objects are difficult to interpret while shorter gazes can indicate efficiency of the glance. Bellenkes et al. (1997) researched the visual scanning pattern of expert and novice pilots. Eight important flight instruments were designated AOIs. Experts had overall shorter dwell times and more fixations than novices. Shorter dwell times indicate experts' ability to efficiently extract information and the higher Fixation Frequency, in support of other findings, indicates the importance of the instruments. Novices had significantly longer dwell times on instruments signifying their inefficiency at extracting information (Bellenkes et al., 1997). Gegenfurtner, Lehtinen, and Säljö (2011) found similar results from their meta-analysis of eye-tracking research on expert and novice

comprehension of visual displays. In all domains (sports, medicine, aviation, driving, etc.) experts had shorter fixation durations and more fixations of task-relevant objects and fewer fixations on task-irrelevant objects in comparison to novices. Interestingly, Taukari, Pant, and Garg (2010) found that when resolving a conflict, expert air traffic controllers (ATC) initially had fewer fixations for longer periods of time then switched to a strategy typical of experts to make more frequent but shorter fixations. They concluded that expert ATCs scattered attention initially to gain more information and then focused attention to process important and relevant information to complete the task. This is in line with previous research that expert strategies are superior in selective attention allocation and more efficient at extracting relevant information (Chapman & Underwood, 1998).

2.7.4 Scan Path

Scan paths identify the order, location, and duration of each fixation and are useful to identify the order in which items are looked at on a display. Circles represent fixations and lines represent saccades. The diameter of each circle is representative of the duration of the fixation. Order is obtained when the scan path is played back; it plots the path of the eye in real-time. Scan paths follow the pattern of fixation-saccade-fixation-saccade, etc. and can be analyzed both quantitatively and qualitatively. Quantitative analysis requires in-depth evaluation such as the string-edit method (Takeuchi & Habuchi, 2007). This technique requires the graphic to be partitioned into multiple grid-like areas (more areas yield more accurate results) and recording the location of successive fixations. Each area is assigned a letter, A-Z. The scan path is denoted by marking the letter for the series of fixations, creating a string (i.e. ABCBDIFGFGHESJDH). The string is further analyzed using a cost function. Quantitative analysis of a scan path does not necessarily analyze the plot but rather the events (fixations, saccades, etc.) that occur during the duration of the scan being plotted.

Qualitative analysis involves reviewing the scan path plot to identify any trends among participants. Videos can be played back slowed or in real-time and watch as fixations and saccades are plotted. Still images of the complete scan path plotted can identify areas with large clusters of fixations

or identify areas with little or no fixations. Sullivan, Yang, Day, and Kennedy (2011) used both quantitative and qualitative methods to analyze the influence expertise had on visual scan patterns of helicopter pilots. Experts were defined by the number of total flight hours (TFH) completed. TFH negatively correlated with mean dwell time indicating that with increased expertise, mean dwell time will decrease indicating efficient data extraction. Qualitatively, expert pilots not only looked at upcoming areas on the map but also looked back at areas of the map they had just flown over. Post-experiment comments strongly supported that pilots looked back to gain confidence in their flight orientation. This is insightful information regarding expert scanning strategy that can be utilized to train novice pilots.

Scan patterns have also been used to explain behavioral differences between expert and novice surgeons examining x-rays. Expert surgeons first examined areas of the x-ray where tumors were most likely to appear while novices tended to examine the x-ray more evenly (Kundel & LaFollette, 1972). With more experienced observers, it appears that searching is often driven by the expectancy of where the target will most likely be found.

2.8 Variability of Eye-Tracking Metrics

Some eye-tracking metrics can be interpreted differently depending on the type of task being performed. For an encoding task, a higher Fixation Frequency indicates higher importance or saliency, whereas during a search task, a higher Fixation Frequency indicates uncertainty and difficulty extracting information (Jacob & Karn, 2003). When considering gaze duration, longer gazes during an encoding task indicate difficulty interpreting the object, but for a search task, longer gazes indicate that the item is more engaging. Table 1 summarizes the common metrics and their meaning based on context.

Eye-tracking has been applied in an array of industries. In health care, researchers used AOI Fixation Frequency of surgeons performing extracorporeal circulation tasks during cardiovascular surgery to interpret attention allocation and found that expert perfusionists distributed their attention to relevant areas better than the intermediates and novices (Tomizawa et al., 2012). In aviation, researchers used

Table 1 - Eye-Tracking Metrics and What They Measure Based on Task Type

Eye-Tracking Metric	Encoding Task	Search Task
Number of Fixations per AOI	More fixations indicate higher importance or is more salient (Jacob & Karn, 2003)	More fixations indicates more uncertainty in recognizing item (Jacob & Karn, 2003)
Gaze duration mean per AOI	Higher duration indicates more difficulty interpreting information (Just & Carpenter, 1976)	Higher duration indicates the item is more engaging (Just & Carpenter, 1976)
	Encoding & Search Tasks	
Gaze Percentage per AOI	A positive correlation can indicate difficulty to interpret as well as the relative importance of the element (Fitts et al., 1950)	

eye-tracking to analyze pilot automation monitoring patterns and performance. Through AOI dwell time and fixation frequencies, they pointed out deficiencies in pilot's monitoring strategies (Sarter et al., 2007). Currently, there is extensive research using eye-tracking to investigate expert strategy of aviation pilots and medical surgeons, but there is little research in the area of supervisory control system operators such as petrochemical operators. The demands on operators are similar to pilots and surgeons in that they both need to make decisions based on real-time information that must be extracted from their consoles and/or surroundings and it is here that they differ. Pilots control their aircraft with the board of instruments in the cockpit and receive feedback from their immediate surrounding (i.e. pilots are able to look out the window to aid in orientation discrimination or to determine their heading). Surgeons extract information through physical feedback (i.e. excessive bleeding, patient seizures, etc.) as well as from monitor screens. Petrochemical operators monitor and manipulate the process from a control center through their console, which consists of sets of computer screens. Not all process information is displayed concurrently and operators have to navigate through screen displays to locate and access information and controls. These operators get little feedback from visual contact with the physical system that they control. In this way, the needs of a petrochemical operator are different which is why the current research intends to use eye-tracking methods to investigate the strategy of expert and novice operators of SCADA systems.

CHAPTER 3. METHODS

3.1 Background on Previous Research

The current research was an extension of previous research conducted at Louisiana State University, Mechanical and Industrial Engineering Department by Koffskey, Ikuma, and Harvey (2013), sponsored by the Center for Operator Performance. In the previous study, two SCADA interfaces were evaluated using performance (speed and accuracy), situation awareness, perceived workload ratings, and eye-tracking metrics during three 30-minute crude refinement simulations at three different workload levels (easy, medium, and hard). Each participant encountered only one interface that was randomly assigned. Four expert operators and twenty-four novice students participated. Fixation Frequency per AOI and overall Fixation Frequency were the eye-tracking measurements used and three AOIs were defined (Main display, Alarm Banner, and Faceplate). Each session involved a 30-minute simulation of a crude refining unit in which the operator was instructed to maintain the system and take action when failures occurred. The two types of failures were pumps shutting off (referred to as pump failures) and mode and setpoint malfunctions (referred to as cascade failures). The crude refinement process was simulated using DeltaV, a process management software produced by Emerson Process Management.

The current research focused on the relationship between expertise, eye-measurements and performance when using the two SCADA displays. Two different interfaces (black and grey) at two expertise levels (expert operator and novice student) were evaluated during active and inactive alarm periods. Specific eye measurements and performance metrics were collected. Table 2 summarizes the independent and dependent variables for the previous and current research.

3.1.1 Simulation and Interface Design

The boundaries of the simulated crude refinement process began with the crude storage tanks and ended at the overhead receiver. The simulation consisted of seven displays that the participant could navigate through: overview, crude storage tanks, desalter, heater, tower bottom, tower top, and overhead

receiver (See Appendix 1). The two interfaces (grey and black) had both functional and cosmetic differences as displayed in Table 3.

Table 2 - Independent and Dependent Variables for Previous and Current Research

Variables	Previous Research	Current Research
Independent Variables	Interface type (Black/ Grey) Workload Level (Easy, Medium, Hard) Expertise (Expert/ Novice)	Interface type (Black/ Grey) Expertise (Expert/ Novice) Alarm Activity (Active/ Inactive)
Dependent Variables	Performance (Speed/ Accuracy) Eye Metrics Fixation Frequency per AOI Overall Fixation Frequency Situation Awareness Perceived Workload Ratings	Performance (Speed/ Accuracy) Eye Metrics Fixation Frequency per AOI Gaze Duration Mean per AOI Gaze Percentage per AOI Scan path

Table 3 - Comparison of the Design Changes for the Grey and Black Interface

	Grey	Black
Appearance	Eliminated unnecessary bends in the pipelines Aligned data displays to limit scattering Increased font size and color on navigation buttons	Increased intersecting and bending pipelines Low contrast between font and background color Color coordination not related to operation
Functionality	When alarm was selected the display and faceplate open Out of mode components were indicated by a blue box Colored boxes in the Main display for indicators that a component has alarmed	Clicking on alarm only opened the display not the faceplate No indicator for components out of mode or alarmed on Main display Pumps were only identified through the faceplate

3.2 Experiment Overview and Design

The focus of the current research investigated the strategy and behavior of expert and novice petrochemical operators during the experiment sessions described in section 3.1 using eye-tracking and performance metrics. For the current research, three experts and six randomly selected novice

participants' results were analyzed. A 2x2x2 experiment design was used (See Table 4) to manipulate interface type, expertise level, and alarm activity. Interface type was a between-subjects variable evaluated using two different interfaces: the black interface and the grey interface. Expertise level was a between-subjects variable evaluated at the novice level and the expert level. Alarm activity was a within-subjects variable evaluated for active and inactive periods. Alarm activity refers to periods during the simulation when alarms are active in the Alarm Banner (active) and when there are no alarms active in the Alarm Banner (inactive). The dependent variables were Fixation Frequency per AOI, gaze duration mean per AOI, Gaze Percentage per AOI, scan path, speed and accuracy.

Table 4 - Experiment Design and Predictor Variables

Predictor B: Interface	Predictor A: Expertise			
	Novice		Expert	
	Predictor C: Alarm Activity			
	Active	Inactive	Active	Inactive
Grey	Participants 1, 4, 6		Participant 9	
Black	Participants 2, 3, 5		Participants 7, 8	

Since eye behavior was investigated at different alarm activity levels (active/ inactive), only the first fifteen minutes of the hard level was used for analysis in order to acquire evenly balanced active and inactive periods; it had approximately 8 minutes active and 7 minute inactive periods. The first fifteen minutes contained five failures (three pump failures and two cascade failures). The last fifteen minutes had a large number of failures that occurred within minutes of each other and would have an overwhelming number of active periods if the operators fell behind in their corrective actions. For this reason, the last half of the hard workload level was excluded from analysis. The number, type, and timing of events for the first 15 minutes of the hard workload level is illustrated in Figure 1.

Figure 1 - Sequence and timestamps of failures for the hard workload level

The specific goals of the current research were as follows.

1. Quantify differences in Fixation Frequency, Gaze Duration, and Gaze Percentage per AOI between expert and novice operators during active and inactive times.
2. Identify differences between scan patterns of expert and novice operators during active and inactive times.
3. Quantify differences in Fixation Frequency, Gaze Duration, and Gaze Percentage per AOI between two different interfaces.
4. Identify differences between scan patterns of operators using two different interfaces.
5. Correlate operator performance (speed and accuracy) with eye behavior specifically Fixation Frequency, Gaze Duration Mean, and Gaze Percentage per AOIs.

3.3 Participants

A total of 9 participants were used for the study: 3 experts and 6 novices. Four expert operators and twenty-four novice students participated in the previous experiment but due to poor traceability of one expert who wore glasses, eye-tracking data was only collected for three experts. Since only three experts were included, only six of the twenty-four novices were included to keep the sample sizes similar, three were randomly selected from the black and three from the grey. Complete eye-tracking data included those with a tracking ratio of 95% or higher for the eye-tracker and completed simulations in

which all scheduled failures occurred. Tracking ratio is defined as the “number of non-zero gaze positions divided by sampling frequency multiplied by run duration expressed in percent” (SensoMotoric Instruments, 2011a). Participants with sessions considered to have “glitched” because one or more failures did not occur due to a malfunction of the software were excluded. Console operators were voluntarily recruited from Flint Hills Resources in Corpus Christi, Texas as experts. Novices were recruited from students at South Central Louisiana Technical College in Reserve, Louisiana and Lambton College in Sarnia, Ontario, Canada. Students were recruited from the process technology and instrument technician programs. They had experience with process variables and controls but lacked control room experience, which made them suitable to represent the novice population. Demographic information for novices included number of semesters completed in school, work/ internship experience in the industry, degree in pursuit of, and experience with DeltaV. Expert demographic information included years of experience as a console operator, years with Flint Hills, years in the oil refining industry, and years as a field engineer. Demographic details for novice and experts are captured in Tables 5 and 6.

Table 5 - Demographic Information Collected for Novice Students

	Lambton Students (n = 12)	SCLTC Students (n = 12)	<i>P</i> -value
Semesters completed (%)	69.4 (44.1)	69.3 (22.7)	0.989
Work experience (months)	4.67 (4.58)	6.50 (20.1)	0.597

Note. Mean (SD); significant at $p < 0.05$

Table 6 - Demographic Information Collected for Expert Control Room Operators

	Participant #				Average
	26	27	28	29	
Years of experience in the oil and gas industry	23	27	21	23.5	23.6
Years of experience as a console operator	21	21	6	7	13.8
Years of experience with Flint Hills Resources	23	24	18	23.5	22.1
Years of experience outside the control room	10	3	12	16	10.3
DeltaV experience	No	Yes	No	No	--
Experimental Interface Design	Black	Gray	Black	Gray	--

Note. Participant # 27 was excluded from eye-tracking analysis

3.4 Eye-Tracking

Eye-tracking was used to evaluate eye behavior and scanning strategies of novice and expert console operators. Head-mounted SMI ETG, SensoMotoric Instruments Eye-tracking Glasses, (SensoMotoric Instruments, Inc., Boston, MA) recorded and analyzed eye-movement (Figure 2). SMI ETG are portable and allow the head and neck to move freely. SMI ETG are binocular tracking devices that have two small cameras on the rim of the glasses (one camera recording the left eye and one camera recording the right eye) and an HD scene camera on the front of the glasses. The two small cameras capture eye movements and map the point-of-regard (where the eye is focused) into the scene video (SensoMotoric Instruments, 2011b). The SMI ETG system uses iView (SensoMotoric Instruments, Inc., Boston, MA) to record eye-tracking data. SMI's BeGaze software was used for analysis.

Figure 2 - SMI eye-tracking glasses used during the experiment (<http://eyetracking.tistory.com/14>)

3.5 Experimental Procedure (Previous Research)

The previous research followed IRB-approved experimental procedures, and the steps pertaining to the current research are reviewed here. During the first session, participants trained to use the DeltaV simulator by reading a training document and practicing a simulated pump failure and cascade failure. Upon training completion, the eye-tracking glasses (ETG) were positioned on the participant and

calibrated using 3-point calibration. 3-Point Calibration, the highest level of calibration available in BeGaze, was used for all participants because it is more accurate than 0- or 1-Point Calibration. It requires three triangular calibration points in the same plane for optimal accuracy. To calibrate the ETG, the participant needs be wearing the ETG. The experimenter clicked the “3-Point Calibration” button, then instructed the participants to focus their eyes on a particular object. Once the participant fixated there, the experimenter clicked on the object in the scene and the participant was instructed to look at the next object. The three calibration objects were the top left corner of the screen, the top right corner of the screen, and a green button at the bottom edge of the screen. Eye-tracking data was recorded continuously during the entire 30-minute simulation that followed. The remaining two sessions consisted of a quick reminder of how to address failures followed by the 30-minute simulation. In total, participants completed three sessions using a single interface type, one at each workload level (Table 2).

3.5.1 How to Address Failures Within Scenarios

The operator’s goal was to maintain the system’s steady state so when failures occurred the operator should adjust values to match the original steady state values. The original steady state value was the value at the start of the simulation. When the simulation began, the system was in a steady state meaning all values were in range and no alarms were occurring. The original steady state values could be obtained by opening the process history, which gave live charting of values as well as a history. The two types of failures that occurred were pump and cascade failures. Pumps are used in the crude refining process to push crude through the system. To fix a failed pump, the operator had to acknowledge the alarm, restart the failed pump, and readjust the flow through the down-stream valve to within +/- 50 barrels per day (bpd) of the original steady state value. Cascade failures were failures within a cascade loop, which consisted of two components that were directly affected by each other. The slave component’s mode is cascade and looks to the master component to make adjustments. The master is in auto and is directly affected by changes in the slave. For example, the heater crude temperature is the master component in the cascade loop with the fuel source (slave). The fuel valve looks to the temp to

see what it should be set at. If the temperature is too high, the valve will close to decrease the fuel and if the temperature is too low the valve will open more allowing more fuel into the furnace. To address a failure within a cascade loop, the flow needs to be adjusted to +/- 50 barrels per day (bpd) of the original steady state value and the slave component's mode must be put back into cascade (CAS).

3.6 Experimental Procedure (Current Research)

The current experimental procedure consisted of the actions required to analyze the eye-tracking data for the 15-minute segment of the hard difficulty level simulation. The AOI Editor and Event Statistic features of BeGaze were used.

1. Participants were randomly selected to be included in the study. Three of the four expert participants were included; the fourth expert with no eye-tracking data was excluded. Six student participants from the eligible participant pool were randomly selected using the random function generator in Excel. The eligible participant pool excluded participants with a tracking ratio lower than 95%, participants who encountered glitched simulations, and participants whose eye-tracking recording went corrupt or froze in the middle of recording. In order to make sure the novice participants were split evenly between interface types, three novices were randomly selected from the black interface pool and three from the grey interface pool. Novice participants 1, 6, 8, 15, 23, and 25 were selected. Participants 1, 15, and 25 used the grey interface and participants 6, 8, and 23 used the black interface.
2. AOIs were created and saved using AOI Editor. Each stimulus was opened in the AOI Editor, and the drawing tools were used to create and label the AOIs. Separate AOI's were created to distinguish between active and inactive periods. AOIs are further discussed in section 3.8.
3. AOIs were mapped for each stimulus video. For each stimulus, relevant AOIs were opened and irrelevant AOIs were made invisible. AOI mapping was where AOIs were resized and repositioned frame by frame to fit the corresponding area of the scene video. AOI mapping is further discussed in section 3.9.

4. Event Statistics were opened to retrieve eye metric data. The AOI Detailed Statistics template under Event Statistics was opened to calculate Fixation Frequency, Gaze Duration Mean, and Gaze Percentage. Certain stimuli and AOIs were highlighted to limit the displayed statistics to specific stimuli and AOIs.
5. Statistics were recorded in the data collection sheet. Statistics collected from Event Statistics were assembled in an Excel spreadsheet and analyzed using SPSS.

The following procedure outlines the steps required to create scan path plots for all participants using Semantic Gaze Mapping.

1. Semantic Gaze Mapping was used to map each fixation to a reference image for each stimulus. Semantic Gaze Mapping allows the user to map each fixation onto a reference image in order to create a scan plot. Semantic Gaze Mapping is further discussed in Section 3.7.4.
2. Scan path plots were created. The scan path plot could only be created after gaze mapping had taken place. Once gaze mapping was complete, the scan path feature was opened which automatically displayed the scan path plot.

3.7 Obtaining Eye Metrics through BeGaze Software

The following sections describe how each eye metric was obtained through BeGaze and what detailed statistics were required to construct each metric.

3.7.1 Fixation Frequency per AOI

Fixation Frequency per AOI for each participant-AOI combination was obtained from the AOI Detailed Statistic template. There was no Fixation Frequency parameter so “Fixation Count” was divided by the AOI visible time to calculate Fixation Frequency. The visible time was the measured time that the AOI was visible and collecting statistics. Active periods and inactive periods were considered separately.

$$\text{Fixation Frequency per AOI} = \frac{\text{Fixation Count per AOI}}{\text{Visible time (ms) per AOI}}$$

3.7.2 Gaze Duration Mean per AOI

Gaze Duration Mean per AOI could not be obtained directly. Instead, it was obtained using dwell time and fixation count. Dwell time “starts at the moment the AOI is fixated and ends at the moment the last fixation on the AOI ends = the sum of durations from all fixations and saccades that hit the AOI” (SensoMotoric Instruments, 2011a). Fixation count is the number of fixations inside the AOI. By dividing the dwell time by the fixation count for each participant for a specific AOI, Gaze Duration Mean per AOI could be obtained.

$$\text{Gaze Duration Mean per AOI} = \frac{\text{Dwell time (ms) per AOI}}{\text{Fixation Count per AOI}}$$

3.7.3 Gaze Percentage per AOI

Gaze Percentage per AOI was obtained by dividing the total dwell time for an AOI by the total time. Total time was the entire length of each participant’s eye-tracking video clip which was 15 minutes or 900,000 ms.

$$\text{Gaze Percentage per AOI} = \frac{\text{Dwell time (ms) per AOI}}{900,000 \text{ (ms)}}$$

3.7.4 Scan Path

Two scan paths were created for each participant from the hard scenario. During testing, it was observed that expert operators were able to detect a setpoint change quickly. To investigate why operators detected the setpoint failure quickly and whether there were differences between experts and novices, both scan plots centered on the setpoint change for C-TIC-13807 that occurred 17 minutes into the hard workload scenario. The setpoint changed from 650°F to 400°F. Figure 3 illustrates the location and subtleness of the setpoint change. Scan paths were only plotted for 1-minute long segments; any longer would yield a cluttered, unreadable plot. The first scan path mapped the eye movement during the 60 seconds after the setpoint for the Crude Temperature decreased from 650°F to 400°F. The second scan path mapped the 30 seconds before and after the faceplate (C-TIC-13807) for Crude Temperature was opened after the setpoint change. This event was very subtle; there were no immediate alarm indicators

or alarm boxes that indicated a change. How quickly and in what manner the expert and novice participants noticed the setpoint change was investigated. To construct a scan plot, semantic gaze mapping was utilized. Semantic gaze mapping maps the location of each fixation onto a reference image. The reference image was a screenshot of the Main display with an opened faceplate. Appendix 3 illustrates the interface of semantic gaze mapping. The video recording plays on the right and freezes at each fixation. The fixation location is then mapped on the reference image on the left. This process was tedious but for accurate results, it is important that the mapping was precise.

Figure 3 - Location of the setpoint failure

3.8 Areas of Interest

AOIs were first created using the AOI editor feature in BeGaze (see Appendix 4 for screenshot of the AOI editor). AOIs are a useful tool allowing certain regions or defined areas to be distinctively analyzed. Figure 4 illustrates how AOIs were used to include and exclude data for analysis. See Table 7 for a complete list of the nine AOIs. The main eye measurements used to analyze the eye-tracking data were Fixation Frequency per AOI, Gaze Duration Mean per AOI, Gaze Percentage per AOI, and scan path. BeGaze software (SensoMotoric Instruments, Inc., Boston, MA) was used to playback and analyze recordings. The defined AOIs were Main display, Alarm Banner, Temp and Flow (T/F), Mode on

Faceplate (MOF), Alarm on Faceplate (AOF), and Readings on Faceplate (ROF). See Appendix 2 for an illustration of each AOI. When selecting and defining AOIs, the function and importance of the objects on-screen were considered. The faceplate AOIs were selected because this is where operators extract information about current levels, flow, etc. and also input changes. The Alarm Banner AOI is where operators extract information regarding alarms and their current status. The temp and flow AOI was also included because the crude temperature and crude flow are two crucial parameters to be aware of and is why the interface designers included those parameters on each display.

Figure 4 - Diagram illustrating how AOIs are used to include or exclude data for analysis

The steps required to retrieve detailed statistics regarding AOIs is as follows:

1. In AOI Editor, create and save the nine AOIs.
2. The AOIs are mapped for each stimulus or recording in AOI Editor. A stimulus is considered to be a single participant's video recording. AOI mapping is described in more detail in 3.9.
3. Open the Event Statistics tab and select the AOI Detailed Statistics template. This template displays statistics as one row per AOI-stimulus combination.
4. Select the appropriate stimuli and AOIs. Table 8 represents the output for the following selection: Participants 1, 2, and 3 fixation count for the Main display during active and inactive periods.

Table 7 - AOI List for Active and Inactive Periods

Active AOIs	Inactive AOIs
Main Display – active Alarm Banner – active Temp and Flow – active Alarm on Faceplate – active Mode on Faceplate – active Readings on Faceplate – active	Main Display – inactive Alarm Banner – inactive Temp and Flow – inactive

Table 8 - Sample Output for Event Statistics

Stimulus	Area of Interest	...Fixation Count
P1	Main Display_Active	343
P1	Main Display_Inactive	562
P2	Main Display_Active	542
P2	Main Display_Inactive	432
P3	Main Display_Active	154
P3	Main Display_Inactive	198

Note. Sample output for novice fixation count using the grey interface for the Main display AOI.

3.9 AOI Mapping

Once the AOIs were created, they were mapped over the recordings. Since the ETG is non-stationary and the computer display is stationary, the orientation of the display changes in the scene video depending on the movement of the head wearing the ETG. This complicated the mapping of the AOIs. AOI mapping required frame by frame manipulation of the AOI size and orientation. If the head was relatively still for a series of frames, the AOIs did not require resizing, but once the head tilted to the left, right, or closer to the computer display, the orientation of the computer display changed in the scene video and the AOIs needed to be manipulated. Another consideration during AOI mapping was ensuring that the correct AOIs were visible. Since there were different AOIs for active and inactive times, the active AOIs needed to be visible and the inactive AOIs needed to be invisible. Visibility refers to the

capability of the AOI to record data, so when an AOI is visible, fixations that are located in the area of interest are recorded. When an AOI is invisible, the AOI is no longer visible on screen and does not record fixation data. Visibility of the AOI was also considered when the faceplate was not open. All faceplate AOIs were invisible when the faceplate was closed and visible when opened. When faceplates were opened that did not contain the mode buttons or the level readings, the particular AOIs would be turned invisible for the period that the faceplate was opened. Another consideration was the dynamic interface. The system had seven different screens but only one display could be viewed at a time. Different Main display AOIs were not created for each display, instead the same AOI was used for all seven displays. Each display functions quite similarly to the others; operators primarily monitor and scan for information using the displays. Since activities were similar between displays there was no need to use different AOIs so each individual display was considered in terms of a general “main display” AOI.

3.10 Alarm Activity

In order to analyze eye behavior based on alarm activity, separate AOI's were created for active and inactive periods. Inactive AOI's included Main, Alarm Banner, and Temp & Flow. Active AOI's included Main, Alarm Banner, Temp & Flow, AOF, MOF, and ROF. The effect that alarm activity (active/ inactive) had on Fixation Frequency per AOI, Gaze Duration Mean per AOI, and Gaze Percentage per AOI were analyzed. Speed and accuracy could not be assessed during inactive times because there were no failures that occurred during inactive periods. During inactive times when there were no alarms, operators only needed to monitor the process.

3.11 Performance

Performance, in terms of speed and accuracy, was included in the study to test for correlation with eye behavior. Significant differences detected between expert and novice eye behavior and performance could indicate which eye behavior and strategy is more efficient. It can also be an indicator of which interface type is more efficient in terms of allowing users to complete tasks faster and more accurately. Speed was measured in seconds required to complete tasks. Tasks were defined as the

required actions to correct and maintain the system. Task time was measured from the time the alarm was triggered to event completion. Each event had certain requirements to be considered complete. The simulation was preprogrammed to indicate a time stamp when each event was completed, if completed. Accuracy was measured as a percentage using two different metrics: correct steps completed out of the total number of steps required and the total number of events completed out of the total number of events that occurred. Previous analysis found speed to be significantly slower for the black ($M = 233s$, $SD = 147s$, $p < 0.05$) interface versus the grey interface ($M = 180s$, $SD = 123s$, $p < 0.05$). Accuracy for the grey interface ($M = 74.4\%$, $SD = 26.2\%$, $p < 0.05$) was significantly higher than for the black interface ($M = 58.6\%$, $SD = 26.5\%$, $p < 0.05$) in regard to the total number of events completed. The sample of 6 novices included in the current experiment were tested for their representativeness of the entire initial group of novices using a t-test (Table 9). % Complete, % Correct, and TTF showed no significant differences between groups. TTC showed a significant difference because of the large deviation in TTC among the six thesis participants.

Table 9 - Independent T-Test between Thesis Participants and Entire Sample Size

	Thesis Sample			Entire Sample		
	<u>Mean (SD)</u>	<u>n</u>		<u>Mean (SD)</u>	<u>n</u>	<u>p-value</u>
% Complete	.494 (.241)	6		.486 (.216)	18	0.751
% Correct	.626 (.181)	6		.613 (.107)	18	0.184
TTF	118 (65.8)	6		93.5 (43.3)	18	0.309
TTC	299 (158)	6		287 (68.4)	18	0.016*

Note. *Significant at $p < 0.05$

3.12 Hypothesis

The main hypothesis predicted that there would be a significant difference between expert and novice eye behavior.

H_0 : Expert Eye Metrics = Novice Eye Metrics

H_1 : Expert Eye Metrics $\neq$ Novice Eye Metrics

Individual hypotheses predicted that experts would have higher fixation frequencies than novices, experts would have shorter mean dwell durations than novices, and experts would have more optimal scan paths than novices.

3.13 Data Analysis

A MANOVA was conducted to test for significance due to main effects of expertise, interface, and alarm activity using IBM SPSS Statistics Version 21, a statistical analysis software. A significance level of 0.05 was used. Due to the small sample size cells had very low number of observations (i.e. expert/grey interface, $n = 1$) a test for interaction could not be completed; only differences due to main effects were analyzed. The criterion variables were Fixation Frequency per AOI, Gaze Duration Mean per AOI, Gaze Percentage per AOI, speed and accuracy. The criterion variables were assessed for the first fifteen minutes of the hard scenario between experts and novices, black and grey interfaces, and active and inactive periods.

Due to the nature of the alarm activity variable, two models were used for data analysis. Main Display, Alarm Banner, and Temp & Flow AOI's were visible for the black and grey interfaces, for both expert and novice operators, and during both active and monitoring periods. AOF, MOF and ROF were only visible during active alarm periods for all participants. This means that there is no participant data during monitoring periods for these three AOI's. Since it was intended to investigate the eye behavior of experts and novices while looking at the faceplate two models were used to analyze data separately. Each model is described in Table 10.

Fixation Frequency, Gaze Duration Mean, and Gaze Percentage were tested for normality (Shapiro-Wilk's statistic, $\alpha = .05$) and homogeneity of variance (Levene's Statistic, $\alpha = .05$). There were several variables that did not meet the normality assumption. Since MANOVA is typically robust to non-normality only criterion variables that were significant for 2 or more independent variables were

Table 10 - Independent and Dependent Variables Analyzed Within Each Model

Model 1		Model 2	
<u>Independent Variables</u>	<u>Dependent Variables</u>	<u>Independent Variables</u>	<u>Dependent Variables</u>
Alarm Activity	FF Main Display	Expertise	FF Main Display
– Active	FF Alarm Banner	– Expert	FF Alarm Banner
– Monitoring	FF Temp & Flow	– Novice	FF Temp & Flow
Expertise		Interface Type	FF AOF
– Expert	GD Main Display	– Black	FF MOF
– Novice	GD Alarm Banner	– Grey	FFROF
Interface Type	GD Temp & Flow		
– Black			GD Main Display
– Grey	GP Main Display		GD Alarm Banner
	GP Alarm Banner		GD Temp & Flow
	GP Temp & Flow		GD AOF
			GD MOF
			GD ROF
			GP Main Display
			GP Alarm Banner
			GP Temp & Flow
			GP AOF
			GP MOF
			GP ROF

Note. FF: Fixation Frequency, GD: Gaze Duration Mean, GP: Gaze Percentage per AOI

transformed. Fixation Frequency: Alarm Banner, Gaze Duration: Main, and Gaze Percentage: Alarm Banner were exponentially transformed. Gaze Duration: Alarm Banner and Gaze Percentage: AOF did not meet the homogeneity of variance assumption. Since the participant data did not meet the homogeneity of variance assumption an additional non-parametric Independent-Samples Median Test ($\alpha = .05$) was conducted. A non-parametric analysis does not assume normality and compares medians instead of population means. This analysis was also used to account for the small sample size where averages can be greatly impacted by extreme data points. Pearson correlations were conducted for Fixation Frequency, Gaze Duration Mean, Gaze Percentage, Speed (TTF, TTC) and Accuracy (% Correct, % Complete). Only correlation coefficients $|r| \geq 0.75$ were considered.

The qualitative analysis included scan paths, which mapped eye behavior of participants. A qualitative scan plot of the eye behavior during the event was created for each participant using SMI BeGaze software, an eye-tracking analysis software by SMI. Goldberg and Kotval (1998) describe optimal scan paths during a search task to contain more straight lines to the target and relatively short fixation durations at the target. Concerning the quantitative analysis, the method described by Takeuchi and Habuchi (2007) was not utilized because of the difficulty in using the gridded areas with mobile ETG. Since the video for the current study is dynamic and the orientation of the computer display is constantly changing in the scene video, it would be extremely difficult to map the grids accurately. Instead, the scan paths were subjectively evaluated. For each scan path, the video was reviewed to note what the participant was doing (i.e. searching for a component, monitoring, diagnosing an alarm, etc.) during the 1-minute segment. Process conditions were noted such as the number of alarms active or whether any new alarms were triggered during the segment. These notes aided in understanding participant scan paths. Scan paths were also separated and subjectively evaluated based on certain criteria. Experts and novices, black and grey interface types, Louisiana participants and Sarnia participants, and participants that recognized the setpoint change within 2 minutes and those that took longer were compared. Scan paths were grouped in order to recognize any similarities or patterns between groups.

CHAPTER 4. RESULTS

4.1 Quantitative Analysis

The raw averages for the novice and expert groups (Table 11) did not reveal any substantial differences. Novices had higher fixation frequencies on the Main, Alarm Banner, AOF, and ROF AOIs. Experts had higher fixation frequencies on the Temp & Flow and MOF. Novices had longer gaze durations on Alarm Banner and ROF. Experts had longer gazes on Main, Temp & Flow, AOF, and MOF. Novices had higher gaze percentages for Alarm Banner, AOF and ROF. Experts had higher gaze percentages for Main, Temp & Flow, and MOF.

Table 11 - Descriptive Statistics for Expert and Novice AOI Fixation Details

		Novice (n = 6)	Expert (n = 3)
Fixation Frequency	AOI	Mean (SD)	Mean (SD)
(fixation/sec)	Main	1.31 (.421)	1.26 (.178)
	Alarm Banner	.196 (.139)	.122 (.143)
	Temp & Flow	.0969 (.0615)	.142 (.0341)
	AOF	.227 (.0828)	.198 (.0615)
	MOF	.341 (.220)	.398 (.0861)
	ROF	1.18 (.553)	.744 (.241)
Gaze Duration Mean	AOI	Mean (SD)	Mean (SD)
(ms)	Main	281 (26.7)	313 (3.47)
	Alarm Banner	262 (17.2)	233 (52.2)
	Temp & Flow	346 (128)	366 (105)
	AOF	265 (58.4)	289 (21.9)
	MOF	285 (67.1)	305 (36.9)
	ROF	329 (95.0)	312 (25.9)
Gaze Percentage	AOI	Mean (SD)	Mean (SD)
	Main	.176 (.0622)	.190 (.0474)
	Alarm Banner	.0272 (.0191)	.0136 (.0165)
	Temp & Flow	.0113 (.00636)	.0244 (.00756)
	AOF	.0172 (.0193)	.0156 (.00563)
	MOF	.0247 (.0196)	.0349 (.0177)
	ROF	.109 (.0635)	.0796 (.0539)

A MANOVA was conducted on Model 1, which tested Fixation Frequency, Gaze Duration Mean, and Gaze Percentage per Main Display, Alarm Banner and Temp & Flow for main effects due to interface type, expertise, and alarm activity (Table 12). There were no significant differences between interface type ($F(9) = 2.377, p = 0.331$), expertise ($F(9) = 0.166, p = 0.978$), and alarm activity ($F(9) = 2.972, p = 0.277$). The MANOVA analysis for Model 2 analyzed Fixation Frequency, Gaze Duration Mean, and Gaze Percentage for all AOI's for main effects due to interface type and expertise (Table 13). Neither expertise ($F(5) = 0.354, p = 0.750$) nor interface type ($F(5) = 0.376, p = 0.836$) resulted in any significant differences.

Table 12 - Model 1: MANOVA Main Effects Results

<u>Variable</u>	<u>F-value</u>	<u>df</u>	<u>Error df</u>	<u>P-value</u> (Wilks' Lambda)	<u>Observed Power</u>
Interface	2.377	9	2	0.331	0.158
Expertise	0.166	9	2	0.978	0.058
Alarm Activity	2.972	9	2	0.277	0.184

Note $\alpha = .05$

Table 13 - Model 2: MANOVA Main Effects Results

<u>Variable</u>	<u>F-value</u>	<u>df</u>	<u>Error df</u>	<u>P-value</u>	<u>Observed Power</u>
Interface	0.376	5	1	0.836	0.059
Expertise	0.354	5	1	0.846	0.058

Note $\alpha = .05$

The results from the Model 1 single ANOVA comparisons are in Tables 14, 15, and 16. Since MANOVA was not significant, each single ANOVA was investigated with an adjusted alpha level. The Bonferroni approach was used, which uses a significance level of α / n , where n is the number of independent variables (Johnson, 1998). A significance level of $.05 / 3 = 0.0167$ was used to evaluate individual ANOVAs. For Model 1, there were no significant results due to interface type and expertise. For alarm activity, Fixation Frequency: Main AOI ($p = 0.000$) was significantly higher during monitoring periods than active alarm periods. Gaze Percentage: Main AOI ($p = 0.011$) was also significantly higher for monitoring periods.

Table 14 - Single ANOVA Results for Model 1: Main Effects Due to Interface

		Black		Grey			
Fixation Frequency	AOI	# of Observations	Mean (SD)	# of Observations	Mean (SD)	P-value	Observed Power
(fixation/ms)	Main	5	.00178 (.000585)	4	.00194 (.000682)	0.380	0.133
	Alarm Banner	5	.000174 (.000119)	4	.000188 (.000222)	0.733	0.088
	Temp & Flow	5	.000101 (.0000532)	4	.000141 (.0000684)	0.322	0.157
Gaze Duration Mean	AOI	No. of Observations	Mean (SD)	No. of Observations	Mean (SD)	P-value	Observed Power
(ms)	Main	5	370 (30.8)	4	279 (16.8)	0.141	0.304
	Alarm Banner	5	248 (36.2)	4	260 (26.1)	0.455	0.109
	Temp & Flow	5	360 (106)	4	342 (90.5)	0.754	0.06
Gaze Percentage	AOI	No. of Observations	Mean (SD)	No. of Observations	Mean (SD)	P-value	Observed Power
	Main	5	.256 (.0884)	4	.266 (.144)	0.888	0.052
	Alarm Banner	5	.0207 (.0150)	4	.0231 (.0269)	0.775	0.079
	Temp & Flow	5	.0173 (.0102)	4	.0213 (.0121)	0.625	0.074

Note. $\alpha = .0167$

Table 15 - Single ANOVA Results for Model 1: Main Effects Due to Expertise

		Novice		Expert			
Fixation Frequency (fixation/ms)	<u>AOI</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>P-value</u>	<u>Observed Power</u>
	Main	6	.00185 (.000593)	3	.00186 (.000717)	0.807	0.056
	Alarm Banner	6	.000209 (.000183)	3	.000121 (.000123)	0.306	0.201
	Temp & Flow	6	.000116 (.0000684)	3	.000125 (.0000519)	0.775	0.058
Gaze Duration Mean (ms)	<u>AOI</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>P-value</u>	<u>Observed Power</u>
	Main	6	292 (33.3)	3	300 (18.3)	0.557	0.093
	Alarm Banner	6	259 (24.0)	3	242 (44.0)	0.483	0.101
	Temp & Flow	6	362 (96.6)	3	331 (103)	0.581	0.081
Gaze Percentage	<u>AOI</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>P-value</u>	<u>Observed Power</u>
	Main	6	.261 (.127)	3	.259 (.0866)	0.986	0.05
	Alarm Banner	6	.0254 (.0223)	3	.0144 (.0152)	0.307	0.198
	Temp & Flow	6	.0183 (.0114)	3	.0205 (.0107)	0.737	0.061

Note $\alpha = .0167$

Table 16 - Single ANOVA Results for Model 1: Main Effects Due to Alarm Activity

		Monitoring		Active			
Fixation Frequency (fixation/ms)	<u>AOI</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>P-value</u>	<u>Observed Power</u>
	Main	9	.00237 (.000274)	9	.00133 (.000347)	0.000*	1
	Alarm Banner	9	.000145 (.000115)	9	.000215 (.000208)	0.629	0.063
	Temp & Flow	9	.000137 (.0000646)	9	.000101 (.0000569)	0.355	0.142
Gaze Duration Mean (ms)	<u>AOI</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>P-value</u>	<u>Observed Power</u>
	Main	9	297 (32.3)	9	292 (26.602)	0.852	0.054
	Alarm Banner	9	255 (32.8)	9	252 (32.6)	0.444	0.112
	Temp & Flow	9	351 (82.4)	9	353 (114)	0.983	0.05
Gaze Percentage	<u>AOI</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>No. of Observations</u>	<u>Mean (SD)</u>	<u>P-value</u>	<u>Observed Power</u>
	Main	9	.340 (.0979)	9	.181 (.0551)	0.011*	0.802
	Alarm Banner	9	.0208 (.0233)	9	.0227 (.0185)	0.938	0.051
	Temp & Flow	9	.0225 (.0120)	9	.0156 (.00909)	0.341	0.148

Note $\alpha = .0167$

The results from the Model 2 single ANOVA comparisons are in Tables 17 and 18. For Model 2, interface type yielded no significance. Expertise yielded no significance but some approached significance. Novice participants had a greater Fixation Frequency: ROF AOI ($p = 0.067$) and a lower Gaze Percentage: Temp & Flow ($p = 0.069$) than experts.

The non-parametric analysis revealed no significant differences due to interface type (Table 19) and expertise (Table 20). However for alarm activity (Table 21), Fixation Frequency: Main ($p = 0.000$) and Gaze Percentage: Main ($p = 0.000$) had significantly higher medians during monitoring periods. These results are similar to the Single ANOVA results.

Pearson correlations of participant data were calculated for Fixation Frequency, Gaze Duration, Gaze Percentage, speed and accuracy (Tables 22, 23, and 24). Since alarm activity was a within-subjects variable, speed and accuracy were considered for both active and monitoring AOIs. As Fixation Frequency on the Alarm Banner increased, participants had a higher Fixation Frequency on the alarms on the faceplate ($r = 0.862$; $p \leq 0.01$). As participants looked more frequently on the Temp & Flow AOI, they looked less frequently on the readings on the faceplate ($r = -0.831$; $p \leq 0.01$). Participants that had longer Gaze Durations on the MOF also had longer gazes on the ROF ($r = 0.815$; $p \leq 0.01$). Also as participants had longer gazes on the MOF, they had a higher Gaze Percentage on the Main display ($r = 0.801$; $p \leq 0.01$). The correlations showed that participants that had longer gazes on the ROF also had higher Gaze Percentages on ROF ($r = 0.831$; $p \leq 0.01$).

As Fixation Frequency on the Alarm Banner increased, Gaze Duration on the MOF decreased ($r = -0.770$; $p \leq 0.05$). Participants that had lower Fixation Frequencies on the ROF had higher Gaze Percentages on Temp & Flow ($r = -0.752$; $p \leq 0.05$). Decreased Gaze Durations on the MOF was correlated with higher percentage of events correctly completed ($r = -0.757$; $p \leq 0.05$).

Table 17 - Single ANOVA Results for Model 2: Main Effects Due to Interface

		Black		Grey			
Fixation Frequency	AOI	No. of Observations	Mean (SD)	No. of Observations	Mean (SD)	P-value	Observed Power
(fixation/ms)	Main	5	.00131 (.000421)	4	.00135 (.000289)	0.951	0.05
	Alarm Banner	5	.000196 (.000139)	4	.000237 (.000298)	0.71	0.062
	Temp & Flow	5	.0000969 (.0000615)	4	.000106 (.0000593)	1	0.05
	AOF	5	.000227 (.0000828)	4	.000469 (.000578)	0.67	0.066
	MOF	5	.000341 (.000220)	4	.000436 (.000315)	0.806	0.055
	ROF	5	.00118 (.000553)	4	.000956 (.000140)	0.513	0.089
Gaze Duration Mean	AOI	No. of Observations	Mean (SD)	No. of Observations	Mean (SD)	P-value	Observed Power
(ms)	Main	5	303 (27.1)	4	279 (22.0)	0.397	0.118
	Alarm Banner	5	253 (41.4)	4	251 (23.5)	0.703	0.063
	Temp & Flow	5	368 (109)	4	334 (134)	0.609	0.073
	AOF	5	267 (52.2)	4	280 (51.3)	0.836	0.054
	MOF	5	317 (49.2)	4	261 (56.3)	0.434	0.108
	ROF	5	352 (91.6)	4	287 (36.9)	0.444	0.105
Gaze Percentage	AOI	No. of Observations	Mean (SD)	No. of Observations	Mean (SD)	P-value	Observed Power
	Main	5	.193 (.0576)	4	.165 (.0552)	0.745	0.059
	Alarm Banner	5	.0237 (.0167)	4	.0214 (.0231)	0.552	0.082
	Temp & Flow	5	.0180 (.0118)	4	.0126 (.00369)	0.263	0.178
	AOF	5	.0107 (.00350)	4	.0242 (.0221)	0.382	0.123
	MOF	5	.0248 (.0166)	4	.0323 (.0226)	0.357	0.132
	ROF	5	.110 (.0727)	4	.0854 (.0416)	0.91	0.051

Note $\alpha = .0167$

Table 18 - Single ANOVA Results for Model 2: Main Effects Due to Expertise

		Novice		Expert			
Fixation Frequency (fixation/ms)	AOI	No. of Observations	Mean (SD)	No. of Observations	Mean (SD)	P-value	Observed Power
	Main	6	.00137 (.000418)	3	.00126 (.000178)	0.695	0.064
	Alarm Banner	6	.000261 (.000231)	3	.000122 (.000143)	0.336	0.141
	Temp & Flow	6	.0000801 (.0000563)	3	.000142 (.0000341)	0.202	0.224
	AOF	6	.000402 (.000463)	3	.000198 (.0000615)	0.526	0.087
	MOF	6	.000376 (.000315)	3	.000398 (.0000861)	0.955	0.05
	ROF	6	.00124 (.000393)	3	.000744 (.000241)	0.067	0.471
Gaze Duration Mean (ms)	AOI	No. of Observations	Mean (SD)	No. of Observations	Mean (SD)	P-value	Observed Power
	Main	6	281 (26.7)	3	313 (3.47)	0.104	0.364
	Alarm Banner	6	262 (17.2)	3	233 (52.2)	0.287	0.164
	Temp & Flow	6	346 (128)	3	366 (105)	0.994	0.05
	AOF	6	265 (58.4)	3	289 (21.9)	0.618	0.072
	MOF	6	285 (67.1)	3	305 (36.9)	0.543	0.084
	ROF	6	329 (95.0)	3	312 (25.9)	0.785	0.056
Gaze Percentage	AOI	No. of Observations	Mean (SD)	No. of Observations	Mean (SD)	P-value	Observed Power
	Main	6	.176 (.0622)	3	.190 (.0474)	0.744	0.059
	Alarm Banner	6	.0272 (.0191)	3	.0136 (.0165)	0.289	0.163
	Temp & Flow	6	.0113 (.00636)	3	.0244 (.00756)	0.069	0.462
	AOF	6	.0172 (.0193)	3	.0156 (.00563)	0.991	0.05
	MOF	6	.0247 (.0196)	3	.0349 (.0177)	0.349	0.135
	ROF	6	.109 (.0635)	3	.0796 (.0539)	0.696	0.063

Note $\alpha = .0167$

Table 19 - Non-parametric Results: Interface Type

Fixation Frequency (fixation/ms)	AOI	Black (min—median—max)	Grey (min—median—max)	P-value
	Main	.000879-- .00199 --.00257	.00105-- .00186 --.00281	1.00
	Alarm Banner	.0000327-- .000200 --.000312	.0000287-- .0000829 --.000673	0.637
	Temp & Flow	.0000284-- .000105 --.000162	.0000552-- .000128 --.000257	1.00
	AOF	.000133-- .000231 --.000339	.000159-- .000190 --.00134	0.524
	MOF	.0000377-- .000334 --.000600	.0000457-- .000451 --.000797	1.00
	ROF	.000475-- .00125 --.00190	.000789-- .000951 --.00113	0.524
Gaze Duration Mean (ms)	AOI	Black	Grey	P-value
	Main	270-- 307 —355	265-- 270 --312	0.153
	Alarm Banner	190-- 258 --291	219-- 254 --298	1.00
	Temp & Flow	171-- 353 --485	241-- 327 --534	0.637
	AOF	178-- 276 --314	223-- 275 --347	1.00
	MOF	263-- 318 --390	186-- 268 --323	0.524
	ROF	281-- 321 --513	254-- 286 --321	0.524
Gaze Percentage	AOI	Black	Grey	P-value
	Main	.123-- .250 --.400	.0821-- .230 -- .563	1.000
	Alarm Banner	.00307-- .0226 --.0393	.00341-- .00936 --.0753	0.637
	Temp & Flow	.00133-- .0194 --.0307	.00776-- .0185 --.0463	1.000
	AOF	.00588-- .0104 --.0157	.00842-- .0160 --.0563	1.000
	MOF	.000706-- .0264 --.0464	.00104-- .0368 --.0545	0.206
	ROF	.0368-- .0831 --.215	.0400-- .0806 --.140	1.000

Note $\alpha = .05$

Table 20 - Non-parametric Results: Expertise

Fixation Frequency	AOI	Novice (min—median—max)	Expert (min—median—max)	P-value
<u>(fixation/ms)</u>				
	Main	.000879-- .00203 --.00261	.00114-- .00174 --.00281	1.00
	Alarm Banner	.0000327-- .000189 --.000673	.0000287-- .0000494 --.000287	0.620
	Temp & Flow	.0000291-- .0000959 --.000228	.0000284-- .000147 --.000162	0.620
	AOF	.000133-- .000224 --.00134	.000162-- .000163 --.000269	1.00
	MOF	.0000377-- .000388 --.000797	0.000334-- .000364 --.000496	1.00
	ROF	.000789-- .00119 --.00190	.000475-- .000816 --.000940	0.167
Gaze Duration Mean	AOI	Novice	Expert	P-value
<u>(ms)</u>				
	Main	265-- 275 --355	270-- 307 --317	0.620
	Alarm Banner	206-- 266 --290	190-- 231 --298	0.620
	Temp & Flow	241-- 344 --534	171-- 338 --485	1.00
	AOF	178-- 273 --347	272-- 281 --314	0.524
	MOF	186-- 277 --390	263-- 323 --330	0.524
	ROF	254-- 316 --513	281-- 321 --330	0.524
Gaze Percentage	AOI	Novice	Expert	P-value
	Main	.0821-- .250 --.563	.139-- .260 --.379	1.00
	Alarm Banner	.00307-- .0186 --.0753	.00341-- .00496 --.0353	0.620
	Temp & Flow	.00358-- .0169 --.0463	.00133-- .0237 --.0307	0.620
	AOF	.00588-- .0106 --.0563	.00999-- .0157 --.0212	0.524
	MOF	.000706-- .0300 --.0464	.0199-- .0305 --.0545	1.00
	ROF	.0400-- .0854 --.215	.0368-- .0619 --.140	1.00

Note $\alpha = .05$

Table 21 - Non-parametric Results: Alarm Activity

Fixation Frequency	<u>AOI</u>	<u>Monitoring</u> (min—median—max)	<u>Active</u> (min—median—max)	<u>P-value</u>
<u>(fixation/ms)</u>	Main	.00201-- .00228 --.00281	.000879-- .00117 --.00197	0.000*
	Alarm Banner	.0000327-- .0000986 --.000349	.0000287-- .000180 --.000673	1.00
	Temp & Flow	.0000284-- .000143 --.000257	.0000291-- .0000753 --.000190	0.347
	AOF	.000133-- .000217 --.00134	--	--
	MOF	.0000377-- .000364 --.000797	--	--
	ROF	.000475-- .000963 --.00190	--	--
<u>Gaze Duration Mean</u>	<u>AOI</u>	<u>Monitoring</u>	<u>Active</u>	<u>P-value</u>
<u>(ms)</u>	Main	265-- 287 --355	267-- 275 --335	1.00
	Alarm Banner	206-- 250 --298	190-- 257 --291	1.00
	Temp & Flow	171-- 355 --484	241-- 289 --534	0.347
	AOF	178-- 276 --347	--	--
	MOF	186-- 282 --390	--	--
	ROF	254-- 317 --513	--	--
<u>Gaze Percentage</u>	<u>AOI</u>	<u>Monitoring</u>	<u>Active</u>	<u>P-value</u>
	Main	.252-- .318 --.563	.0821-- .192 --.249	0.000*
	Alarm Banner	.00307-- .0113 --.0753	.00341-- .0207 --.0542	1.00
	Temp & Flow	.00133-- .0210 --.0463	.00358-- .0149 --.0307	0.347
	AOF	.00588-- .0108 --.0563	--	--
	MOF	.000706-- .0305 --.0545	--	--
	ROF	.0368-- .0831 --.215	--	--

Note $\alpha = .05$

Table 22 - Correlation Between Fixation Frequency and Gaze Duration

		Fixation Frequency						Gaze Duration					
		<u>Main</u>	<u>Alarm Banner</u>	<u>Temp & Flow</u>	<u>AOF</u>	<u>MOF</u>	<u>ROF</u>	<u>Main</u>	<u>Alarm Banner</u>	<u>Temp & Flow</u>	<u>AOF</u>	<u>MOF</u>	<u>ROF</u>
Fixation Frequency	<u>Main</u>	1	-0.301	0.093	-0.210	-0.026	0.424	0.005	0.132	-0.112	0.169	0.142	-0.158
	<u>Alarm Banner</u>		1	0.092	.862**	-0.477	-0.018	-0.359	0.3	-0.053	0.332	-.770*	-0.512
	<u>Temp & Flow</u>			1	0.539	-0.262	-0.831**	-0.173	-0.005	0.144	0.59	-0.426	-0.26
	<u>AOF</u>				1	-0.524	-0.155	-0.386	0.227	-0.316	0.594	-.688*	-0.38
	<u>MOF</u>					1	-0.151	-0.169	-0.467	0.299	-0.583	0.026	-0.194
	<u>ROF</u>						1	-0.287	0.053	-0.245	-0.267	0.37	0.258
Gaze Duration	<u>Main</u>							1	0.12	0.298	0.222	.716*	.681*
	<u>Alarm Banner</u>								1	0.337	0.068	-0.265	0.058
	<u>Temp & Flow</u>									1	0.077	0.248	0.197
	<u>AOF</u>										1	-0.12	-0.114
	<u>MOF</u>											1	.815**
	<u>ROF</u>												1

Note ** significant at a level of 0.01

* significant at a level of 0.05

Table 23 - Correlation Between Gaze Percentage and Dependent Variables

		Gaze Percentage					
		<u>Main</u>	<u>Alarm Banner</u>	<u>Temp & Flow</u>	<u>AOF</u>	<u>MOF</u>	<u>ROF</u>
Fixation Frequency	<u>Main</u>	.811**	-0.082	0.108	-0.416	-0.302	-0.315
	<u>Alarm Banner</u>	-0.212	.863**	-0.064	.698*	-0.66	-0.493
	<u>Temp & Flow</u>	-0.086	-0.038	.830**	0.669	-0.261	-0.521
	<u>AOF</u>	-0.609	.699*	-0.1	.931**	-0.607	-0.429
	<u>MOF</u>	0.041	-0.443	-0.022	-0.43	.764*	0.05
	<u>ROF</u>	0.434	0.152	-.752*	-0.368	-0.126	0.427
Gaze Duration Mean	<u>Main</u>	0.212	-0.269	0.002	-0.19	0.169	0.422
	<u>Alarm Banner</u>	0.145	0.399	0.085	0.061	-0.602	-0.082
	<u>Temp & Flow</u>	-0.03	-0.003	0.453	-0.278	0.088	0.075
	<u>AOF</u>	0.077	0.181	0.36	0.599	-0.593	-0.429
	<u>MOF</u>	.801**	-.680*	0.173	-0.617	0.302	.713*
	<u>ROF</u>	0.454	-0.497	0.184	-0.319	0.094	.831**
Gaze Percentage	<u>Main</u>	1	0.194	0.162	-0.644	0.094	0.286
	<u>Alarm Banner</u>		1	0.03	0.482	-0.651	-0.448
	<u>Temp & Flow</u>			1	0.072	0.017	-0.145
	<u>AOF</u>				1	-0.35	-0.314
	<u>MOF</u>					1	0.453
	<u>ROF</u>						1

Note ** significant at a level of 0.01

* significant at a level of 0.05

Table 24 - Correlation Between Speed and Accuracy and Dependent Variables

		<u>Inactive (Monitoring)</u>				<u>Active</u>			
		<u>Accuracy</u>		<u>Speed</u>		<u>Accuracy</u>		<u>Speed</u>	
		% Correct	% Complete	TTF	TTC	% Correct	% Complete	TTF	TTC
Fixation Frequency	Main	0.305	-0.089	0.179	-0.029	-0.457	-0.588	0.098	0.05
	Alarm Banner	0.294	0.468	-0.171	0.194	0.151	0.402	-0.035	0.057
	Temp & Flow	0.149	0.47	-0.262	-0.053	0.305	0.458	-0.351	0.499
	AOF	-	-	-	-	0.075	0.29	0.111	-0.023
	MOF	-	-	-	-	-0.175	0.334	-0.634	0.352
	ROF	-	-	-	-	-0.435	-0.654	0.581	-0.564
Gaze Duration	Main	-0.444	-.839**	.873**	-0.670*	0.003	-0.332	0.26	-0.052
	Alarm Banner	0.321	-0.215	.676*	-0.439	0.154	0.131	0.265	-0.374
	Temp & Flow	0.135	-0.145	0.413	-0.43	-0.301	0.064	-0.107	0.092
	AOF	-	-	-	-	-0.164	-0.223	0.249	0.164
	MOF	-	-	-	-	-0.366	-0.757*	0.504	-0.374
	ROF	-	-	-	-	-0.264	-0.565	.675*	-.716*
Gaze Percentage	Main	0.037	0.007	0.182	0.039	-0.521	-.876**	0.372	-0.155
	Alarm Banner	0.249	0.405	-0.022	0.102	0.112	0.319	-0.036	0.049
	Temp & Flow	0.115	0.473	-0.244	0.077	0.1	0.179	-0.281	0.427
	AOF	-	-	-	-	0.255	0.401	0.071	0.034
	MOF	-	-	-	-	0.209	0.222	-0.298	0.131
	ROF	-	-	-	-	-0.073	-0.37	0.620	-0.748*
Inactive	% Correct	1	0.578	-0.196	0.168	-	-	-	-
	% Complete		1	-0.685*	0.421	-	-	-	-
	TTF			1	-.847**	-	-	-	-
	TTC				1	-	-	-	-
Active	% Correct					1	0.578	-0.196	0.168
	% Complete						1	-0.685*	0.421
	TTF							1	-.847**
	TTC								1

Note ** significant at a 0.01 level

* significant at a level of 0.05

Participants with a higher Fixation Frequency on the Main display also had a higher Gaze Percentage on the Main display ($r = 0.811$; $p \leq 0.01$). This correlation was also true for the Alarm Banner ($r = .863$; $p \leq 0.01$), Temp & Flow ($r = .830$; $p \leq 0.01$), AOF ($r = .931$; $p \leq 0.01$), and MOF ($r = .764$; $p \leq .05$). During monitoring periods, as participants had longer gaze durations on the Main display, their % Events Complete decreased ($r = -0.839$; $p \leq 0.01$) and their Time to Faceplate increased ($r = 0.873$; $p \leq 0.01$). During active periods, as participants had higher Gaze Percentages on the Main display, the lower percentage of events were completed ($r = -0.876$; $p \leq 0.01$).

4.2 Scan Plot Qualitative Analysis

The following section describes each participant's two scan plots including the operator's actions during the 1-minute segments (searching, monitoring, task-oriented) and how the scan plot correlated with their actions. The first figure mapped each participant's eye movement 60 seconds after the setpoint failure. The second figure mapped eye behavior 30 seconds before and after C-TIC-13807 faceplate was opened after the setpoint failure. Each circle represents a fixation and lines represent saccades. The diameter of each circle represents the fixation duration. Larger diameters represent longer dwells.

Scan paths have certain patterns depending on the task. Scan paths during search tasks have a number of short fixations spread across the display; the quantity of fixations depends on how long it takes to find the object. Most fixations are short during a search task, but longer fixations indicate difficulty reading or extracting necessary information. Monitoring scan plots look similar to search task plots with fixations over a large part of the display but may often have longer fixations. During monitoring, participants may take longer to recognize that values are correct or normal rather than making a quick fixation, recognizing it is not what is being looked for and moving on as in a search task. Task-oriented scan plots have fixations mostly in isolated areas that pertain to the task. There will also be more fixations on the faceplate since most tasks involve actions using the faceplate.

Figure 5 is the scan plot for novice Participant 1 during the 60 seconds after the setpoint change. During the setpoint change, Participant 1 was monitoring the Heater display then switched to the Overview display; there were no active alarms during this period. The scan plot is representative of a monitoring scan pattern with several short fixations over the entire display as well as some longer fixations on certain components. Figure 6 is the scan plot during the 30 seconds before and after C-TIC-13807 faceplate was opened. At this point, Participant 1 had noticed the setpoint change and opened the faceplate to fix it. The scan plot has very localized fixations on the areas relevant to the Crude Temperature (fuel sources, Temp & Flow AOI, C-TIC-13807, and faceplate). Participant 1 was triggered to open C-TIC-13807 faceplate after answering the Situation Awareness questions. There is a question that asks “Which component’s setpoint drastically dropped to 400?” Participant 1 vocalized in the video that she “knew what to do” after reading this question. She had been having trouble with the fuel sources and could not diagnose why they were shutting down. After closing the Situation Awareness questions, Participant 1 went directly to the Heater display and looked at C-TIC-13807 setpoint. This is why Participant 1’s eye movement are very deliberate and localized typical of a task-oriented scan pattern.

Figure 5 - Participant 1: 60 seconds after setpoint failure

Figure 6 - Participant 1: 30 seconds before and after opening FP

During the 60 seconds after the setpoint change for novice Participant 2, there were two active alarms (C-FAL-13833 and C-XA-2047). The scan plot illustrates Participant 2 searching for C-FAL-13833 (Figure 7). There are also many fixations at the second alarm in the Alarm Banner where C-FAL-13833 alarm was located. Figure 8 illustrates the scan plot 30 seconds before and after opening C-TIC-13807. Participant 2 was triggered to open the faceplate from the Alarm Banner. After clicking on the alarm, the Tower Bottom display opened and the operator had to search for C-TIC-13807. This scan plot is representative of a searching scan pattern. Once the faceplate was opened, Participant 2 acknowledged the alarm and opened the process history view.

Figure 7 - Participant 2: 60 seconds after setpoint failure

Figure 8 - Participant 2: 30 seconds before and after opening FP

Figure 9 is another scan plot representative of a search scan pattern; there are several short fixations over the entire display. There were two active alarms present during the 60 seconds after the setpoint failure (C-FAL-13833 and C-XA-2047). Novice Participant 3 was searching for C-FAL-13833. After locating the alarm on the display, the operator acknowledged the alarm and then proceeded to investigate the fuel oil and gas components which is why there is a cluster of fixations in the lower left corner. During the 1-minute period plotted in Figure 10, the Alarm Banner was full with alarms; the scan plot has fixations along the length of the Alarm Banner. Participant 3 was prompted to open C-TIC-13807 faceplate from the Alarm Banner. It does not take the operator long after opening the display to locate C-TIC-13807 (not many fixations over the display). After opening the faceplate, the operator acknowledges the alarm and tries to change the setpoint from the SP slider several times but the component is in the wrong mode so it will not change. After several attempts to change the setpoint, the operator abandons the task and turns her attention to the other alarms in the Alarm Banner.

Figure 9 - Participant 3: 60 seconds after setpoint failure

Figure 10 - Participant 3: 30 seconds before and after opening FP

During the 60 seconds after the setpoint failure, novice Participant 4 was adjusting the flow for C-FIC-10125 (Figure 11). The scan pattern is task-oriented having most fixations on the faceplate and task relevant components. While waiting for the flow to adjust for C-FIC-10125, the operator does glance down to look at the crude temperature. After closing the faceplate for C-FIC-10125, Participant 4 starts to investigate the crude temperature setpoint failure (Figure 12). The operator looks at the inlets and outlets on the far left side of the screen and then opens the faceplate for C-TIC-13807. The operator only opened the faceplate for C-TIC-13807 after noticing the setpoint; there were no active alarms to trigger opening the faceplate.

Figure 11 - Participant 4: 60 seconds after setpoint failure

Figure 12 - Participant 4: 30 seconds before and after opening FP

There were two active alarms (C-FAL-13833 and C-XA-2047) present when the setpoint changes for novice Participant 5. The operator had already searched for the alarming components and had the alarm screen up when the setpoint changed. The operator acknowledged the alarms from the alarm display. After closing the alarm display, the operator searched for C-FIC-13803 which is illustrated in Figure 13. It is indicated that the operator understood the relationship between the two alarms and the fuel sources since he immediately searched for C-FIC-13803 (Fuel Oil) after acknowledging the alarms. After opening C-FIC-13803, the operator looked down at C-TIC-13807 which is affected by the fuel sources. There were 2 long fixations on the setpoint value for C-TIC-13807 which may have indicated

that the operator had difficulty understanding the meaning of a 400°F setpoint. Regardless, the operator did open the faceplate for C-TIC-13807 after noticing the setpoint. Participant 5 was quick to notice the setpoint change so the two scan path plots do overlap about 30 seconds. Figure 14 illustrates the latter part of the searching for C-FIC-13803 and picks up when the faceplate for C-TIC-13807 is opened. Participant 5 was prompted to open C-TIC-13807 after investigating the shutdown to fuel oil alarm indicating an understanding of the relationship the fuel sources have with the crude temperature.

Figure 13 - Participant 5: 60 seconds after setpoint failure

Figure 14 - Participant 5: 30 seconds before and after opening FP

Figure 15 plots a task-oriented scan pattern. Novice Participant 6 was addressing C-FIC-10128 flow problem when the setpoint changed. Most fixations are on the faceplate; the operator was staring at the SP slider waiting for the value to rise. The operator did look away from the faceplate and fixated on C-TIC-13807. The operator noticed the setpoint change and after the value for C-FIC-10128 reached its normal value, the operator immediately opened the faceplate for C-TIC-13807. Figure 16 is another task-oriented scan pattern. Most of the operator's fixations were on the faceplate of C-TIC-13807 investigating the setpoint and how to address it. The operator opened the Process History View to find the correct setpoint but was hesitant to change the setpoint. The operator's actions are interrupted when an active alarm appears in the Alarm Banner.

Figure 15 - Participant 6: 60 seconds after setpoint failure

Figure 16 - Participant 6: 30 seconds before and after opening FP

The two scan plots for expert Participant 7 (Figure 17 and 18) are the same because the 60 seconds after the setpoint change are the same time as 30 seconds before and after opening the faceplate. There were no alarms active when the setpoint changed for Participant 7. The operator was monitoring values on the Overhead Receiver display when the setpoint changed. The operator then opened the Heater display and monitored the fuel sources. After looking at the fuel sources, the operator's gaze turned to the Temp & Flow AOI particularly the Crude temperature (C-TIC-13807). The operator noticed the setpoint and opened the faceplate. After checking the Process History View for the correct setpoint value, the operator changed the setpoint to 650°F. Participant 7 was prompted to open C-TIC-13807 through monitoring, but both scan plots are characteristic of task-oriented scan patterns.

Figure 17 - Participant 7: 60 seconds after setpoint failure

Figure 18 - Participant 7: 30 seconds before and after opening FP

The scan plot in Figure 19 begins while expert Participant 8 was monitoring the Heater components. Thirteen seconds into monitoring, two alarms called the operator's attention to the Alarm Banner. When the operator clicked on the alarms, the fuel alarm display opened and the operator acknowledged the alarms. The Heater display was then opened and the operator focused attention on the two fuel sources. After looking around the fuel source area, the operator opened C-FIC-13803 (Fuel Oil) faceplate. The scan plot in Figure 20 picks up at the end of the first scan plot. The operator opened the faceplate for C-TIC-13807 from the bottom of C-FIC-13803 faceplate. Both scan plots are characteristic of task-oriented scan patterns. The operator knew there was a problem somewhere within the cascade loop and focused attention between the fuel sources, C-TIC-13807 and faceplate.

Figure 19 - Participant 8: 60 seconds after setpoint failure

Figure 20 - Participant 8: 30 seconds before / after opening FP

Both scan plots for expert Participant 9 (Figures 21 and 22) are the same since the operator noticed the setpoint change within 30 seconds. The setpoint changed while the operator was monitoring and read off the five monitoring items, which C-TIC-13807 is one of them. The operator noticed the setpoint change while monitoring. Right after reading off the last value on the Tower Top display, the operator immediately opened the Heater display and opened C-TIC-13807 faceplate. The operator immediately changed the setpoint to 650°F. The operator opened the Process History View to check the correct value after he had already changed the setpoint. After closing C-TIC-13807 faceplate, the operator looked to the fuel sources to check their values.

Figure 21 - Participant 9: 60 seconds after setpoint failure

Figure 22 - Participant 9: 30 seconds before and after opening FP

CHAPTER 5. DISCUSSION

The objective of the current research was to investigate eye behavior of expert and novice control room operators by analyzing and comparing Fixation Frequencies, Gaze Duration Means, Gaze Percentages per AOIs and scan paths. The criterion variables were assessed regarding expertise, interface type and alarm activity. It was predicted that experts would have higher fixation frequencies than novices, experts would have shorter mean dwell durations than novices, and experts would have more optimal scan paths than novices.

5.1 Model 1

The original hypothesis presumed that experts would have higher fixation frequencies and shorter mean dwell durations than novices. The MANOVA for Model 1 revealed no significant differences due to expertise, interface or alarm activity which is not surprising because of the small sample size and large number of variables. The MANOVA also had extremely low power, which is a consequence of the small sample size. The quantitative analysis would have benefited from a larger sample size, but due to the tediousness of the eye-tracking analysis it was not feasible.

The single ANOVAs did reveal two significant results. Both expert and novice participants had a significantly higher Fixation Frequency on the Main display during monitoring periods than during active alarm periods. This means that during inactive periods when there were no alarms in the Alarm Banner, participants spend more time looking at the Main display than when there were alarms in the Alarm Banner. This is supported by the findings for Gaze Percentage, which showed that participants had a significantly higher Gaze Percentage on the Main display during monitoring periods. This is expected because when there are active alarms, participants' attention is divided among the Alarm Banner, faceplate, and Main display. When there are alarms, participants mostly focused on opening faceplates, acknowledging alarms through the faceplate and making process adjustments through the faceplate as well as looking up steady-state values in the Process History View. These are all aspects that take attention away from the Main display during active alarm times which explains why the Fixation

Frequency and Gaze Percentage on the Main display would be lower during active times. The results from the non-parametric analysis also coincide with the single ANOVA results. Participants had a significantly higher Fixation Frequency and Gaze Percentage on the Main display during monitoring periods than during active alarm periods.

There is no statistic to show that the faceplate is where the participant's attention is since the experiment was designed so that the faceplate AOI's (AOF, MOF, and ROF) were not visible during monitoring times, but after reviewing the eye-tracking videos, there is evidence that participants fixated on the faceplate during active alarm periods. Figure 23 shows a scan plot of 30 seconds before and after opening C-TIC-13807 faceplate for Participant 3. The Alarm Banner was full of alarms (as indicated by the fixations along the length of the Alarm Banner) and within the 1-minute segment, most fixations were on the faceplate and the Alarm Banner. The scan plot illustrates how the participant's attention had many more demands during active alarm periods which drew attention away from the Main display.

Figure 23 - Participant 3 scan path: 30 seconds before and after opening C-TIC-13807 faceplate

5.2 Model 2

The MANOVA for Model 2 did not reveal any significant difference for Fixation Frequency, gaze duration mean, and Gaze Percentage for Main, Alarm Banner, Temp & Flow, AOF, MOF, and ROF due to expertise and interface type. This is mostly attributable to the small sample size. The single ANOVA for expertise did reveal that Fixation Frequency: ROF and Gaze Percentage: Temp & Flow approached significance. Novices had a higher Fixation Frequency on ROF (readings on faceplate) than experts. The ROF AOI contains the PV, the setpoint slider, and the output. The ROF is where the participant looks to find the current process values for the component.

The second variable to approach significance was Gaze Percentage: Temp & Flow. Experts had a higher Gaze Percentage on Temp & Flow revealing that experts allocated more of their attention to the crude temperature and crude flow readings than novices. The Temp & Flow AOI contains two of the most critical process parameters (crude temperature and crude flow) which is why they were located on all displays except the overview in the bottom right corner. If one of these parameters is operating outside of normal conditions, operators have a hint that there is a problem somewhere in the system. Expert operators would have this knowledge and understand the importance of keeping crude temperature and crude flow within normal bounds. This is supported by previous findings that experts scan to vital instruments more frequently (Bellenkes et al., 1997). In a study on surgeon vigilance, experts glanced at the anesthetic monitor more during times when the patient was stable than novices did (Zheng et al., 2011). Novice surgeons tended to focus more on the surgical task (Zheng et al., 2011). In another study on expert and novice eye behavior while driving, experts allocated more attention to areas of high-risk (Fisher & Pollatsek, 2007). This could explain why expert operators allocated more attention than novices did to the critical crude temp and flow values.

5.3 Correlation

As participants had a higher Fixation Frequency on the Alarm Banner, they also had a higher Fixation Frequency on the alarm area on the faceplate. All alarms had to be acknowledged and were

acknowledged through the faceplate, so it is expected that as participants fixated more on the Alarm Banner, they would also have more fixations on the AOF. Participants that looked more frequently at the crude temperature and flow controllers also looked less frequently at the readings on the faceplate. A similar correlation showed that participants that spent a larger percentage of time looking at the Temp & Flow looked less frequently at the readings on the faceplate. This is consistent with expert eye behavior. Experts had higher Gaze Percentages on the Temp & Flow and lower on the ROF. Experts looked more frequently to the Temp & Flow because of its importance. The readings on the faceplate tell the operator what the component's current value is but once the expert operators saw the values, they did not frequently look back to check the readings.

As participant's gaze durations increased on the mode buttons on the faceplate, their gaze lengths also increased on the readings (PV, SP, and Output) on the faceplate. The mode buttons were directly next to the values and level readings on the faceplate and these areas were where most attention was focused on the faceplate (MOF and ROF had the highest Gaze Percentages of the faceplate AOI's). This could imply that operators had a consistent style of extracting information from the faceplate. The longer their fixations were looking at the modes buttons, the longer their fixations were when looking at the readings. Longer fixations on the faceplate likely meant that the operator was unsure about current values and hesitant toward making decisions about changing those values. Shorter fixations indicated a more efficient extraction of information (Bellenkes et al., 1997) and confidence in their decision. The correlation between gaze duration on the MOF and ROF indicates a consistency between the operator's ability to extract information from important areas on the faceplate.

Participants that had longer Gaze Durations on MOF had a lower Fixation Frequency on the Alarm Banner and a lower percentage of events completed. Longer gazes represent difficulty extracting information which meant operators were either unsure of the current mode or unsure what mode the component should be placed into. Lower fixations on the Alarm Banner indicate the perceived importance of the Alarm Banner to the operator. It makes sense that operators with uncertainty of mode

and low perceived importance of the Alarm Banner would have a lower performance indicated by the lower number of events correctly completed.

There was a strong correlation between Fixation Frequency and Gaze Percentage. As participants' Fixation Frequency on the Main display increased, their Gaze Percentage on the Main display also increased. This was also true for the Alarm Banner, Temp & Flow, AOF, and MOF which is five of the six AOI's. This has strong implications towards the necessity of using both metrics for analysis since one of the downfalls to using solely Gaze Percentage is that it is difficult to tell whether fixations were quick and frequent or longer but fewer. Gaze Percentage is a function of dwell duration, but Gaze Percentage was highly influenced by Fixation Frequency. As the number of fixations went up, the Gaze Percentages increased as well. This meant that there was a large number of fixations having small durations. It is important to know which type of fixations are represented because they mean two different things. By testing for correlation between the dependent measures (Fixation Frequency, Gaze Duration Mean, and Gaze Percentage per AOI), it was possible to discover the type of fixations. Since Gaze Percentage and Fixation Frequency for specific AOIs were correlated, most fixations on those AOIs were short and frequent. Fixation Frequency and Gaze Duration Mean are the most common statistics used in eye-tracking analyses because they reveal the importance and ease of interpretation of certain elements (Fitts et al., 1950) but Gaze Percentage reveals important fixation information in relation to other AOIs. Gaze Percentage has typically been used to explore attention allocation. Schriver et al. (2008) used Gaze Percentage to investigate attention allocation to critical areas among aviation pilots but this was the sole eye-statistic used and did not explore what the type of fixations meant. In the current research, the positive correlation between Fixation Frequency and Gaze Percentage identified that participants had a higher number of quick fixations. It can be concluded that when using Gaze Percentage to investigate attention allocation, Fixation Frequency and Gaze Duration should also be included for more revealing results.

Participants that had longer gazes on the ROF also had higher Gaze Percentages on ROF. ROF was the one AOI that Gaze Percentage and Fixation Frequency were not correlated indicating that fixations on the readings on the faceplate were not quick and frequent. Instead, Gaze Percentage was influenced by Gaze Duration meaning participants that spent more time looking at the readings mostly stared at them. This was noticed during testing that some participants would stare at the PV to wait for the value to fall within acceptable range.

Participants that had longer gazes on the Main display while monitoring completed a lower percentage of events correctly and took them longer to open the faceplate. Participants that looked longer at the Main display may have had difficulty finding controllers to open the correct faceplates and had difficulty interpreting the process components on-screen. If participants had trouble understanding how components were related to each other, they would have had difficulty correctly completing events which would yield a lower completion percentage.

During active periods, as participants spent more time looking at the Main display, they were less accurate and completed a lower percentage of events. Participants that spent more time looking at the Main display while there were active alarms present may indicate that they had difficulty diagnosing the problem and were confused on how to fix it. They spent a high percentage of time looking around the Main display trying to find the cause of the problem. The more difficulty participants had diagnosing failures, the fewer number of events they were able to correct.

Also as participants had longer gazes on the mode buttons on the faceplate, they had a higher Gaze Percentage on the Main display. Participants that had long fixations on the mode buttons showed signs of hesitation toward either changing the mode or confusion toward the correct mode. Also, participants that had longer gazes on the MOF also correctly addressed fewer failures ($r = -0.757$; $p = .05$). As previously stated, participants with higher Gaze Percentages on the Main display also had lower completion percentages because they had difficulty diagnosing problems. This explains why participants

with lower completion percentages would have longer gazes on the mode buttons and spend more time looking at the Main display.

Overall, correlation data supported previous research conducted by Fitts et al. (1950) that showed that fixation frequency is associated with the perceived importance on the component and gaze duration is indicative of difficulty interpreting information. Experts had a higher fixation frequency on the Temp & Flow AOI which corresponds to the high significance of these process parameters. While monitoring, the participants with longer gaze durations on the Main display completed a lower number of events correctly which indicates a difficulty extracting and understanding items on-screen.

5.4 Setpoint Change

Seventeen minutes into the hard workload level scenario, there was a subtle setpoint change that occurred for C-TIC-13807 from its regular value of 650°F to 400°F. When the setpoint changes to 400°F, the fuel sources (oil and gas) decrease flow to allow the crude temperature to decrease to 400°F. When the fuel sources decrease to 0 GPM (valve is closed) the normal operating conditions are violated since the valves should never remain closed during normal operations. Not immediately, but within 1 – 2 minutes, this violation triggers C-XA-2047 and C-FAL-13833 alarms which indicate that there is no flow.

Both scan paths for each participant centered on this setpoint change. For comparison of scan plots, participants were split into two groups: those that noticed the setpoint change and opened C-TIC-13807 faceplate within 2 minutes and those that opened the faceplate after 2 minutes from the setpoint change (See Table 25 for elapsed times). All experts (participants 7, 8, and 9) and novice participants 4, 5, and 6 recognized the setpoint change and opened C-TIC-13807 faceplate within 2 minutes.

Monitoring percentages were investigated in order to explain why some participants were able to catch the setpoint change early on (within two minutes). During the simulation, participants were asked to monitor and read off to the researcher five values every five minutes. C-TIC-13807 was one of those values. Participants with higher monitoring percentages may have looked to the C-TIC-13807 controller

Table 25 - Participant Elapsed Time to Open Faceplate and Correct Setpoint

<u>Participant</u>	<u>Opens C-TIC-13807*</u> (min:sec)	<u>Corrects Setpoint**</u> (min:sec)
1	9:41	9:55
2	11:05	11:48
3	14:51	--
4	1:19	--
5	0:51	1:14
6	1:05	--
7	0:29	0:47
8	1:23	2:16
9	0:28	0:34

Note. *Elapsed time from the setpoint event until C-TIC-13807 is opened

**Elapsed time from the setpoint event until the setpoint is corrected to 650°F

-- Did not change setpoint

more often and could explain why they were able to catch the setpoint change quickly. After looking at the monitoring percentages, there was no relationship between the elapsed time to open the faceplate and the monitoring completion percentages. All participants had percentages of 100 except Participants 3 and 7 with percentages of roughly 80.

Experts looked at the Temp and Flow AOI for a significantly larger percentage of time than novices did. This may explain why all three experts were able to detect the setpoint change so quickly. Two of the experts detected the change within 30 seconds. The larger percentage of attention to the Temp & Flow area indicated that experts were aware of the significance of these values and allocated attention to them effectively.

After reviewing the eye-tracking videos, it was evident that those that discovered the setpoint within two minutes found it through their own monitoring and diagnosis whereas all three participants that it took longer than 9 minutes to find the setpoint change were triggered by the C-TIC-13807 alarm in the Alarm Banner or the Situation Awareness questions. Participant 1(novice) was only prompted to change the setpoint after reading the situation awareness question that referred to that failure. This participant used it as an aid to diagnose the problem. Participants 2 and 3 (novices) both opened the faceplate and found the setpoint change only because of the alarm in the banner. Through other process

conditions within the system, the crude temperature had risen to an alarming level that triggered the C-TIC-13807 alarm. Participants 2 and 3 had several alarms in the Alarm Banner and only investigated the setpoint after being prompted to open the faceplate by the alarm.

The task that the operator was doing when the setpoint change occurred also seemed to affect how quickly the setpoint was detected. All three experts were actively monitoring the system when the change occurred. Two of the three novice participants (4 and 6) that detected early were completing a task when the setpoint changed. They were both adjusting flows for other components and while waiting for values to reach a goal value, both participants scanned down to the C-FIC-13807 controller. It is apparent that they noticed the setpoint change because immediately after finishing their original task, they opened C-TIC-13807 faceplate. They demonstrated a scan pattern typical of experts. They allocated attention to the high-risk area even while in the middle of a task. This is unlike novices who tend to intently focus on the task at hand (Zheng et al., 2011). Participants 2 and 3 who found the change late were searching for specific components that needed alarms to be acknowledged. They were engrossed in the search task and did not notice the change.

Interface type also had an effect on early detection of the setpoint change. Four of the five participants using the black interface discovered the setpoint change through some type of active alarm in the Alarm Banner. Two of the participants that detected early on, diagnosed that there was a problem with C-TIC-13807 from the fuel shutdown alarm (C-FAL-13833). This alarm is triggered within 1-2 minutes from the setpoint change. This alarm points the participant to the fuel sources. It was the participant's understanding of the system that triggered them to open C-TIC-13807 and it was then that the setpoint change was noticed. The two participants that took 11 and 15 minutes to open C-TIC-13807, did not realize the setpoint had changed until the crude temperature alarmed. All three participants using the grey interface that caught the setpoint change early detected it because they had looked at the C-TIC-13807 controller and noticed the change. It may have been easier to detect the change (without a cue from an alarm) on the grey interface rather than the black interface because it may have been more visible

and easier to notice. Previous research conducted by Michalski and Grobelny in 2008 investigated the effects that object background color had on search times. They found that certain color contrasts had an effect on task time as the objects were less salient but they found higher contrasting colors had shorter task times over the grey panel. This is contrary to the current research findings which found the grey interface to be more salient. This difference is most likely due to the selection of contrasting colors. Compared to the highly contrasting neon colors in the current study, the colors used in 2008 are of a medium more optimal contrast.

It should also be noted that the three novices that detected the setpoint change early were all students from South Central Louisiana Technical College in Reserve, La. The students from Lambton College in Sarnia, Ontario took the longest to notice the setpoint change. It was also the two novice participants from SCLTC that demonstrated eye behavior similar to expert behavior (i.e. attention allocation). Operator behavior and monitoring style may also explain the difference. Participants 1, 2, and 3 had a more relaxed monitoring style. When nothing was happening, they would have the Overview display up and monitor from there; Participant 1's 60 sec after setpoint change scan plot is when there are no alarms and the operator is watching the Overview. Participants 4, 5, and 6 had a more active monitoring style similar to the behavior of the experts as they flipped through displays to monitor and diagnose problems. Active monitoring style also contributed to early detection of the setpoint change.

When looking at eye-behavior of such a small sample size, it is important to also include a qualitative analysis. Quantitative analyses are useful because they are objective and can give direction to a qualitative analysis, although on a small sample size can yield results with very low power. Qualitative analyses are easier to perform but may not hold as much strength as a statistical analysis since they are more subjective to the reviewer. Sullivan et al. (2011) included both types of analysis to investigate the influence of flight expertise on visual scan patterns of helicopter pilots. Scan plots were explored to explain correlation differences between eye scan data. Similarly in the current study, the weaker quantitative results were strengthened by a qualitative analysis of the scan plots. The scan plots were able

to support some of the statistical results as well as provide additional findings. It is recommended when working with a small sample to include a qualitative analysis as well to strengthen findings.

5.5 Future Research and Limitations

Future research can investigate specific search patterns of expert and novice operators. This research focused on scan paths centered on a single event rather than a specific search task. Studying where experts start searching and how they move between objects can help designers create displays that support operator search patterns. It also would be beneficial to look at how monitoring strategies differ between expertise levels. Looking at eye behavior and scan plots while operators monitor a process may reveal what information is attended to and how often operators scan back to recheck values.

Limitations include experiment design which only used one display monitor for the simulation which may have hindered the external validity since operator control rooms in industry have multiple monitors displaying process information. The experimental monitor only allowed the operator to have one screen open at a time. The study was also limited by the simulation software. There were technical issues within the software that caused bogus alarms to trigger occasionally and may have caused additional stress or confusion to the participant thus altering the workload level and affecting internal validity. The experiment was also limited by the use of the eye-tracking glasses. Some participants did not yield good eye-tracking data due to eye glasses or certain eye features and were excluded from the analysis. Only participants with a tracking ratio of 95% or higher were included. The eye-tracking equipment was still a limiting factor because the actual tracking of the eye was not as accurate as it should have been. There were times when fixations were tracked 1-2 centimeters off of where the participant was actually looking. Most likely, only the statistics for the faceplate AOIs (AOF, MOF, and ROF) were affected since the AOIs were small and close together. Due to the proximity of the AOIs and accuracy of the eye tracker some fixation data may have been recorded for the wrong AOI; although it is unlikely that the results were significantly impacted. The largest limitation was the small sample size of expert operators and only having one expert operator with eye-tracking data for the grey interface. This

weakened the MANOVA results. Also, due to the small sample size interaction effects could not be measured. It would have been interesting to study the differences between how experts and novices behaved on each interface.

CHAPTER 6. CONCLUSION

6.1 Expertise

Overall, expert and novice operators did have different strategies toward monitoring the crude refinement process. Eye-tracking revealed that expert operators scanned more frequently to critical areas. Most notably, this facilitated the early detection of the setpoint change and follows previous research showing that a characteristic of experts is to scan to high-risk areas more often (Zheng et al., 2011). With a higher Gaze Percentage on the Temp & Flow, experts demonstrated their ability to allocate attention more effectively. This is applicable toward novice training. Training novice operators to scan to high-risk areas often could lead to earlier detection of failures and malfunctions.

Encouraging an active monitoring style may also lead to early detection. Operators that waited on the Overview Display for alarms to trigger and relied on the alarm banner to direct their actions took longer to detect the setpoint change. It was the operators that actively monitored and took action when they noticed something was wrong before alarms were triggered that detected the setpoint change early. It is necessary to note that this simulation was only 30 minutes long which allows the operator to actively monitor the entire time. It is unrealistic to expect this same behavior from operators working an 8-12 hour shift. Rather than having operators wait for alarms to trigger action, one solution may be to have operators do a system “walk-through” intermittently and monitor process values on a scheduled basis to improve early detection.

6.2 Interface Design

Eye-tracking also revealed that the operator’s attention was in much higher demand when there were alarms present. This was expected since operators had to divide their attention between the Main display, alarms in the Alarm Banner, and values on the faceplate. With such a demand on the operator, the interface should be designed in a way that adds no additional strain on the operator’s attention. It may be concluded that display color scheme played a role in reducing the attentional demand on the operator. Operators using the grey, more muted color scheme, were able to detect the setpoint failure more quickly

than operators using the black, high contrast color scheme display. The grey color scheme allowed the subtle setpoint change to grab the attention of the operator more quickly. The highly contrasting colors on the black display may have been distracting thereby diverting the operator's attention. Although there are other contributing factors, display color scheme did play a role in early detection.

In conclusion, eye-tracking metrics were capable of detecting differences between the strategies of expert and novice control room operators while monitoring a Crude Refinement simulation. Expertise, interface background color, and alarm activity affected how operators interacted with the process which have strong implications toward effective novice training and interface design.

REFERENCES

- Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, 51(4), 355-365.
- Bellenkes, A.H., Wickens, C.D., & Kramer, A.F. (1997). Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. *Aviation Space and Environmental Medicine*, 68(7), 569-579.
- Burns, C. M. (2000). Navigation strategies with ecological displays. *International Journal of Human-Computer Studies*, 52(1), 111-129.
- Cellier, J. M., Eyrolle, H., & Marine, C. (1997). Expertise in dynamic environments. *Ergonomics*, 40(1), 28-50.
- Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, 27(8), 951-964.
- Chase, W. G., & Simon, H. A. (1973). PERCEPTION IN CHESS. *Cognitive Psychology*, 4(1), 55-81.
- Cooke, L. (2005). Eye tracking: How it works and how it relates to usability. *Technical Communication*, 52(4), 456-463.
- Cyr, D., Head, M., & Larios, H. (2010). Colour appeal in website design within and across cultures: A multi-method evaluation. *International Journal of Human-Computer Studies*, 68(1-2), 1-21.
- D'Souza, M. E., & Greenstein, J. S. (1999). The design of a visual display for the presentation of statistical quality control information to operators on the plant floor. *Human Factors*, 41(4), 619-627.
- Ebbesen, E. B., & Konecni, V. J. (1975). Decision making and information integration in the courts: The setting of bail. *Journal of Personality and Social Psychology*, 32(5), 805-821.
- Ericsson, K. A. (2005). Recent Advances in Expertise Research: A Commentary on the Contributions to the Special Issue. *Applied Cognitive Psychology*, 19(2), 223-241.
- expert. (n.d.). In Merriam-Webster.com. Retrieved March 22, 2013, from <http://www.merriam-webster.com/dictionary/expert>.
- Fisher, D.L., & Pollatsek, A. (2007). Novice Driver Crashes: Failure to Divide Attention or Failure to Recognize Risks. In A. F. Kramer, D. A. Wiegmann & A. Kirlik (Eds.), *Attention: From Theory to Practice* (pp. 134-156). Oxford, UK: Oxford University Press.
- Fitts, P.M., Jones, R.E., & Milton, J.L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review*, 9(2), 24-29.
- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, 23(4), 523-552.

- Glaser, R., Chi, M.T.H., Farr, M.J. (1988). *The Nature of Expertise*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, (xv-xxiii).
- Goldberg, J.H., & Kotval, X.P. (1998). Eye movement-based evaluation of the computer interface. In S. K. Kumar (Ed.), *Advances in Occupational Ergonomics and Safety*. Amsterdam: ISO Press.
- Gramopadhye, A. K., Melloy, B. J., Nair, S. N., Vora, J., Orhan, C., Duchowski, A. T., & Kanki, B. (2002). The use of binocular eye tracking in virtual reality for aircraft inspection training. *International Journal of Industrial Engineering-Theory Applications and Practice*, 9(2), 123-132.
- Hampton, D. C. (1994). Expertise - The True Essence of Nursing Art. *Advances in Nursing Science*, 17(1), 15-24.
- Jacob, R.J.K., & Karn, K.S. (2003). Eye tracking in Human-Computer Interaction. In R. Radach (Ed.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (471-606). Amsterdam: Elsevier.
- Johannsen, G. (1997). Conceptual design of multi-human machine interfaces. *Control Engineering Practice*, 5(3), 349-361.
- Johnson, Dallas E. (1998). *Applied Multivariate Methods for Data Analysts*. Pacific Grove, CA: Duxbury Press.
- Just, M. A., & Carpenter, P. A. (1976). Eye Fixations and Cognitive Processes. *Cognitive Psychology*, 8(4), 441-480.
- Klein, Gary. (2008). Naturalistic Decision Making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3), 456-460.
- Koffskey, C., Ikuma, L., & Harvey, C. (2013). *Performance Metrics for Evaluating Petro-Chemical Control Room Displays*. Paper presented at the Human Factors and Ergonomic Society Annual Meeting, San Diego, CA, 57(1717-1721).
- Lavine, R. A., Sibert, J. L., Gokturk, M., & Dickens, B. (2002). Eye-tracking Measures and Human Performance in a Vigilance Task. *Aviation Space and Environmental Medicine*, 73(4), 367-372.
- MacDorman, K. F., Whalen, T. J., Ho, C. C., & Patel, H. (2011). An improved usability measure based on novice and expert performance. *International Journal of Human-Computer Interaction*, 27(3), 280-302.
- Mele, M. L., & Federici, S. (2012). Gaze and Eye-Tracking Solutions for Psychological Research. *Cognitive Processing*, 13, S261-S265.
- Murphy, E. D., & Mitchell, C. M. (1986). Cognitive Attributes - Implications For Display Design In Supervisory Control-Systems. *International Journal of Man-Machine Studies*, 25(4), 411-438.
- Oskamp, S. (1962). The relationship of clinical experience and training methods to several criteria of clinical prediction. *Psychological Monographs: General and Applied*, 76(28), 1-27.

- Ottati, W.L., Hickox, J.C., & Richter, J. (1999). *Eye scan patterns of experienced and novice pilots during Visual Flight Rules navigation*. Paper presented at the Human Factors and Ergonomics Society, Houston, Texas, 43(66-70).
- Poole, A., & Ball, L. J. (2005). Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future. In C. Ghaoui (Ed.), *Encyclopedia of Human-Computer Interaction* (211-219). Pennsylvania: Idea Group, Inc.
- Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors*, 49(3), 347-357.
- Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, 50(6), 864-878.
- SensoMotoric Instruments. (2011a). SMI BeGaze Manual.
- SensoMotoric Instruments. (2011b). SMI EYE TRACKING GLASSES: WORLD'S FIRST BINOCULAR GLASSES-TYPE EYE TRACKER RELEASED. from http://www.smivision.com/en/gaze-and-eye-tracking-systems/about-smi/news/article.html?tx_ttnews%5Btt_news%5D=431&cHash=55a11b84aee94d23f929d79540e1fafe.
- Shanteau, James. (1992). The Psychology of Experts: An Alternative View. In G. Wright & F. Bolger (Eds.), *Expertise and Decision Support* (11-23). New York: Plenum Press.
- Snizek, J. A., & Henry, R. A. (1990). Revision, weighting, and commitment in consensus group judgment. *Organizational Behavior and Human Decision Processes*, 45(1), 66-84.
- Sullivan, J., Yang, J. H., Day, M., & Kennedy, Q. (2011). Training Simulation for Helicopter Navigation by Characterizing Visual Scan Patterns. *Aviation, Space, and Environmental Medicine*, 82(9), 871-878.
- Takeuchi, H., & Habuchi, Y. (2007). *A Quantitative Method for Analyzing Scan Path Data Obtained by Eye Tracker*. Paper presented at the Computational Intelligence and Data Mining, Honolulu, Hawaii, 283-286.
- Taukari, A., Pant, R. S., & Garg, A. K. (2010). A comparative study of cognitive strategies used to resolve air traffic conflict between air traffic controllers and general population using eye tracking machine. *Journal of Psychosocial Research*, 5(2), 125-132.
- Tomizawa, Y., Aoki, H., Suzuki, S., Matayoshi, T., & Yozu, R.. (2012). Eye-tracking analysis of skilled performance in clinical extracorporeal circulation. *Journal of Artificial Organs*, 15(2), 146-157.
- Tzanidou, E., Petre, M., Minocha, S., & Grayson, A. (2005). Combining eye tracking and conventional techniques for indications of user-adaptability. In M. F. Costabile & F. Paterno (Eds.), *Human-Computer Interaction - Interact 2005, Proceedings* (Vol. 3585, pp. 753-766).
- U.S. Chemical Safety and Hazard Investigation Board. (2007). BP America Refinery Explosion. <http://www.csb.gov/assets/1/19/CSBFinalReportBP.pdf>.

- U.S. Chemical Safety and Hazard Investigation Board. (2008). Pesticide Chemical Runaway Reaction Pressure Vessel Explosion. http://www.csb.gov/assets/1/19/Bayer_Report_Final.pdf.
- Vicente, K. J. (1992). Memory recall in a process control system: A measure of expertise and display effectiveness. *Memory & Cognition*, 20(4), 356-373.
- Wang, Y.S. & Hwang, S.L.. (1995). An experimental study on the information design of CRT display in process control systems. *International Journal of Industrial Ergonomics*, 15(6), 459-469.
- Wu, F.G., Chen, C.Y., Lee, Y.J., & Chen, R. (2010). Effects of color sample display and color sample grouping on screen layout usability for customized product color selection. *Computers in Human Behavior*, 26(1), 51-60.
- Zheng, B., Tien, G., Atkins, S.M., Swindells, C., Tanin, H., Meneghetti, A., Qayumi, K.A., Panton, O.N.M. (2011). Surgeon's vigilance in the operating room. *The American Journal of Surgery*, 201(5), 673-677.

APPENDIX 1. GREY AND BLACK DISPLAYS

Figure - Grey overview display

Figure - Black overview display

Figure - Grey desalter display

Figure - Black desalter display

Figure - Grey heater display

Figure - Black heater display

Figure - Grey tower bottom display

Figure - Black tower bottom display

Figure - Grey tower top display

Figure - Black tower top display

Figure - Grey overhead receiver display

Figure - Black overhead receiver display

APPENDIX 2. AOIS

Figure - AOF: Alarm on Faceplate; MOF: Mode on Faceplate; ROF: Reading on Faceplate; TAF: Temp and Flow

APPENDIX 3. SEMANTIC GAZE MAPPING

Figure - Semantic Gaze Mapping is a feature of BeGaze that allows the user to map fixations onto a reference image

APPENDIX 4. AOI EDITOR

Figure - AOI Editor is a feature of BeGaze where AOI's are drawn and shifted/ resized when necessary to match the area in the scene video.

VITA

Christina Koffskey is a native from Kenner, Louisiana. Christina received her undergraduate degree in Industrial Engineering from Louisiana State University in 2010. In August 2011, after a semester break, she was eager to enroll at her Alma Mater, Louisiana State University, to attain her Masters of Science in Industrial Engineering. Her passion for people led her to focus her degree in Human Factors. Christina expects her diploma in May 2014. Her interests include Healthcare process improvement, Human-Computer Interaction, Usability, and Interface Design.