UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ANALYSIS OF EYE TRACKING DATA TO MEASURE SITUATIONAL AWARENESS IN OFFSHORE DRILLING OPERATIONS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

JIWON JEON
Norman, Oklahoma
2018

ANALYSIS OF EYE TRACKING DATA TO MEASURE SITUATIONAL AWARENESS IN OFFSHORE DRILLING OPERATIONS

A THESIS APPROVED FOR THE
GALLOGLY COLLEGE OF ENGINEERING

BY

Dr. Ziho Kang, Chair

Dr. Randa L. Shehab

Dr. Sudarshan Dhall

This thesis is dedicated

to my parents, Nambae Jeon and Sookhee Hong,

whom I admire the most,

to my sister, Heewon Jeon,

whom I appreciate the most,

and to my brother, Seunghoon Jeon,

whom I am proud of the most.

Acknowledgements

I appreciate all those who have supported me to complete this research and thesis successfully. This thesis is not only my work in the lab, but also priceless experience that I have gained through the school years. I could not have achieved this without the help and encouragement of everyone.

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Ziho Kang. He provided me with a great opportunity to work on this project and has always been supportive, for allowing me to cultivate academic knowledge and passion. His unstinting guidance and constructive feedback improved my research ability and professional skills.

I also would like to thank my committee members, Dr. Randa Shehab and Dr. Sudarshan Dhall, for their valuable time and consideration to my research. Thanks to their advice, I was able to learn the subject in depth. I am also grateful to Dr. Saeed Salehi, Dr. Edward Cokely and Dr. Muhammad Azeem Raza for their generous support and contribution to establish a wonderful environment for this research project. I also appreciate the National Academy of Sciences for the continuous funds for the research.

Furthermore, I sincerely appreciate Ms. Nicola Manos for her tremendous help during the graduate program. I also thank my research mates Raj Kiran, Vimlesh Bavadiya, Syed Ali Mehdi, Vincent Ybarra, Saptarshi Mandal, Abdulrahman Khamaj, Salem Naeeri, Amin Alhashim, other HFS members for their help to have interesting discussion on the analysis. I would like to express my special thanks to Yuji Kim, Sekar Rachmawati, Ankita Sinha, and all of my friends for their encouragement.

Lastly, I convey my love and gratitude to my family for their constant support.

Table of Contents

List of Tables

List of Figures

Abstract

In complex, high-stakes tasks such as offshore oil and gas drilling where substantial number of monitoring parameters involve in the operation, the analyses of human operator’s situational awareness (or situation awareness, SA) become more important to avoid severe incidents initiated by the poor cognitive performance. Numerous SA measurement practices have been proposed in the previous researches, however, most of them employed the verbal and behavioral response analyses which are subjective to the researchers’ notions.

In this study, an integrated approach combining subjective measures (e.g. verbal responses) with physiological metrics (e.g. eye fixation data) was investigated to seek the benefits for SA analyses in the field of offshore oil and gas drilling. A pre-existing incident based experimental test in a high-fidelity simulator facility was designed for real-time log indicators monitoring tasks, and a set of eye tracking devices collected verbal responses and oculomotor information simultaneously during the real-time tasks. To quantify the verbal responses, scoring metrics were newly developed for this study. The metrics assigned the points to the participants’ verbal responses based on their uttered keywords (abnormal, kick or blow-out) reacting to the situation.

Quantitative statistical analyses were applied to ocular observations and verbal response scores collected from the predesigned Areas of Interests (AOIs) on the monitoring screen, using one-way ANOVA and Friedman test, respectively. The analyses provided unique and complementary insights that mapped the measures from both metrics to the level of situation awareness and that helped understand the cognitive process in time critical decision-making tasks in offshore oil and gas field. In addition

to the statistical investigation, data mining approach using time series clustering technique was introduced to group the participants' scanning pattern with respect to the temporal sequences, and to find the correlation of the scanning pattern to the quantified situation awareness.

According to the analysis results, the expertise of the participants affected on their cognitive mechanisms to identify and respond to the situations to some extent. The content and timing of the situation also served as one of the important factors to determine the level of situation awareness. The participants' scan patterns were clustered into four groups and suggested a potential correlation between the visual scanning pattern and the quantified situation awareness (i.e. verbal response scores). It was found that the vertical attending tendencies to the individual logs might lead a higher comprehension of the situation than the horizontally transitional attending tendencies between different logs.

Chapter 1: Introduction

With the recent advancement in oil and gas operations technology, the immense, complicated and interrelated information requires the concurrent analyses for immediate decision-making to prevent catastrophic incidents. Despite the automated system in the high-risk tasks, human operators' cognitive performance is still considered to be one of the most critical factors to ensure reliable operations, as we have experienced in several well explosion incidents (US Chemical Safety and Hazard Investigation Board., 2010).

One popular way to assess the operator's cognitive performance is to measure how accurately the person is aware of the surrounding situations. Previous researches observed that situation awareness (SA) of the operators is critical to the safety of the oil and gas operations (Roberts et al., 2015; Sneddon et al., 2006; Sneddon et al., 2013). It is also well known that poor SA prevents the proper decision making from the operators and diminishes the operation's efficiency in any dynamic system (Endsley, 1995).

To measure the situation awareness, several techniques such as the Situation Awareness Global Assessment Technique (SAGAT) or the Situational Awareness Rating Technique (SART) were developed, however, these methodologies require to pause the experiment in the middle of the task to ask some queries to the test participants to obtain explicit SA (Endsley et al., 1998). To reduce the hindrance to a time-critical task where immediate decision-making is necessary, this study proposes a mixed approach combining verbal response evaluation with eye tracking analyses as a possible alternative to the conventional SA measurements. When a person could express their comprehension of the situation with verbal description, our eye movements may reinforce the evidence of comprehension by visually interrogating the information simultaneously. Since the eye movements can be continuously captured during the experiment, this approach requires no pause of the test, therefore, might be effective to holistically study the SA throughout the task.

This study aims to draw a preliminary assessment of SA by employing the proposed mixed approach. To quantify the analysis outcome, the verbal responses and eye movement data collected from the pre-designed incident indicator monitoring tasks are statistically investigated using conventional parametric (within-subject design one-way ANOVA) and nonparametric tests (Friedman test). The results are then converged to understand the factors affecting the participants' performance and eventually evaluate the level of SA. In addition to the statistical assessment, a data mining technique, "Time Series Clustering (a combination of Dynamic Time Warping and Density-Based Spatial Clustering of Applications with Noise)", was applied as a supplementary method to identify the eye movement trajectories with respect to temporal sequence.

Chapter 2 provides a brief background knowledge about situation awareness and the measurements in high risk tasks. This chapter also states theoretical explanation of the data mining techniques to be employed in the visual scanning trajectory analyses. Chapter 3 presents a proposed way to conduct the experimental tasks and analyses. In this study, we develop new verbal response scoring metrics to quantify the participants' comprehension, and adapt a combinatorial time series clustering model to analyze the pattern of indicator observation. The results of analyses and the corresponding discussion are included in Chapter 4. Chapter 5 concludes the study with some recommendations for future researches.

Chapter 2: Background

This chapter introduces the situation awareness (SA) and its most common measurements. Two main algorithms of the time series clustering method are introduced to provide an idea about the adapted scan pattern identification model described in Chapter 3.

2.1 Situational Awareness

Situation awareness (SA) is generally defined as a combination of three cognitive levels (Endsley, 1995). First level includes "perception" of the surrounding events, which indicates the acceptance of the presented information. In this study, perception can be considered to relate a participant's attention to the given situation or to acknowledge the incident. Second level is "comprehension" of the observed events, which is associate with the understanding and retention of the accepted information. Third level is regarded as "projection" to make correct decisions based on the previous levels of SA.

Among several measures to evaluate SA, the Situation Awareness Global Assessment Technique (SAGAT) and the Situational Awareness Rating Technique (SART) are most widely used (Endsley et al., 1998; Selcon and Taylor, 1990). In SAGAT, the experiment is paused at random times to provide several queries to the participants to assess their knowledge of the test at the point when the experiment is frozen. The queries are customized to the tasks and typically include the contents to measure the three levels of SA. Participant's verbal or written answers are compared with "ground truth" information or "expert's answers" and scored. This method is an objective, unbiased measure in that the participants cannot anticipate the queries arbitrarily asked during the task. However, it is intrusive since it requires the simulation experiment to be frozen to collect the participant's responses. SART is a post-trial multidimensional scaling technique based on the operator's subjective opinion. In SART method, the participants are asked to scale the degree of their perception to what they observed during the task. The rating is performed after the test, and each dimension of the queries are usually scaled by seven points (1 = low, 7 = high). The queries cover general contents, therefore need not be customized. Original SART was composed of ten dimensions to measure the test participant's SA. Unlike SAGAT, this method does not disturb the simulation experiment, however, it embeds the problems of post-trial information gathering process. For example, participants may forget some periods of the task when they had a poor level of SA while relatively well remember the periods when they showed a better level of SA.

A variety of studies revealed that high risk tasks depend the task eligibility and reliability on the SA of the operators (Bhavsar et al., 2017; Deacon et al., 2010; Härtel, et al., 1989; Helmreich, 2000; Sanfilippo, 2016; Sharma et al., 2016; Williams et al., 2013). The task domain includes aviation, offshore oil and gas operation, process plant control room management, clinical operations, etc. Most researches employed the traditional behavioral measures to investigate the participants' SA, however, recent studies started adapting the physiological measures such as eye movement analyses and claimed that the utility of eye tracking devices provided objective and qualitative approaches to assess the initial perception level of SA (Bhavsar et al., 2017; Sanfilippo, F., 2016; Sharma et al., 2016; Williams et al., 2013;).

2.2 Time Series Clustering

Clustering is a data mining technique to group a set of objects based on their similarity to each other without prior knowledge of the group's nature (Rai and Singh, 2010). The objects in the same group (i.e. "cluster") have the maximum similarity with other objects within the group. Clustering has been extensively used in data analytics to identify the structures in unlabeled data and organize the dataset into smaller subsets for in-depth analyses (Aghabozorgi et al., 2015).

Time series clustering is a special type of clustering to accommodate temporal elements into similarity evaluation when grouping the data. If a sequence of data is enumerated with respect to time, the data is considered "time series". Since time series data is a prevalent type in real world, clustering time series data becomes a ubiquitous analysis in various fields (Liao, 2005). Researchers found that time series data was generally characterized with high dimensions and large size, and that it was difficult to measure similarity to determine clusters due to noise, shifts or different sequence length of each observation (Keogh and Kasetty, 2003; Lin et al., 2004; Rani and Sikka, 2012). Hence, time series data limits the implementation of conventional clustering techniques designed for static data.

The strategy to perform time series clustering to a set of data can be presented in three primary approaches (Aghabozorgi et al., 2015).

- Shape-based approach: original time series data are paired up and compared to match the shape by stretching or shrinking the time axes. Conventional clustering methods are applied afterwards with some modifications in the similarity measurement.

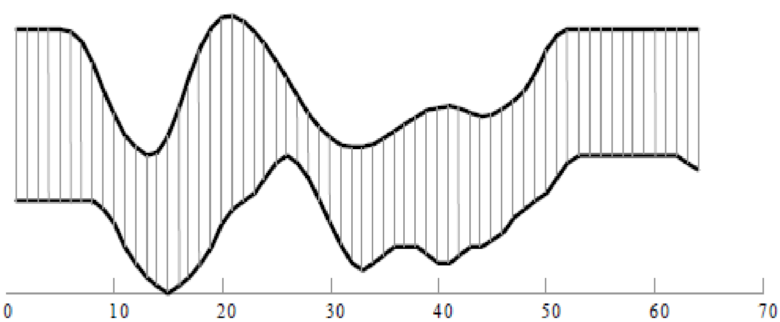
- Feature-based approach: original time series data are transformed into features of lower dimension. Conventional clustering algorithms are applied to the newly computed features later.
- Model-based approach: original time series data are transformed into model parameters (i.e. a parametric model for each dataset). A feasible model distance and conventional clustering algorithm are chosen, and then can be applied to the extracted model parameters.

Due to the ease of application, the shape-based approach is widely used in time series clustering. In this study, Dynamic Time Warping algorithm was chosen among the various techniques for modifying the time axes, and a density-based clustering method was selected as a conventional clustering tool.

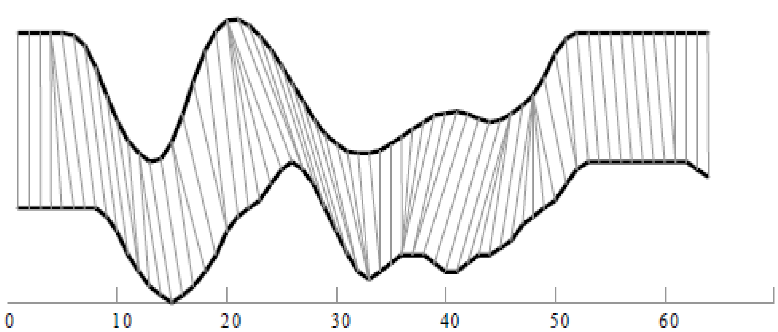
2.2.1 Dynamic Time Warping

Dynamic Time Warping (DTW) is a very useful technique to compare the similarity between two different temporal sequences. The algorithm was originally discussed in the speech recognition studies to normalize the nonlinear fluctuation in speaking rates, hence to match the words and utterances by overcoming variations in timing and pronunciations (Sakoe and Chiba, 1978). DTW was then introduced to the data mining field for reducing the sensitivity to the time component in classification and clustering (Berndt and Clifford, 1994). In DTW, two given time series are stretched or compressed locally (i.e. “warped”) to align in a nonlinear manner so that the global distance between the two series (i.e. the summation of the local distances of individual aligned elements) becomes minimal (Giorgino, 2009; Ratanamahatana and Keogh, 2004). Researchers found that DTW distance measure outperformed the conventional

Euclidean distance metric where the distortion in time axis highly affected the grouping results. Since then, the similarity measurement using DTW has been more applied to classification and clustering in many domains such as gene expressions, medicine, and engineering (Aach and Church, 2001; Bar-Joseph et al., 2002; Caiani et al., 1998; Debrégeas and Hébrail, 1998; Kadous, 1999; Ratanamahatana and Keogh, 2004). Figure 2.1 visualizes the different metrics for measuring distance similarities. Note that the two temporal sequences (solid lines) are not aligned in the time axis (horizontal axis) even though their variation shape is similar. As shown in Figure 2.1(a), Euclidean distance metric maps the ith point of one sequence to the ith point of another sequence locally. Meanwhile, Figure 2.1(b) illustrates that DTW distance metric allows a nonlinear alignment between the points to shift the sequences for intuitive matching.



(a) Euclidean distance metric



(b) DTW distance metric

Figure 2.1. Distance similarity measurement for time series (Chu et al., 2002)

To understand DTW algorithm, assume two time series, A of length n, and B of length m, where

$$\mathrm{A} = a_1, a_2, \cdots, a_i, \cdots, a_n \tag{2.1}$$

$$\mathrm{B} = b_1, b_2, \cdots, b_j, \cdots, b_m \tag{2.2}$$

The two sequences A and B can be aligned by creating an n-by-m matrix where each element (i, j) includes the distance $d(a_i, b_j)$ between the two points a_i and b_j. The distance $d(a_i, b_j)$ is commonly computed using the Euclidean distance, hence,

$$d\left(a_i, b_j\right) = (a_i - b_j)^2 \tag{2.3}$$

The element therefore means the alignment between the points a_i and b_j. A warping path, W, is a combinational sequence of these alignments, represented as a set of adjacent elements mapped on the matrix as shown in Figure 2.2. The k^{th} element of W is defined as $w_k = (i, j)_k$, hence,

$$W = w_1, w_2, \cdots, w_k, \cdots, w_K \quad \text{where } \max(m, n) \leq K < m + n - 1 \tag{2.4}$$

The optimal path aims to minimize the warping cost as stated below:

$$DTW(A, B) = \min(\sqrt{\textstyle\sum_{k=1}^{K} w_k}) \tag{2.5}$$

The path is also constrained by several constraints:

- Monotonicity: the points in W must be monotonically ordered in time, satisfying that $i_k - i_{k-1} \geq 0$ and $j_k - j_{k-1} \geq 0$.
- Continuity: the steps to the next element are limited to the adjacent points including the diagonal points, satisfying that $i_k - i_{k-1} \leq 1$ and $j_k - j_{k-1} \leq 1$.
- Boundary conditions: the warping space is restricted by the endpoint conditions such as w_1 = (1, 1) and w_K = (m, n). From this constraint, the warping path is required to start and finish in diagonally opposite corners of the matrix.

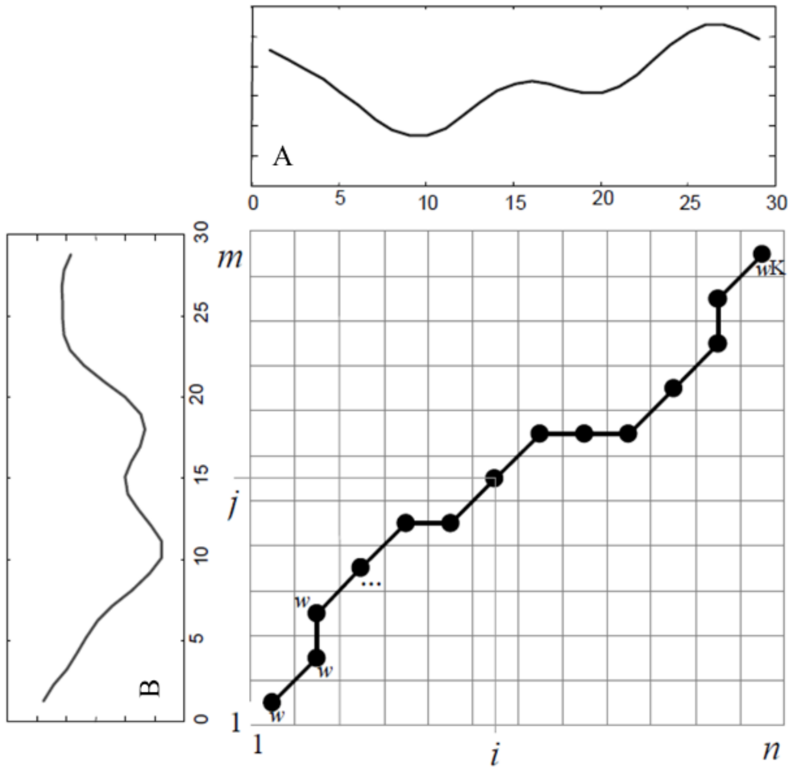


Figure 2.2. An example warping path between two time series A and B on a matrix (Chu et al., 2002)

With the cost function and the constraints, the optimal path can be computed using dynamic programing to evaluate the following recurrence relation, which introduces the cumulative distance, $\gamma(i, j)$, for each element,

$$\gamma(i,j) = d(a_i, b_j) + \min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\} \qquad (2.6)$$

where $d(a_i, b_j)$ denotes the distance found for the current points.

Equation (2.6) indicates that the cumulative distance, $\gamma(i, j)$, is the summation of the current distance ($d(a_i, b_j)$), and the minimum among the cumulative distances of the adjacent elements, ($\min\{\gamma(i\text{-}1, j), \gamma(i\text{-}1, j\text{-}1), \gamma(i, j\text{-}1)\}$).

2.2.2 Density-Based Spatial Clustering of Applications with Noise

In general, density-based clustering divides the data into separate spaces by how much the points are packed over a contiguous region (i.e. high density) and by how much they are sparse (i.e. low density). The subspaces of dense objects are grouped as clusters, while the data objects in low-density subspaces are considered noise or outliers (Kriegel et al., 2011). Density-Based Spatial Clustering of Applications with Noise (DBSCAN) proposed by Ester et al. (1996) is one of the most popular density-based clustering methods to expand a cluster as long as the neighborhood points locate within certain distance thresholds. The method is well-known for its advantages that (1) the number of clusters does not need to be predefined before clustering, which leads to (2) a superior application for finding arbitrarily shaped clusters and identifying outliers, and (3) it has good efficiency on large databases (Ester et al, 1996; Ma and Angryk, 2017).

The basic understanding of DBSCAN starts from specifying "density" for a local data point, p, in data set, D, by two parameters:

- *Eps* (ε): radius to declare the neighborhood of point, p. We set ε-neighborhood, $N_\varepsilon(p)$, as all points within ε from the point p. Hence, for a point q in D,

  $$N_\varepsilon(p) \coloneqq \{q \in D | d(p, q) \leq \varepsilon\} \tag{2.7}$$

  where $d(p, q)$ is a distance function for two points p and q, for instance in this study, Euclidean distance metric is used.

- *MinPts*: minimum number of points in an ε-neighborhood of p to form a cluster.

If an ε-neighborhood of p contains at least *MinPts*, the density of p is considered high and sufficient to make a cluster, otherwise the density is low. Based on the density, data points in D are classified into three categories:

- Core point: a point with high density for its ε-neighborhood. Hence, it locates inside a cluster.
- Border point: a point with low density, but in the ε-neighborhood of a core point. It positions on the edge of a cluster.
- Noise point: any point neither a core point nor a border point.

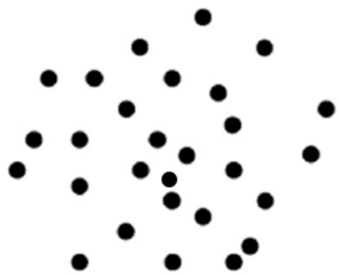
In DBSCAN algorithm, the cluster expands by connecting the points. The connectivity between the points is evaluated by the concept of "density-reachability".

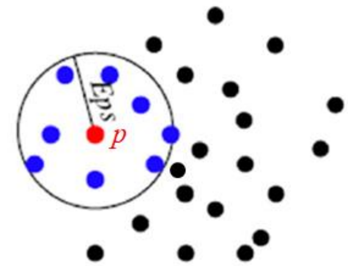
- Directly density-reachable: if a point p is a core point and a point q is in the ε-neighborhood of p, q is directly density-reachable from p. The relation between p and q is not symmetric.
- Density-reachable: if there is a chain of points $p_1, p_2, \cdots, p_n$ where $p_1 = p$, $p_n = q$ such that the p_{i+1} is directly density-reachable from p_i (hence, all p_i's except for q must be core points and form a chain of $p \to p_2 \to \cdots \to p_{n-1} \to q$), then q is density-reachable from p. It is also an asymmetric relation.
- Density-connected: if a pair of points p and q are both density-reachable from another point o, p is density-connected to q. The relation is symmetric.

To sum up, given a data set D and parameter *Eps* and *MinPts*, a cluster in DBSCAN algorithm satisfies the following conditions:

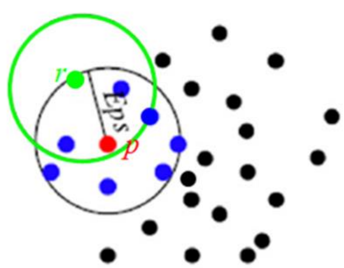
- All points within the cluster are density-connected to each other.
- If p is in a cluster and q is density-reachable from p, then q is in the cluster.
- Border points are considered to be included in the cluster.

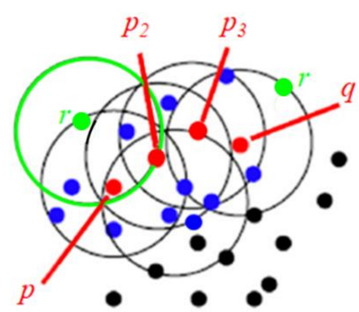
Figure 2.3 illustrates the three data point types and the connectivity concepts with sample data points when *MinPts* is 9. In the figures, the black solid circle groups represent the same cluster. Blue points are non-core points within the cluster.





(a) Data points (b) Core point (red point), *p*, within *Eps*, *ε*





(c) Border point (green point), *r* (d) Density-reachable points (red points)

Figure 2.3. Sample DBSCAN Clusters (Modified from Xue et al., 2018)

When applied to clustering, DBSCAN method first starts with selecting an arbitrary point *p* to create a cluster. If *p* is a core point with respect to predetermined *Eps* and *MinPts*, all density-reachable points from *p* are found to form a cluster. If *p* is not a core point, it is labeled as a noise or an outlier, and a new point is selected to measure the *Eps* and *MinPts* to make a cluster. This process is repeated to visit and label all the points on the data space whether they are clustered or considered noise (Kiware, 2010).

Chapter 3: Methodology

This chapter introduces the experimental test setups and procedures to collect the eye tracking data from a scenario-based offshore drilling monitoring simulation. A high quality offshore virtual simulator facility provided a realistic environment of the drilling exercise. A commercial eye tracking device was used to track the eye movements while the subject participants performed the task of observing the real-time data on the monitoring screen. Participant's verbal responses were also recorded. The analysis software coupled with the eye tracker exported the digitized ocular data. Quantitative statistical analysis was conducted to assess the collected data based on the Areas of Interest (AOIs), and time series clustering method revealed the participant's visual scanning pattern.

3.1 Participants

A group of 23 undergraduate and graduate students was recruited from the Department of Petroleum and Geological Engineering at the University of Oklahoma (OU) to represent the novice participants. Prior to the experiments, the novices learned how to read and interpret the drilling operation logs. Among the 23 novice participants, 2 participant datasets were omitted from the analysis due to the low quality of recording. Another 2 datasets were also excluded to avoid potential outliers since their number of eye fixation were precisely lower than those of other participants. A field expert with 30 years of hands-on experience also participated in the experiment to compare his physiological responses to those of novices. Therefore, the data from 19 novice participants and 1 expert participant involved in the analysis.

3.2 Apparatus

The state-of-the-art National Oilwell Varco (NOV) Drilling Virtual Simulator at OU's Drilling Simulation Center (OUDSC) was used to reflect the real-life offshore drilling operation. Figure 3.1(a) shows the overview of the OUDSC equipped with three NOV Drilling Virtual Simulator chairs. The simulator setup with the control chairs resembled the actual drilling rig floor (Dadmohammadi et al., 2017). The participant sat on the chair and the Tobii TX 300 Eye Tracker (sampling rate of 300 Hz) was placed approximately three feet away from the participant to trail the eye movement. A 23" monitor (1,920 x 1,080 pixels) was fitted above the eye tracker as shown in Figure 3.1(b) to display the real-time drilling operation logs. The Tobii TX300 Eye Tracker collected the gaze information on the monitor screen every 3.3 to 33 ms using 4 built-in cameras (Tobii Technology AB., 2016). The visual angle error of the eye tracker was less than 0.5° for binocular visions (Tobii Technology AB., 2010). The ocular data from the eye tracker were recorded and digitized by Tobii Studio 4 software. The collected data were analyzed for statistical tests and time series clustering using R software (version 3.4.0) (Venables et al., 2004).

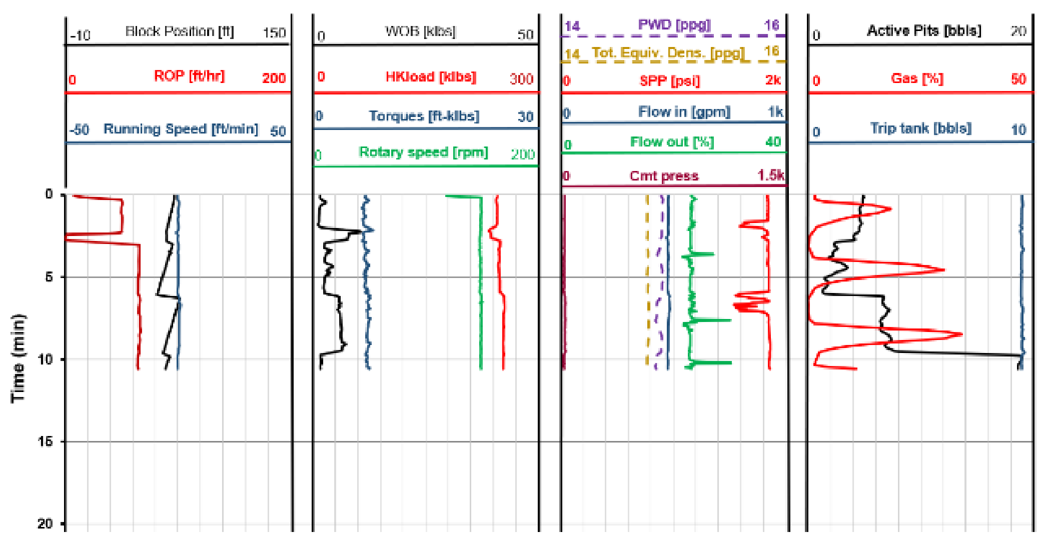
(a) NOV Drilling Virtual Simulator (b) Tobii TX300 Eye Tracker

Figure 3.1. Test environment and apparatus

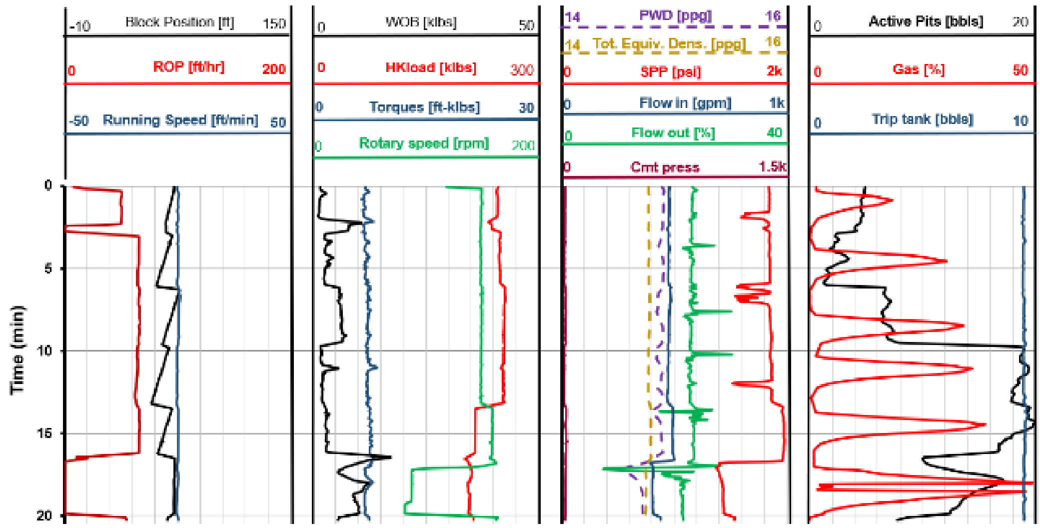
3.3 Scenario

The scenario was designed to collect the physiological responses of the participants while they were monitoring the real-time drilling logs, and to measure their situational awareness (SA). This real-time scenario was based on the preliminary events of an oil well control loss occurred at the North Sea rig (Norwegian Petroleum Directorate Guidance for Drilling and Well Activity, 2004). The actual incident involved a power failure to mud pumps and engaged gas influxes into the well beforehand and afterwards. The scenario focused on the influx indicators prior to the incident (i.e. the precursors). Influx indicators were detected as abnormal tendencies of the drilling logs such that the participants would recognize the well malfunction.

The twenty-minute actual log profiles were shortened into a 150-second duration to save the experimental time. In the scenario, the logs were displayed in real-time. Figure 3.2(a) shows the real-time scenario after 80 seconds and Figure 3.2(b) displays the scenario after the entire 150 seconds run. Both figures include 16 logs to monitor the regular trend and anomalies in drilling operations.



(a) Drilling log scenario after 80 seconds



(b) Drilling log scenario after 150 seconds

Figure 3.2. Real-time drilling log scenario

Among the logs some showed major indicators of well control loss whereas others were supportive indicators to help the interpretation of the major indicators. The major indicators in the scenario are “rate of penetration” (i.e. “ROP”), “flow out” (i.e. “Flow out %”), “gas percentage” (i.e. “Gas %”), and “active pit volume” (i.e. “Active Pits”).

“Rate of penetration (ROP) [ft/hr]” represents the speed of the drill bit to break the rock and deepen the wellbore. It is one of the most important predictive indicators to understand the downhole condition since its value changes mainly based on either the solidity of the subsurface formation or the properties of the fluid/gas within the subsurface. For example, if the formation becomes softer or fluid/gas gets lighter, ROP will increase. However, since a drilling operator cannot conclude whether the increase of ROP is due to gas influx, or tender formation, or other many possible reasons, when understanding the ROP values, the operator should additionally interrogate other predictive indicators such as flow out.

“Flow out [%]” means the return flow of the drilling mud from the well. If both the values of flow out and ROP increase, then we can suspect an influx into the well. Whether we have an influx is checked by shutting down the mud pumps. If the flow out does not decrease even if the mud pumps have been shut down, then it is highly likely that we have an influx.

“Gas [%]” that is represented in percentage is often called “trip gas” and shows the accumulation of gas that flows into the well when connecting or tripping (i.e. the vertical movement) the drill pipe string. Most of the gas in the well is cleared by circulating mud in the hole. Hence, if a consistent increasing trend of the connection gas is observed, it might indicate an influx from the formation.

“Active pits [bbls]” is an important parameter to monitor drilling operation. The pit is a large tank to store drilling mud. Several pits are usually placed to handle the fluctuating fluid volume. During the drilling operation, mud is constantly circulated through the drill pipe string to clean the wellbore and reduce friction. If pit gain (i.e. increase in pit volume) occurs during drilling, it might indicate an influx.

In this scenario, three major abnormal situations can be found in different time windows: in the beginning phase of 0 – 5 minutes window, ROP increased from 50 to 60 feet/hour, which led the increment of active pit volume (Active Pits) in the later time at around 5 minutes and 10 minutes. The ascending trend of the gas percentage (Gas) throughout the whole-time window engaged in the irregular picks of the flow-out (Flow out) from the well.

3.4 Task and Procedure

The task was to obtain the oculomotor information and vocal responses while the participants were monitoring the scenario-based drilling logs and observing any anomalies in the logs.

To ensure a constant level of knowledge within the participant group, each participant completed an online test before the experiment about the basic drilling operation and the relevant logs. After the online test, the participant was provided with a brief introduction to the experimental environment and the use of eye tracker. Participant's eye position from the display screen and oculus detection accuracy were calibrated prior to the monitoring task. To do a calibration, the participant attended to a mobile red dot on the screen for a few seconds and the eye tracker adjusted the detection accuracy according to the distance of the participant from the screen.

After the calibration, the participant was asked to monitor the trend of the logs and diagnose any anomalies while the 150-second scenario was being displayed in real-time. The participant was not informed whether the abnormalities were imbedded in the scenario or not so that the test was not interfered by any biased attention to certain logs. Whenever an abnormality was detected the participant was requested to verbally explain how the log trend looks like, what would be the interpretation of the trend and if any action or decision is necessary. Three keywords ("abnormal", "kick" and "blow-out") were given to the participants to use in their verbal response for the assessment of the situation. The word "kick" and "blow-out" are common terminologies used in drilling operations, implying an influx into the oil well and an explosion of the well, respectively. The eye movements and the verbal responses were concurrently recorded by the eye tracking device with a timestamp throughout the task.

3.5 Data Extraction and Preparation

The ocular data obtained during the log monitoring task were studied in two ways to understand the participant's cognitive performance: a quantitative statistical analysis to relate the visual attention to the level of SA, and a time series clustering to assess the visual scan pattern. In the quantitative statistical analysis, the oculomotor information was corroborated with the verbal data as a mixed approach to evaluate the participant's SA.

3.5.1 Data Extraction

For both approaches, Tobii Studio 4 software was used to digitize and export the oculomotor data collected by Tobii TX300 Eye Tracker. During a recording, raw eye movement data points from the eye tracker were identified with "X, Y" coordinates with respect to a timestamp (Tobii Technology AB., 2016). The software received the coordinate information and processed it further to provide a spatial fixation of the ocular attention on the monitor screen. The spatial fixation was regarded as an "eye fixation". In this experiment, the "eye fixation" was defined as a continuous gaze of the oculus focusing on a certain "X, Y" coordinate pair for more than or equal to 100 ms. Each eye fixation was visualized as a circle overlaid on the screen. The size of each circle indicated how long the participant attended to the fixation point (i.e. duration of an eye fixation), and number in the center of the circle provided the sequential order of the eye fixation throughout the recording.

The eye fixation data were then exported from Tobii Studio 4 software in the excel file format (.xlsx). The software also provided several predefined data settings related to the fixation. In the excel output, the data settings were represented as column domains and each row contained the relevant information of single eye fixation. Among the setting options, five settings were selected for the analyses. The details and the units of the settings are described below (Tobii Technology AB., 2016):

- "Fixation Index" indicates the sequential order in which a fixation event was recorded. The index starts from 1 and is in increasing order.
- "Gaze Event Duration [milliseconds]" represents the time duration of an eye movement event. Since the gaze event was regarded as an eye fixation, this setting implies the duration of an eye fixation.
- "Fixation Point X (MCSpx) [pixels]" provides the horizontal coordinate of the eye fixation on the recorded screen. In the Media Coordinate System pixels (MCSpx), the data point is mapped into a 2D coordinate system on the screen.
- "Fixation Point Y (MCSpx) [pixels]" provides the vertical coordinate of the eye fixation on the recorded screen.
- "AOI [Name of AOI] Hit" reports whether the AOI is active and whether the eye fixation is made inside the AOI. "AOI" is an abbreviated term of "Area of Interest". Since the software generates the spatial fixation on the entire screen area, it is necessary to define AOI to separate and export the eye fixation information based on the particular region of user's interest. The number, shape, size and name of the AOI is determined by the software users per their application design. The users can also decide whether the AOI is activated in

certain time windows according to their research purpose. By selecting the "AOI Hit" setting, the excel output includes one column per AOI. If the AOI cell is empty, it implies that the recorded media for the subject AOI is not activated. If the cell has a value of -1, it means that the subject AOI is inactive. A value of 0 indicates that the subject AOI is active, but the eye fixation is not made within the AOI. Lastly, a value of 1 informs that the subject AOI is active and the eye fixation point is positioned inside the AOI.

Tobii Studio 4 software also provides the audio data recorded during the experiment. Not only the video data of eye tracking but the audio data are contained in the media file and can be exported in AVI (Audio Video Interleaved) format if needed.

3.5.2 Data Preparation for Quantitative Statistical Analysis

To perform the quantitative statistical analysis, two eye tracking metrics were computed from the eye fixations on the particular AOIs within the log monitoring screen: "eye fixation count" and "eye fixation duration". "Eye fixation count" is the total number of eye fixations made on the designated AOI. "Eye fixation duration" means the total time of the individual eye fixations made on the AOI.

Figure 3.3 illustrates three predetermined AOIs for the real-time scenario. The AOIs are shown as square boxes with different colors from each other. These AOIs were designed to represent the major indicators of the well control loss during the task. AOI1 includes the ROP log indicator that increased from 50 ft/hr to 60 ft/hr at around 2 mins and 30 seconds time window, AOI2 covers the sudden picks in the Flow out log corresponding to the rising trend in Gas log, and AOI3 displays the Gas log indicator as well as the increment in Active Pits volume at around 6 mins and 10 mins time slots.

Since the eye tracker had the visual angle error (less than 0.5°), the AOIs were decided to be relatively larger than the actual log indicator area to accommodate the angular deviation.

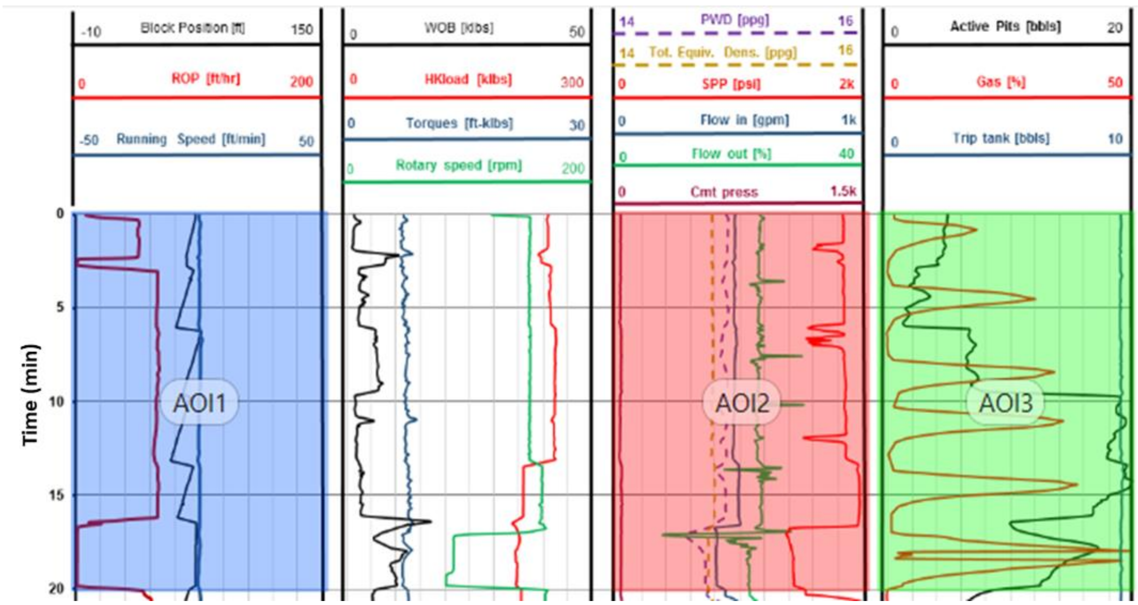


Figure 3.3. AOIs for the real-time drilling log scenario

As a mixed approach to understand the participant's SA, the eye tracking metrics of the participants were coupled with their verbal responses on each AOI. To quantify the verbal response, the audio records obtained from the Tobii Studio 4 software were first transcribed into text format, then scored from 0 to 10 by a "verbal response scoring metric" developed for the study. The metric was prepared based on two aspects: whether the participant correctly understood and interpreted the situation, and whether the verbal reaction to the situation was made fast enough. Table 3.1 shows how the score points were assigned to the AOI.

Table 3.1. Verbal response scoring metric

Score	Score description	Response Time
0	Attention but not related to the indicator log	Early/Late AOI
	No reaction: no comment on the indicator log	Early/Late AOI
1	Incorrect interpretation and incorrect understanding	Early/Late AOI
	Incorrect interpretation and no understanding explanation	Early/Late AOI
	Incorrect understanding	Early/Late AOI
	Attention to the situation	Late AOI
2	Correct understanding, but incorrect interpretation	Early/Late AOI
	Attention to the situation	Early AOI
3	Correct/Decent understanding, but no interpretation	Late AOI
4	Decent understanding of the situation, but no interpretation	Early AOI
5	Correct interpretation and incorrect understanding	Late AOI
	Correct understanding of the situation, but no interpretation	Early AOI
6	Decent interpretation	Late AOI
	Correct interpretation, but curiosity in understanding	Late AOI
	Correct interpretation after understanding in early AOI	Late AOI
7	Correct interpretation and incorrect understanding	Early AOI
8	Decent interpretation	Early AOI
	Correct interpretation and correct understanding	Late AOI
	Correct interpretation and no understanding explanation	Late AOI
9	Correct interpretation, but curiosity in understanding	Early AOI
10	Correct interpretation and correct understanding	Early AOI
	Correct interpretation from the indicators in previous AOIs	Early AOI
	Correct interpretation and no understanding explanation	Early AOI

In the score description, the term “understanding” was considered when the participant verbally explained the status of the log indicators. For example, a participant would indicate a gas issue with specific phrases such as “gas percentage is increasing”, which would be counted as an “understanding” of the issue on the target indicator. “Interpretation” was regarded when the participant verbally reacted with the given keywords (“abnormal”, “kick” and “blow-out”) after assessing the situation based on their “understanding”. For instance, if the trend of gas percentage log was correctly identified as abnormally ascending and the participant concluded the issue as an “abnormal” situation or a “kick”, then the log indicator (gas percentage increment) was correctly interpreted.

The verbal responses were scored based not only on the description of the issue but also on the reaction time with respect to the indicators in the target AOI. In the response time column, “Early AOI” means that the participant reacted to anomalies while the log indicators were presented within the upper half area of the AOI. Hence, the response made on “Early AOI” can be understood as a prompt reaction to the issues. Meanwhile, if the participant verbally announced the anomalies in the logs while they were appearing within the lower half area of the AOI, it was considered as a “Late AOI” reaction. Therefore, if a participant had “Late AOI” with the correct interpretation of the indicator based on the correct understanding of the situation, the verbal response score would be an 8 instead of a perfect score of 10. To sum up, high scores can be related to the correct and/or prompt awareness of the situation, while lower scores relatively imply lack of awareness. If multiple interpretations were made on the same AOI, the maximum score was considered to avoid a replicated evaluation.

3.5.3 Data Preparation for Time Series Clustering

The time series clustering approach utilized the “X, Y” coordinates of the eye fixations on the entire log area instead of segmenting it into several AOIs. The coordinates of the eye fixation points were studied with respect to the fixation index, which provided the sequential order of each eye fixation during the recording. By considering the overall area without an assumption of the predetermined AOIs, the eye movement data required less adjustment prior to the analyses, and the holistic visual scanning trajectory between the log indicators occurred at the same time window was able to be sought (Haass, 2016).

The scanning trajectory consists of three elements: “X, Y” coordinates of eye fixations, and the timestamps that the fixations were made. Tobii Studio 4 software produced Fixation Point X (MCSpx) and Fixation Point Y (MCSpx) to show the “X, Y” coordinates of the individual eye fixation point. These Fixation Points were collected with corresponding timestamps such as Recording Time Stamp and Fixation Index. Since every participant paid different amount of time on each eye fixation, the number of eye fixations were varied among the participants despite the same task time. Hence, the apple-to-apple comparison of the fixation coordinates at the same recording time may not be able to provide the global understanding of the scanning trajectories. Instead, it is more feasible to trace the fixations with their sequential order. Therefore, the Fixation Index was used to alter the actual timestamp and to represent the temporal information in clustering the scan pattern of all participants.

A thorough review of all eye fixation recordings suggested that the overall task time (150 seconds) be quantified approximately 400 fixation indices for most of the participants. Among them, the first 100 fixation indices covered the time window where most of the major log indicators appeared. The “X, Y” coordinates for the fixation index from 1 to 100 were employed for the clustering analyses.

When the log scenario was displayed on the monitor screen, the (0, 0) coordinates for the horizontal and vertical axes of the scenario were deviated for several participants due to an experimental error. Hence, the raw “X, Y” coordinates were examined and reassigned to ensure that the participants were monitoring the same position when attended on the same indicators. The newly-assigned "X, Y" coordinates were then considered as spatial information for the scan pattern clustering.

3.6 Quantitative Statistical Analysis

To evaluate the participant's SA, one-way ANOVA (Analysis of Variance) was applied as a quantitative statistical method. ANOVA is a form of statistical hypothesis testing to seek whether or not the means of three or more sample groups are equal. It is commonly presented in a linear model to relate the effect of independent variables to the values of dependent variables. In case that "only one" independent variable dataset involves in producing the numerical responses (i.e. dependent variables), the ANOVA is considered as "one-way" (Howell, 2012). In ANOVA, the assumption of *null hypothesis* is that the sample means do not show a significant difference between them. In other words, all groups are randomly sampled from the same population, therefore the means of the sample groups may differ only by a random chance. The null hypothesis is held true when the probability (*p*-value) of the model is same as or greater than the actual observed results (Wasserstein and Lazar, 2016). If the *p*-value is less than the threshold (i.e. significance level), the null hypothesis is rejected and implies that the means of the sample groups are significantly different.

The result of ANOVA is generally considered reliable when three assumptions about the probability distribution of the responses are satisfied. The first assumption is that the responses are independent. To achieve this, the observations from one participant must not directly rely on the observations from other participants. The second assumption is that the residuals of the responses are normally distributed. That is, the dependent variable Y for a given independent variable X (i.e. Y|X or the distribution of Y within each X) should have a normal distribution. In this study, the residual normality was validated by Shapiro-Wilk test. If the residual normality is violated, a nonparametric test is required to alter ANOVA. The nonparametric test is often called "ANOVA on ranks" since it transforms the response data value to its rank and uses the rank for analysis of variance. Kruskal-Wallis test is a well-known nonparametric test for between-subject design experiments, while within-subject design experiments need to perform Friedman test for analysis. The third assumption is the homogeneity of variance. This means that the responses should have equal variance across the sample groups. To verify the assumption, Levene's test was applied because it is less sensitive to the normality. To enforce the robustness of the statistical power, non-normal responses were tested using the median while normally distributed responses were tested using the mean (Brown and Forsythe, 1974).

When ANOVA or a nonparametric test reveals that the means of the sample groups have significant difference among them, post hoc analyses are required to scrutinize the impact of the independent variable. Tukey test after ANOVA or Mann-Whitney-Wilcoxon test after the nonparametric test is commonly applied for multiple comparisons within the levels of the independent variable.

The quantitative statistical analyses for this study were performed in two ways. First, the impact of log indicators on the oculomotor responses was examined. Since the predefined AOI1, AOI2 and AOI3 included the major log indicators, the eye fixation count and the eye fixation duration calculated from these AOIs became the dependent variable values corresponding to the independent variable – log indicator. Three levels of the independent variable were determined in accordance with the number of AOIs. Secondly, the impact of log indicators on the verbal responses was investigated. Again, the verbal response scores computed for each AOI became the dependent variable values. In addition, the individual participants provided eye tracking data and verbal response data, therefore the experiment followed within-subject design.

Among the total 20 participants, 19 novice datasets were used for the statistical test. The expert data were separately studied to provide the comparison between the novice response and the expert response. To perform the statistical tests for the oculomotor responses and the verbal responses, three assumptions were verified in advance. Since the 19 participants were independently performed the task, the eye tracking metrics and the verbal scores for a participant were not directly related to others. Moreover, the log indicators simultaneously appeared on the monitor screen and were attended as per the participant's decision. Hence, the order of log indicators was considered random and the physiological responses were deemed to be independent.

The normality and homogeneity of response residuals were then investigated using Shapiro-Wilk test and Levene's test, respectively. It was found that the eye fixation counts were normally distributed and had the equal variance. The eye fixation durations also retained the homogeneity of variance, however, the assumption for normal distribution was not fulfilled for AOI2. For the verbal response scores, both normality and equal variance assumptions were not observed.

Upon the results stated above, one-way ANOVA was applied to the eye fixation count data. Eye fixation duration was also analyzed by one-way ANOVA although the data from AOI2 were not normally distributed. A research found from the extensive investigation that parametric tests performed well with non-normal continuous data with some sample size guidelines (Minitab 17 Support, 2015 and The Minitab Blog, 2015). Especially for the data with 3 levels of a factor and more than 15 observations, one-way ANOVA can be applied.

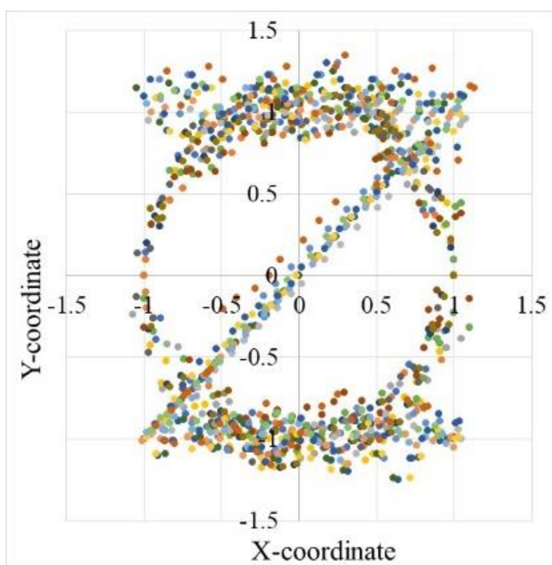
Friedman test was implemented for verbal score data to examine the impact of log indicators. A significance level of $\alpha = 0.05$ was used for every test. To explore further to the outputs of one-way ANOVA and Friedman tests, pairwise comparison tests were performed as a post hoc analysis using a Tukey test for the eye fixation counts and Mann-Whitney-Wilcoxon test for the eye fixation durations and the verbal scores. These multiple comparison tests were adopted with Bonferroni correction.

After the statistical tests were completed for oculomotor information and verbal reaction, the results were analyzed in a combinatorial way to evaluate the SA of the participants during the task.

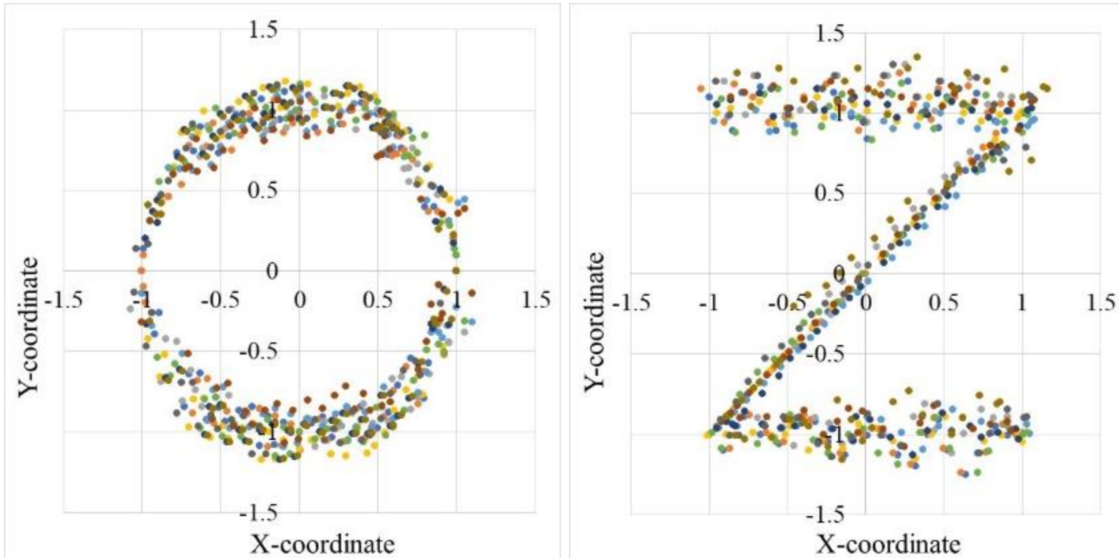
3.7 Adapted Time Series Clustering Model

In addition to the statistical approaches to quantify the level of SA, an analysis on the scan pattern of the subject participants could extend the understanding of the cognitive behavior responding to the stimuli from the surroundings. In this study, the scan patterns of the participants were identified with help of an unsupervised clustering technique; density-based spatial clustering of applications with noise (DBSCAN). The spatial information of the eye fixation for the clustering was obtained from the "X, Y" coordinates of the fixation points. Since the eye fixation occurred in a serial manner with respect to time, the temporal information (fixation index) was also included for the analysis. The temporal information was preprocessed by dynamic time warping (DTW) method prior to the clustering.

Based on the DTW algorithm and the DBSCAN protocol introduced in Section 2.2.1 and 2.2.2, a time series clustering model was developed to apply the real-time data from the experimental test. A model validated by Alcock and Manolopoulos (1999) was referred as a basis. The model was created using a synthesized dataset. Figure 3.4 shows the "X, Y" coordinates of the synthesized data. One dataset consists of 60 fixation indices and the corresponding "X, Y" coordinates. Total 20 sets of mock data were prepared and displayed on the same plane in the Figure 3.4(a). The dataset was generated to manipulate the scanning trajectory on the monitoring screen in two patterns as shown in Figure 3.4(b) and (c): a circular trajectory of 10 data sets and a Z-shaped trajectory of another 10 data sets. As seen in Figure 3.4(a), many of the fixation points were overlaid on the same coordinates, thus it would be hard to distinguish the two scan patterns without considering the timestamp.



(a) Overall data points

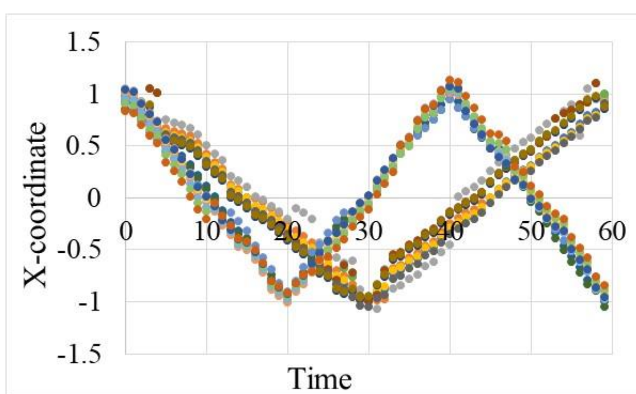


(b) Circular pattern points

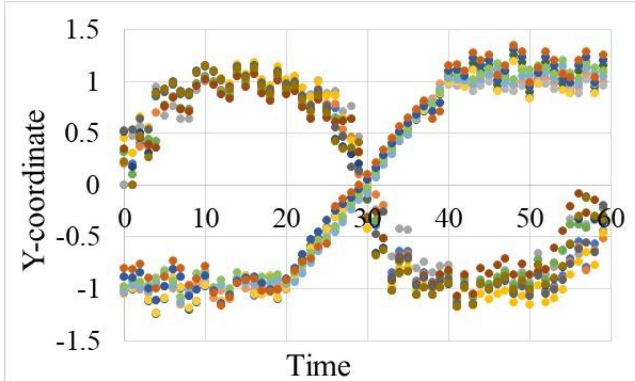
(c) Z-shaped pattern points

Figure 3.4. Synthesized eye fixation data sets on the X-Y coordinate plane

In the proposed model, 2-dimensional DTW method discussed in Section 2.2.1 was first applied using "dtwDist" function in "dtw" package in R software (Giorgino and Giorgino, 2012) to partition the time series data and compute the optimal distances. To simplify the process, the data points on the 3-dimensional X, Y, and Time-axis were segregated into X-coordinate vs Time plane and Y-coordinate vs Time plane. Two separate DTW processes were applied respectively to each data plane. Figure 3.5 shows the profiles of the data points.



(a) X-coordinate vs Time



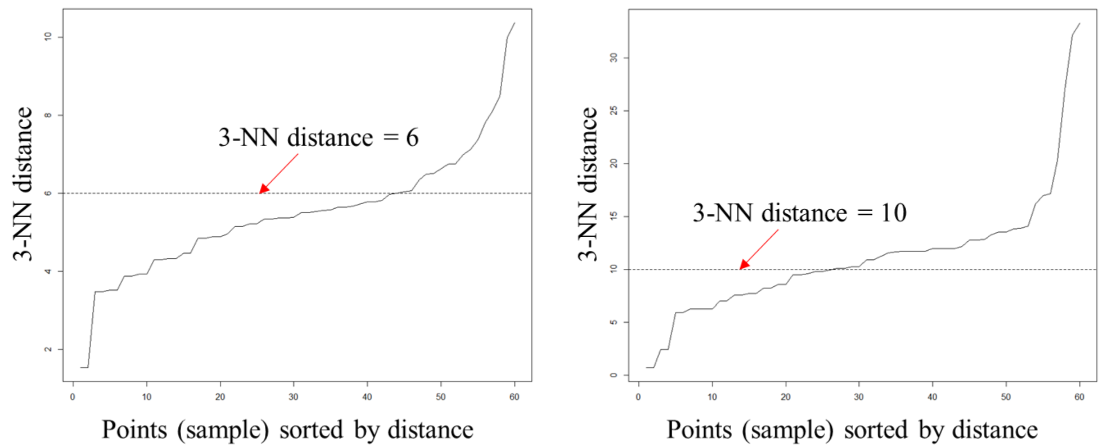
(b) Y-coordinate vs Time

Figure 3.5. Synthesized eye fixation data sets on the X-Time and Y-Time planes

After DTW found optimal alignments between the time series and constructed a 20-by-20 DTW distance matrix (the matrix dimension corresponded to the number of data sets) for each plane, DBSCAN using "dbscan" function in "dbscan" package in R software (Hahsler and Piekenbrock, 2017) was employed to identify the clustering structure.

Prior to the clustering, the DTW distance matrix went through a Kruskal nonmetric multidimensional scaling using "isoMDS" function in "MASS" package in R software (Ripley et al., 2013) to display the 20 objects of the matrix into 20 points on the virtual N-dimensional coordinate space such that the DTW distances between the objects are preserved as well as possible. In this study, the dimension, N, is selected 2.

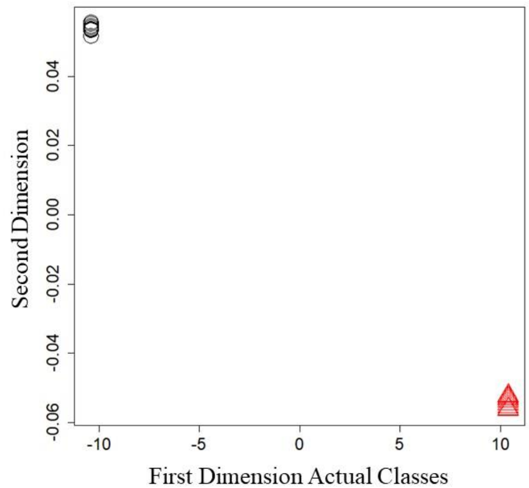
Considering the small size of the objects to cluster, *MinPts* of 3 was selected for "dbscan" function. The optimal *Eps* was determined by computing and visualizing the k-nearest neighbor (kNN) distances using "kNNdistplot" function in "dbscan" package (Hahsler and Piekenbrock, 2017). In kNN method, the average of the distances of every point to its k nearest neighbors was calculated. The value of k should be same as *MinPts*, therefore, k became 3 in this model. Then the average k-distances were plotted in an increasing order to find the "knee" which implied the threshold where a drastic change occurred in the k-distance profile. Figure 3.6 shows the optimal *Eps* found for this model.

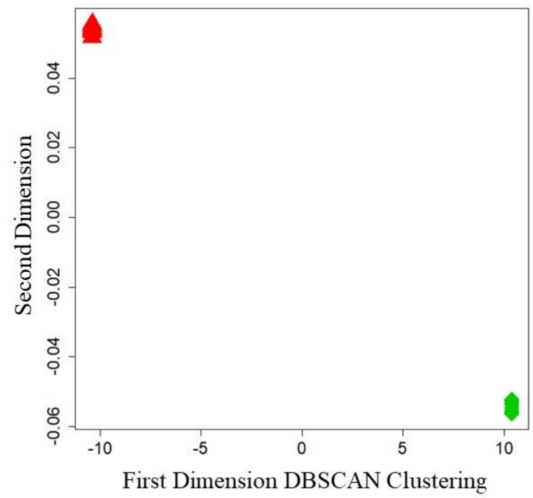


(a) X-coordinate vs Time (*Eps* = 6) (b) Y-coordinate vs Time (*Eps* = 10)

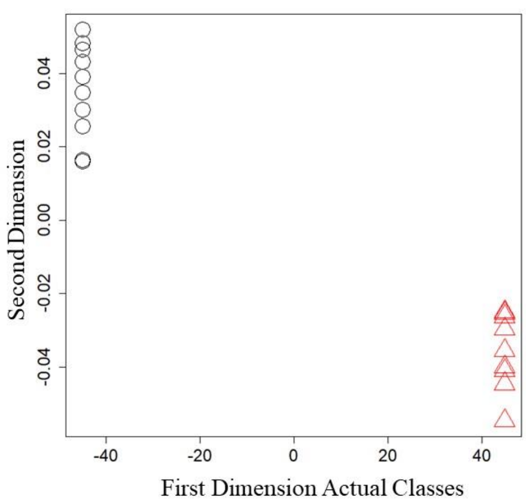
Figure 3.6. Optimal *Eps* for X-Time and Y-Time plane computed by kNN distance

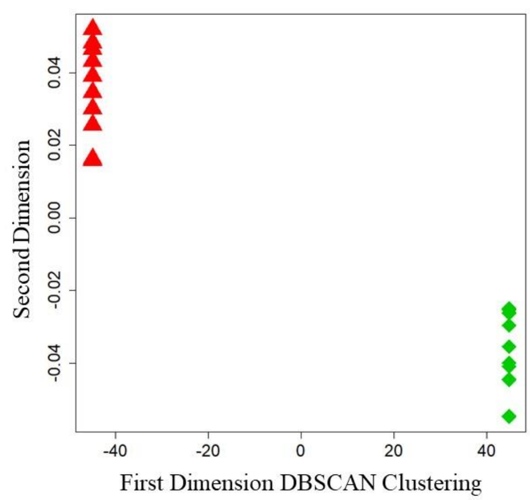
Using the parameters, DBSCAN grouped the 20 objects of time series into two clusters. Figure 3.7 illustrates the results of the clustering. Figure 3.7(a) and (c) visualize the original scanning pattern for X-coordinate vs Time plane data, and Y-coordinate vs Time plane data, respectively. Figure 3.7(b) and (d) show the clustering result for the corresponding plane data. The cluster outputs from both X-coordinate vs Time plane data and Y-coordinate vs Time plane data show a 100% match to the original scanning pattern.





(a) Original groups for X-coordinate vs Time (b) Clusters for X-coordinate vs Time





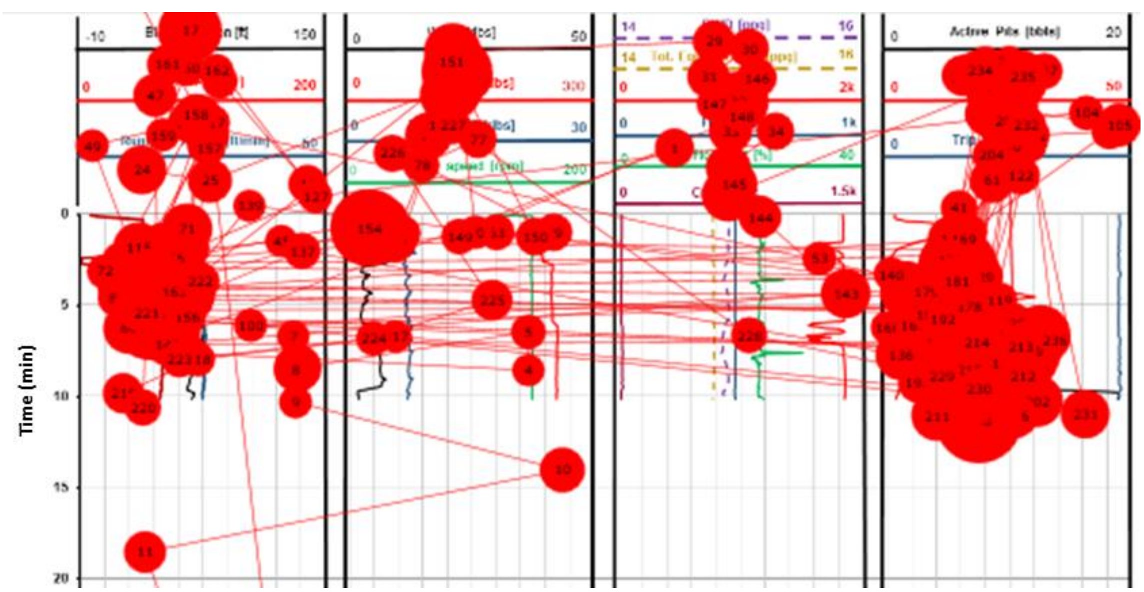
(c) Original groups for Y-coordinate vs Time (d) Clusters for Y-coordinate vs Time

Figure 3.7. DBSCAN results for X-Time and Y-Time plane data

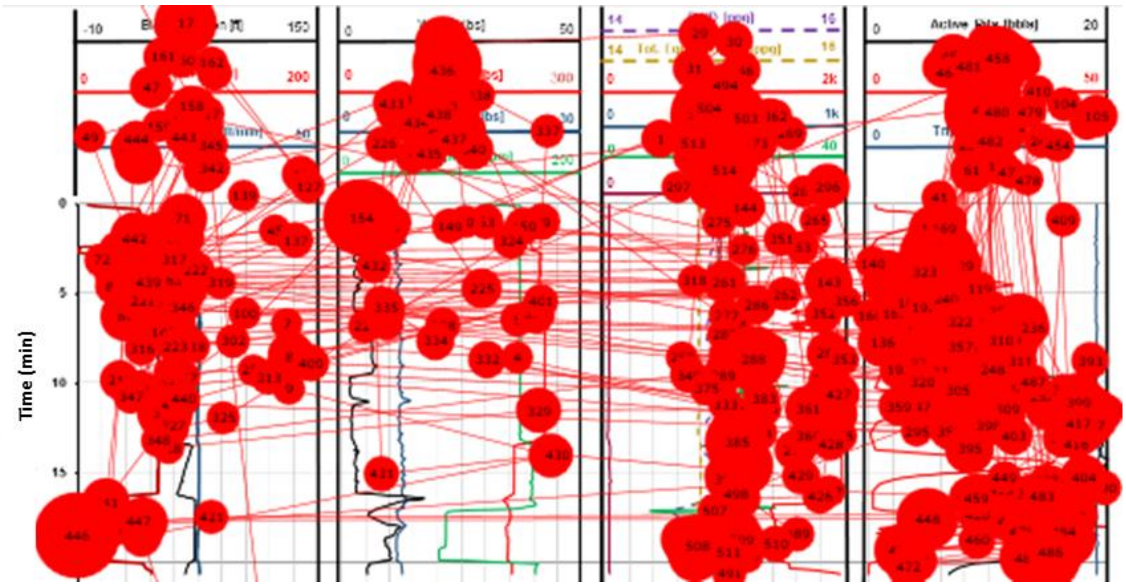
Chapter 4: Results and Discussion

Using the methods described in Chapter 3, the outcomes from the experimental task were investigated. The quantitative statistical test results are introduced in Section 4.1 and the time series scan pattern clustering results are described in Section 4.2.

Figure 4.1 visualizes the sample snapshots of the cumulative eye movements during the task. The eye fixation points are described as red-filled circles and connected by the red lines. These red lines indicate the “saccades” which show the intermediate eye movements between the fixation points. As mentioned before, the number in the center of the circle represents the sequential order of the fixation, and the size of the circle is proportional to the fixation duration. Figure 4.1(a) depicts the eye fixation information after 80 seconds of the experiment and Figure 4.1(b) displays the same information after the full 150-second period. From the figures, we can glance the scanpath of the participant on the monitor screen.



(a) Eye fixations after 80 seconds



(b) Eye fixations after 150 seconds

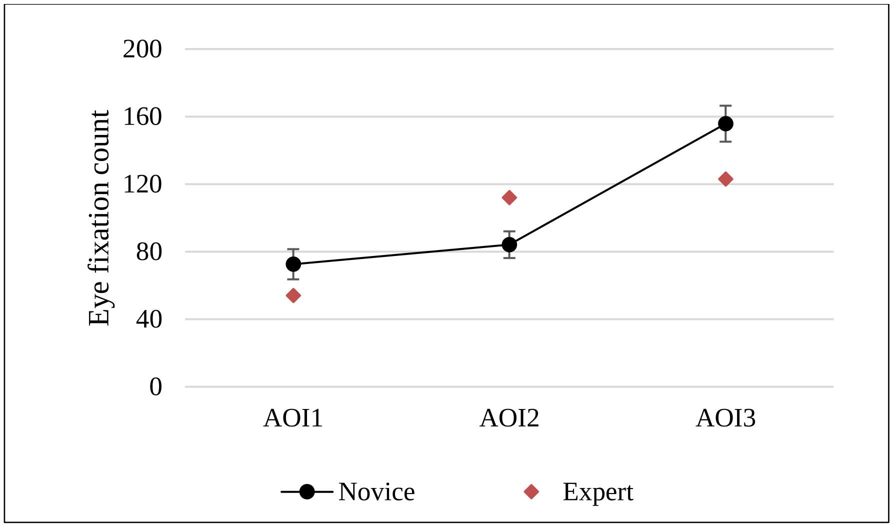
Figure 4.1. Examples of scanpaths during the monitoring task

4.1 Statistical Analysis of interrogated Areas of Interests (AOIs)

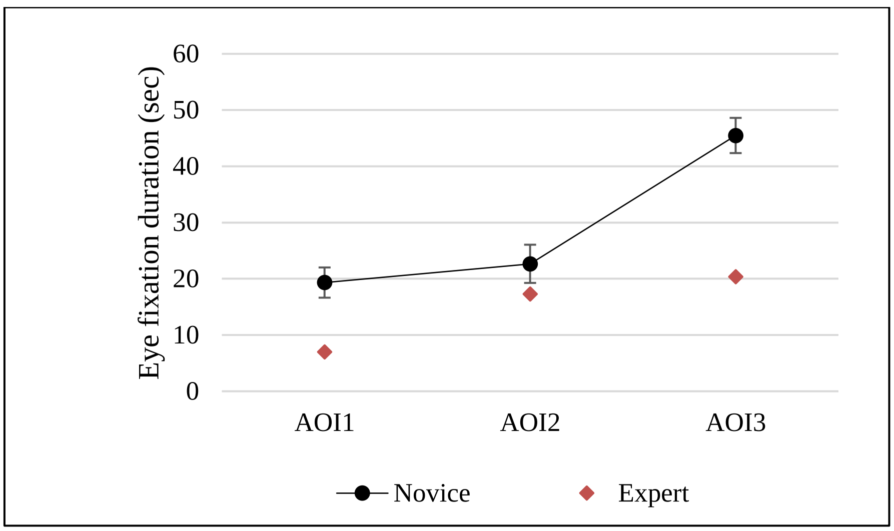
The physiological assessment using the eye movement data and the analysis using verbal responses were conducted separately, then combined to provide a comprehensive understanding on the SA. For statistical evaluation, a significance level of $\alpha = 0.05$ was used.

4.1.1 Oculomotor Response Results

Using the oculomotor data, the average eye fixation count and duration on the individual AOIs were computed and illustrated in Figure 4.2. Note that the eye fixation information is plotted with its standard errors.



(a) Eye fixation count result



(b) Eye fixation duration result

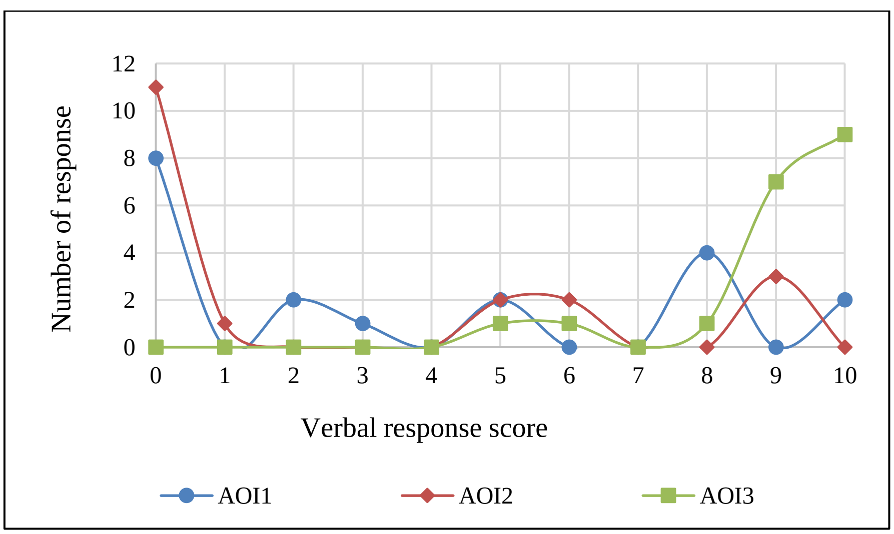
Figure 4.2. Eye tracking metrics from the AOIs

For both metrics, the novices showed an increasing trend from AOI1 to AOI3. A sharp increase in AOI3 (i.e. Gas % and Active pit volume log indicators) implies that the participants paid more attention to AOI3 indicators than those in other AOIs (i.e. ROP increment in AOI1 and Flow out abnormality in AOI2). One-way ANOVA based on the within-subjects design revealed significant differences on both eye fixation count

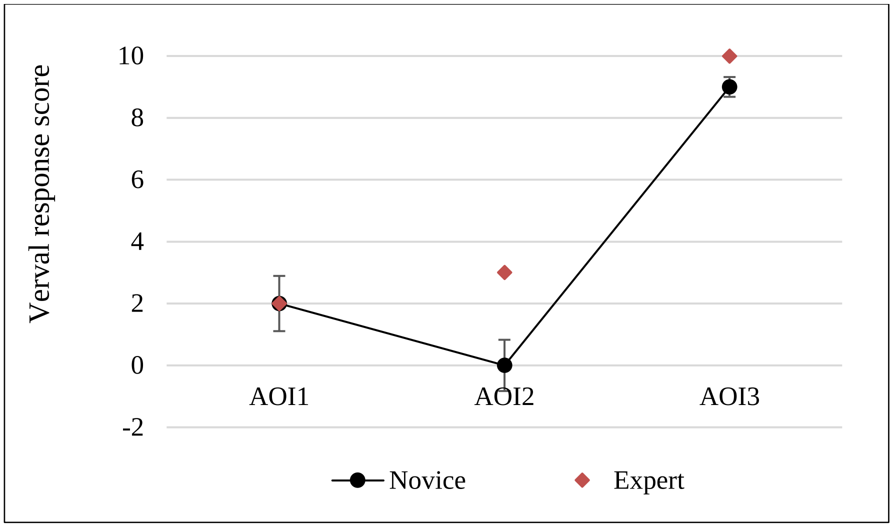
(F = 23.8, $p < 0.001$) and eye fixation duration (F = 21.3, $p < 0.001$). Post hoc analyses were done using the Tukey tests with $\alpha = 0.017$ (or 0.05/3). Significant differences in eye fixation count were found for AOI1 versus AOI3 (z = 6.37, $p < 0.001$) and AOI2 versus AOI3 (z = 5.48, $p < 0.001$). Similar statistical outcome was obtained for eye fixation duration, as significant differences were found for AOI1 versus AOI3 (z = 5.99, $p < 0.001$) and AOI2 versus AOI3 (z = 5.23, $p < 0.001$). The expert showed the similar ocular tendency, however, the difference between AOI1 and AOI2 is comparatively larger than that of novices. Meanwhile, the difference between AOI2 and AOI3 became smaller, indicating that the expert attended to AOI2 and AOI3 to almost the same extent.

4.1.2 Verbal Response Results

The verbal responses were quantified for each participant by the scoring metrics in Table 3.1. Figure 4.3 provides the distribution of the novices' responses and the response score result for each AOI. As the residual distribution was not normal, median response scores are presented in Figure 4.3(b).



(a) Novice's verbal response distribution



(b) Verbal response score result

Figure 4.3. Verbal response distribution and scores for the AOIs

As illustrated in Figure 4.3(a), majority of novices won low scores (0 – 2) in AOI1 and AOI2 while obtaining high scores (8 – 10) in AOI3. This result is well reflected in Figure 4.3(b) with the highest median response score in AOI3.

The expert also showed the highest score in AOI3 suggesting that all the participants recognized the abnormal log indicators in AOI3 (i.e. Gas % and Active pits volume) better than the indicators in other AOIs. A Friedman test for the novice's verbal response indicated that the three AOIs have significant differences in their median (χ^2 = 24, p < 0.001). Post hoc analysis using Mann-Whitney-Wilcoxon test found that the significant difference occurred for AOI1 versus AOI3 (p < 0.01) and AOI2 versus AOI3 (p < 0.01). Interestingly, the expert obtained slightly higher score in AOI2 than AOI1, while novice's median score was higher in AOI1 than AOI2.

4.1.3 Analyses of Situation Awareness from the Mixed Approach

The integrated method using eye fixation information and verbal response scores was investigated for analyses of situational awareness of the human operators in a complex, realistic oil well incident monitoring simulation task. Separate analyses of both metrics suggested two important convergent evidences. First, specific disparities in SA content or timing can be highly related to the differences in human performance, which reinforces the previous researches evaluating the level of operators' SA using only behavioral and verbal response analyses (Endsley, 1995; Roberts et al., 2015; Sneddon et al., 2006; Sneddon et al., 2013). Secondly, the eye tracking measurement functioned as a reliable supplementary analysis to corroborate and possibly improve the understanding of traditional measurements (Guan et al., 2006; Van et al., 2005).

The multi-metric approach provided complementary insights to assess the human decision-making process in the time critical task. For instance, the log indicators in AOI3 were regarded as the most prominent indicators, as both participant groups – the novices and the expert – tended to identify the risk for well loss incident from it. If we considered the verbal response scores alone, the indicators in AOI1 and AOI2 would be concluded not very supportive in detecting the potential incident. However, with the help of the eye movement result, the study found an increasing attentional trend of AOI1 < AOI2 < AOI3. By combining the outputs from verbal response information and eye fixation information, we were able to notify the difference in the cognitive performance of novices and expert. For both the novices and the expert, the eye fixation was made more on AOI2 than on AOI1, implying that AOI2 was more visually observed than AOI1, even though the novices nevertheless less interpreted the situation.

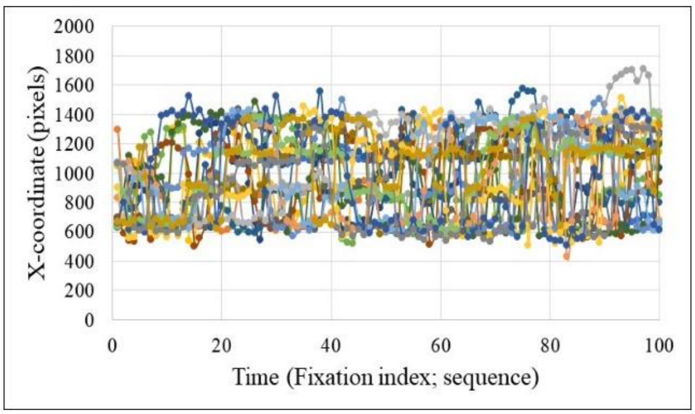
Meanwhile, the expert tried to corroborate the indicators in AOI1 and AOI2 with the indicators in AOI3 to some extent so that the contribution of AOI1 and AOI2 was more appreciated.

The fact that the novices less attended to the log indicators in AOI1 and AOI2 than the indicators in AOI3 might inform of a potential decision vulnerability caused by less visible indicators. It escalates the risk that the human operators may turn their attention from such indicators, thus potentially lose awareness of the critical signals of the situation. In other words, the indicators in AOI1 and AOI2 appeared to be frequently overshadowed by the indicators in AOI3 during the simultaneous monitoring phase of the task, which makes the novices miss the risky elements. The expert, however, put similar amounts of efforts between AOI2 and AOI3, which implies that the expert was able to correlate the abnormal increasing trend of Gas % and Active pits in AOI3 with the Flow out anomalies in AOI2. AOI1 is relatively less attended since the ROP abnormality occurred in the short time in the beginning and kept the constant level throughout the AOI afterwards. The distinction between the novices' and the expert's ocular behavior might be understood by some previous researches claiming that lower-skilled people more opt to concentrate on the "first-mind direction" when they involve in non-routine situations, whereas people with more expertise are likely to integrate more information from multiple sources before making a final decision (Cokely et al., 2016; Van Gog et al., 2005). The study results provided an agreement to this assertion by showing that the novices tended to focus more on the explicit trends of the indicators in AOI3 rather than instead of validating the outputs with other AOIs.

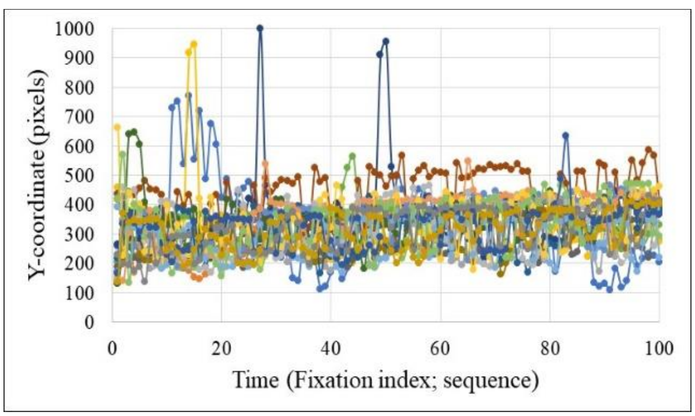
4.2 Scan Pattern Clustering

By applying the time series clustering model introduced in Section 3.7, the real-time monitoring data were analyzed to explore the participant's scan pattern in abnormal well control situation. As stated in Section 3.5.3, the "X, Y" coordinates for the fixation indices from 1 to 100 were subject to the clustering analysis.

The actual "X, Y" coordinate data were very complex to determine the clusters as shown in Figure 4.4. The profile of each eye fixation point changed with respect to time and locally overlapped each other. Hence, the original "X, Y" coordinates were processed with DTW computation to differentiate the time series points.



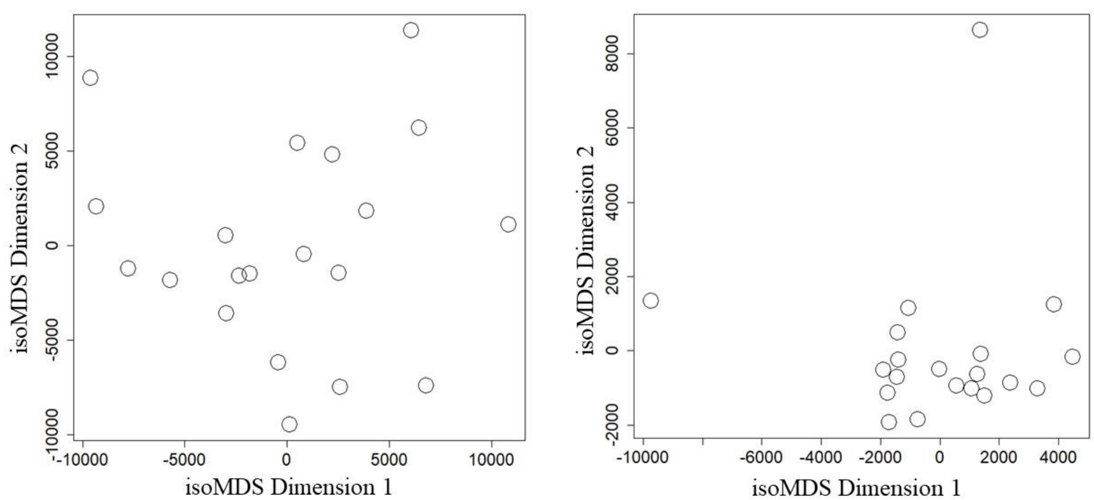
(a) X-coordinate vs Fixation index



(b) Y-coordinate vs Fixation index

Figure 4.4. Actual time series monitoring data

After performing DTW, the Kruskal nonmetric multidimensional scaling displayed the virtual coordinates of the 20 objects from the optimal DTW distance matrix on the 2-dimensional space as depicted in Figure 4.5. From these result, *Eps* of 5000 and *MinPts* of 3 were selected for X-coordinate vs Fixation Index plane data, and *Eps* of 2000 and *MinPts* of 3 were selected for Y-coordinate vs Fixation Index plane data to apply DBSCAN method.



(a) DTW distance matrix objects for X-coordinate time series

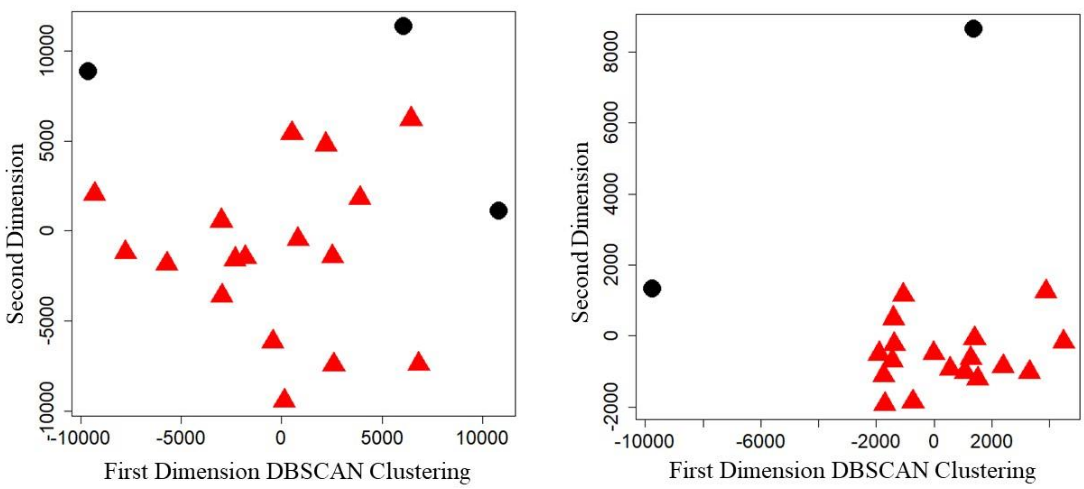
(b) DTW distance matrix objects for Y-coordinate time series

Figure 4.5. Representation of the 20 DTW distance matrix objects on 2-D space

The DBSCAN clustering results of the 20 participants (19 novices and 1 expert) for X-coordinate vs Fixation Index plane data and for Y-coordinate vs Fixation Index plane data are tabulated in Table 4.1 and visualized in Figure 4.6. Since *MinPts* was set to 3, the Cluster ‘0’ in Y-coordinate vs Fixation Index plane should be regarded as noise or outliers. These noise points can also be seen in Figure 4.6(b) illustrated as black circles.

Table 4.1. DBSCAN scan pattern clustering result

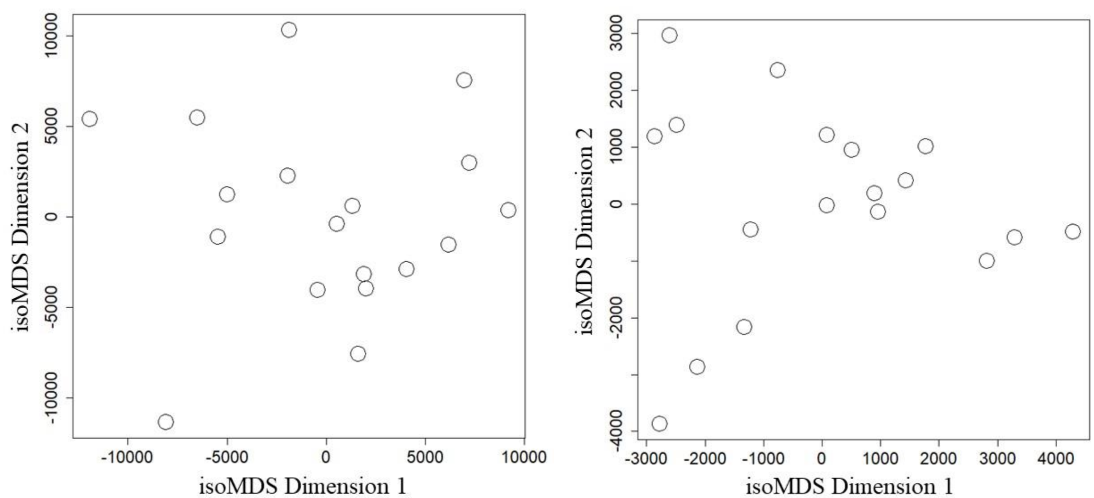
Participant	Cluster for X vs Fix. Index plane	Cluster for Y vs Fix. Index plane
Expert	1	0
Novice 1	1	1
Novice 2	0	1
Novice 3	1	1
Novice 4	1	1
Novice 5	0	1
Novice 6	1	1
Novice 7	1	1
Novice 8	1	1
Novice 10	1	1
Novice 11	1	1
Novice 13	1	1
Novice 14	1	1
Novice 16	0	1
Novice 17	1	1
Novice 19	1	0
Novice 20	1	1
Novice 21	1	1
Novice 22	1	1
Novice 23	1	1



(a) X-coordinate vs Fixation index (b) Y-coordinate vs Fixation index

Figure 4.6. DBSCAN scan pattern clustering result

To robust the analysis, the time series clustering model was re-applied after removing the two outliers (i.e. Expert and Novice 19) from eye fixation data sets. The Kruskal nonmetric multidimensional scaling suggested distribution of DTW distance matrix objects as presented in Figure 4.7. Based on the scaling result, *Eps* of 4000 and *MinPts* of 3 were selected for X-coordinate vs Fixation Index plane data, and *Eps* of 1000 and *MinPts* of 3 were selected for Y-coordinate vs Fixation Index plane data.



(a) DTW distance matrix objects for X-coordinate time series

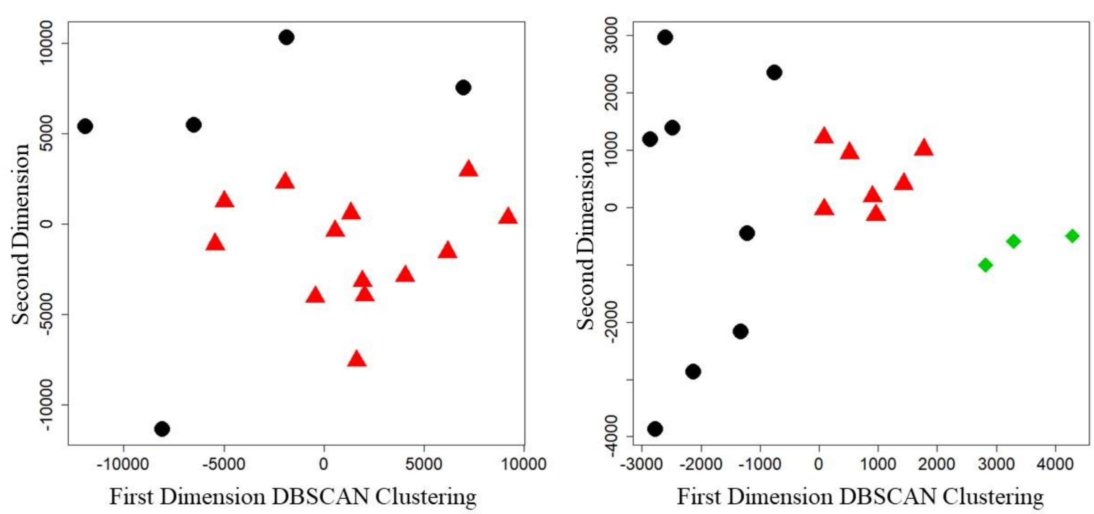
(b) DTW distance matrix objects for Y-coordinate time series

Figure 4.7. Representation of the 18 DTW distance matrix objects on 2-D space

The new result of DBSCAN clustering with 18 participant data are summarized in Table 4.2 and Figure 4.8. Since the clusters were determined in two ways – X-coordinate vs Fixation Index plane, and Y-coordinate vs Fixation Index plane, clusters in each plane were paired up to provide a combined final cluster. For example, if a participant was clustered as '1' in X-coordinate plane and '0' in Y-coordinate plane, the participant's final cluster was determined as 'A'. Table 4.2 includes the final clusters of all 18 participants' scan pattern.

Table 4.2. DBSCAN scan pattern clustering result after removing outliers

Participant	Cluster for X vs Fix. Index plane	Cluster for Y vs Fix. Index plane	Combined cluster
Novice 1	1	0	A
Novice 2	0	0	B
Novice 3	1	0	A
Novice 4	0	0	B
Novice 5	0	0	B
Novice 6	1	2	C
Novice 7	1	0	A
Novice 8	0	1	Outlier
Novice 10	1	1	D
Novice 11	1	1	D
Novice 13	1	2	C
Novice 14	1	1	D
Novice 16	0	0	B
Novice 17	1	0	A
Novice 20	1	1	D
Novice 21	1	2	C
Novice 22	1	1	D
Novice 23	1	1	D



(a) X-coordinate vs Fixation index (b) Y-coordinate vs Fixation index

Figure 4.8. DBSCAN scan pattern clustering result after removing outliers

By employing two-sided approaches of X-coordinate vs Fixation Index and Y-coordinate vs Fixation Index, the model suggested four major patterns of the participant's strategy to monitor the given log indicators: four (4) scanning behaviors from novice 1, 3, 7, and 17 are grouped into a cluster 'A', another four (4) scanning results from novice 2, 4, 5 and 16 comprised a cluster 'B', three (3) novices 6, 13, and 21 formed a distinctive cluster 'C' and the rest of six (6) novices 10, 11, 14, 20, 22 and 23 was considered as a cluster 'D'. Figure 4.9 briefly shows the approximate "X, Y" coordinates of the log area of interest for the clustering.

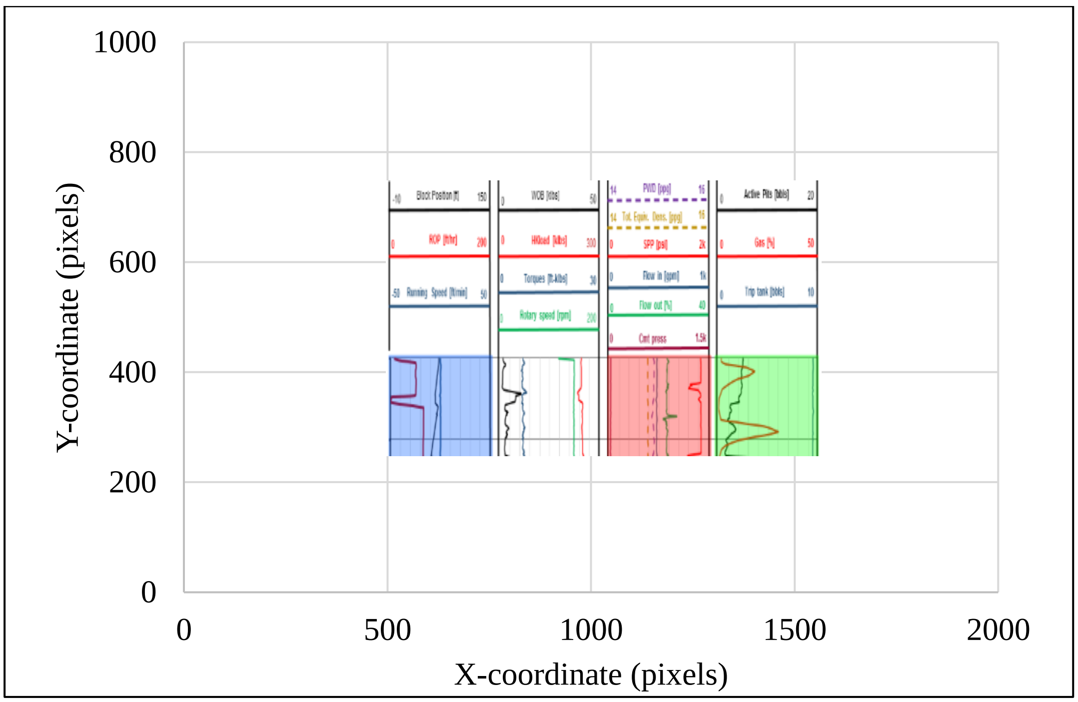
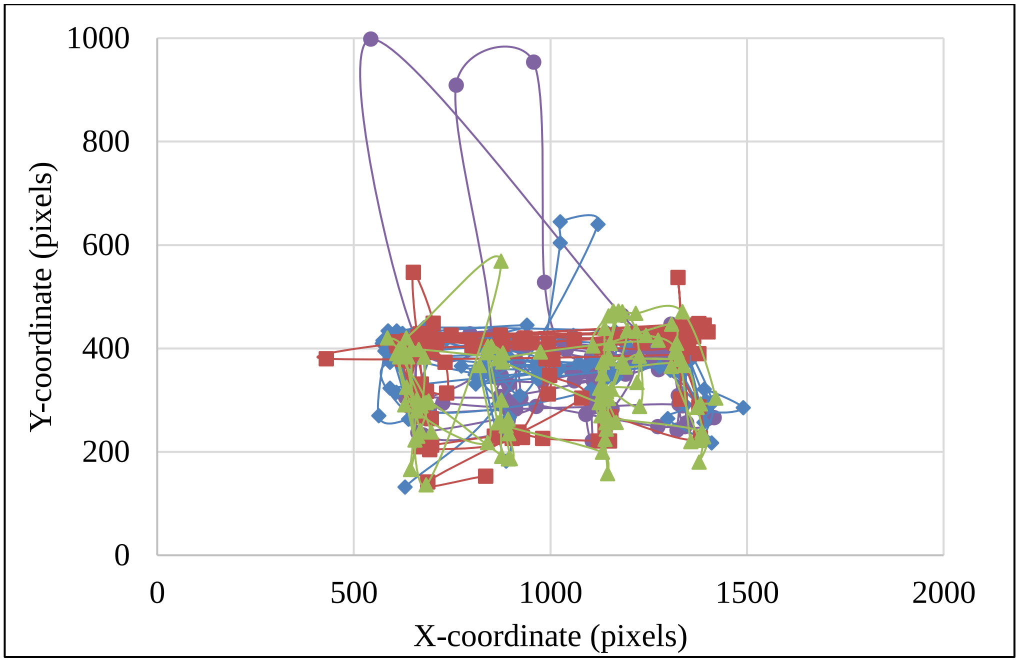


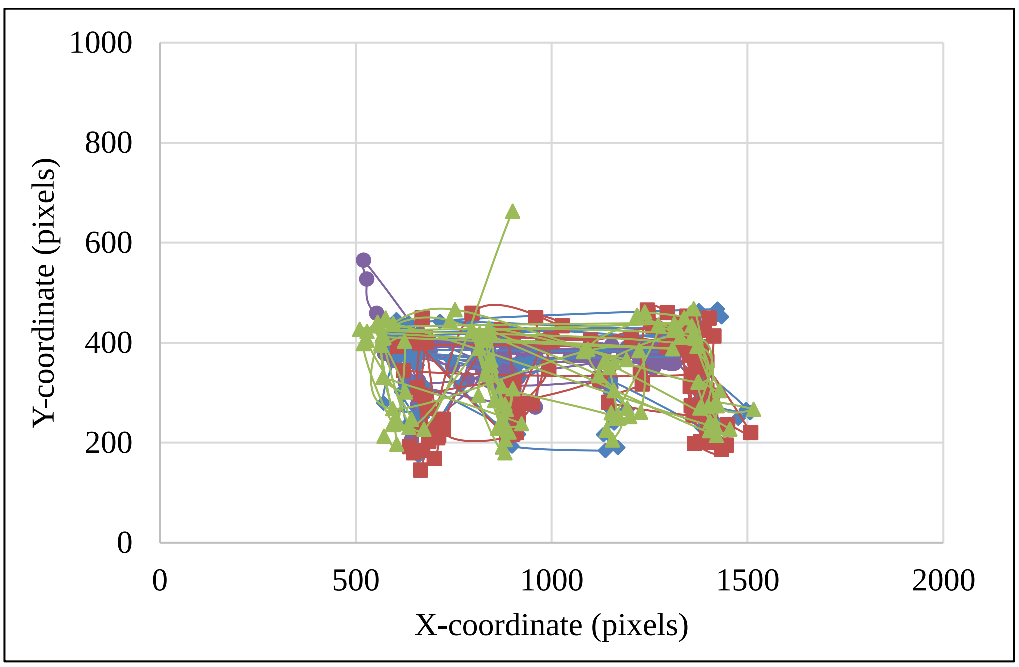
Figure 4.9. Real-time log scenario's "X, Y" coordinates

The real-time log scenario started from the 400-pixel line on the Y-axis and propagated downwards. The log area was vertically partitioned into four regions where AOI1 corresponded to the first region (blue shade), AOI2 were mapped to the third region (red shade), and AOI3 covered the fourth region (green shade).

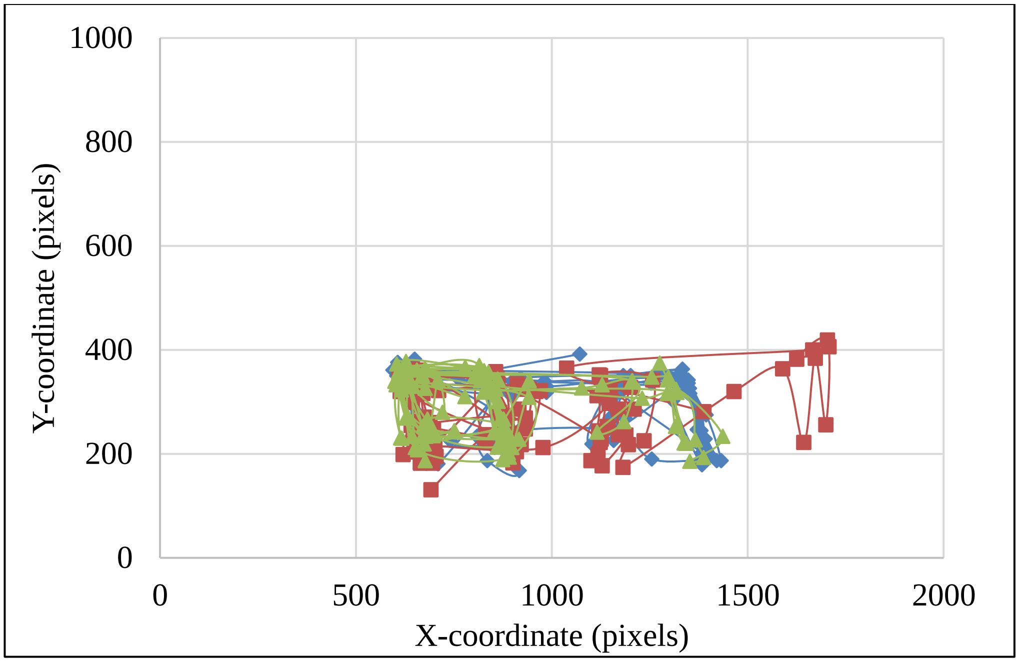
Figure 4.10 visualizes the ocular scanning trajectories of the four clusters projected on the original "X, Y" space.



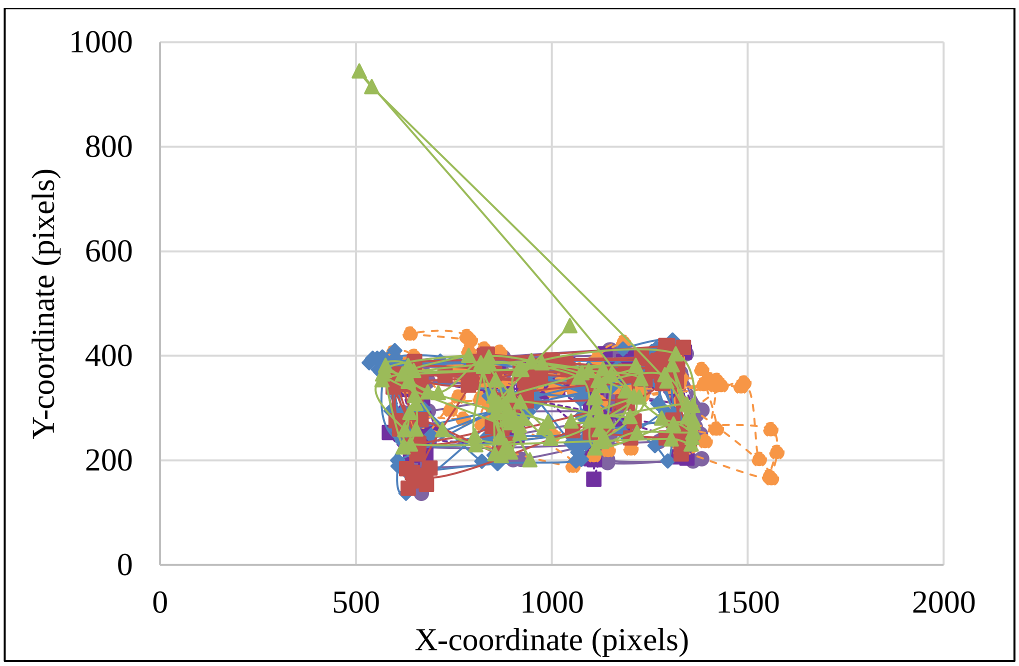
(a) Cluster "A": 4 participants



(b) Cluster "B": 4 participants



(c) Cluster “C”: 3 participants



(d) Cluster “D”: 6 participants

Figure 4.10. Four clusters of scan pattern for real-time monitoring task

Compared to other clusters, pattern ‘A’ showed several scanning deviations to upper part of the monitoring screen. The deviations imply that the participants intended to check the legend of the logs more than the participants in other clusters, as the legend box was displayed on top of the logs. Participants in pattern ‘A’ also tended to switch their attention from one log area to another throughout the monitoring tasks.

Scan pattern ‘B’ had relatively uniform distribution on the four vertical regions of log area. This can be understood in a way that the participants in this group showed a tendency to try to collect the information from all logs. Although the transitions between the logs frequently occurred in the beginning of the task as presented by a number of horizontal trails in Figure 4.10(b), the participants generally focused on each log indicator rather than hopping between the logs during the task. The result that transitions happened mainly in the beginning of the log provides a reasonable inference of the participants’ reaction. Since the ROP log indicator in AOI1 experienced an abnormal change in the early phase of the task, the scan pattern ‘B’ might indicate a high and/or prompt SA of the participants on this event. Indeed, the average verbal response score of cluster ‘B’ to AOI1 was the second highest after cluster ‘C’.

In cluster ‘C’, the logs in the left two regions received higher attention than the right regions of the area. That is, the participants intensively focused on the ROP log indicator in AOI1, which probably led to a better SA on this indicator as proved by the highest average verbal response score for AOI1. In addition, pattern ‘C’ demonstrated that the transitions were often made between the first and second regions, and between the third and fourth region, but barely happened between second and third regions.

Scan pattern 'D' displayed the most scattered attention to every log. The eye fixation points roamed across the entire log area and created complex transitions between the four vertical regions. Hence, compared to other patterns, 'D' showed a lumplike ocular behavior.

The analyses of scan pattern integrated with the multi-metric study discussed in Section 4.1 provided a supportive understanding to assess the SA of human operators. Table 4.3 summarizes the average verbal response score of the clusters to the AOIs. Considering that the verbal response score was used to quantify the SA, it can be concluded that the participants showing the scan pattern B or C had relatively high level of SA than the participants showing the scan pattern A or D. As can be seen in Figure 4.10, the pattern B and C presented discrete attention to the log regions with comparably less transitions between the logs than the pattern A and D indicated. The result might suggest that if an operator retrieved visual stimuli and tended to interpret the meaning with frequent attention changes between the objects of interest, the cognitive performance such as anomaly detection or decision-making may deteriorate, hence leads to a poor SA.

Table 4.3. Average verbal score of the clusters to AOIs

Cluster	AOI1	AOI2	AOI3
A	2.50	2.25	8.75
B	4.50	3.00	9.75
C	6.67	5.00	9.33
D	2.67	2.33	8.50

Chapter 5: Conclusion and Recommendation

5.1 Conclusion

Based on the scenario-based experimental investigation equipped with a high-fidelity virtual simulator and eye tracking device, this study provided a valuable insight to learn human operator's cognitive performance in a time-critical task of interrogating log indicators. The interpretation of the results drew some notable conclusions.

Statistical analysis of interrogated Areas of Interests (AOIs):

(1) Novice operators more focused on the Gas % trends (i.e. AOI3) and less focused on other indicators (i.e. AOIs 1 and 2). The reason seems that the drastic incremental trend of the Gas % indicator drew more visual attention compared to other indicators. Therefore, their visual attention and comprehension (i.e. verbal response scores) of the other trends were low. On the other hand, the expert showed somewhat equally divided visual attention to most of the indicators, hence obtained better comprehension scores for the important indicators.

(2) The proposed mixed approach of analyzing eye tracking metrics along with newly developed verbal response scoring metrics allowed a holistic evaluation of human behavior. In general, eye fixation durations were somewhat related to the participants' comprehension level. The study results showed an agreement to this assertation. Moreover, the proposed approach might be used to support the analysis of situation awareness in high risk operation environment.

Adapted time series clustering of visual scan paths:

(1) The adaptation of time series clustering model deconstructed the lump of eye tracking data into several groups to explore the implicit trends of visual attention and accommodated the temporal sequence into the eye fixation analyses without the need of defining AOIs.

(2) The application of the adapted model in the offshore oil and gas drilling log monitoring task showed that those who had tendencies to visually scan vertically (to spend more time on interrogating the historical trend of each indicator) scored higher in their comprehension evaluations compared to those who had tendencies to frequently go back and forth between multiple indicators in a horizontal manner.

5.2 Limitation and Recommendation

The results of the study are confined for general application by several limitations. Recommendations for future studies based on these limitations are listed.

(1) One limitation is due to the experimental design that there was no available action other than verbal announcement for the participant to take in order to inform their awareness of the situation. To validate the quantification of the participant's SA and evaluate the decision quality, more distinctive measures need to be provided.

(2) The total length of time series data was trimmed to be the same for all data sets for the ease of study. Since the actual length of time series varies with the research interest, the different lengths need to be considered.

(3) The clustering analyses was limited to the early phase of the task to reduce the computation time and to effectively apply the model algorithm. To obtain the complete understanding of the scan pattern on the overall task, the study needs to be extended to the entire test data.

(4) The clustering analyses employed 2-dimensional time warping method separately (i.e. X-coordinate vs timestamp, and Y-coordinate vs timestamp) and combined the outputs. For the investigation on the interactions between the "X, Y" coordinate elements and for robust clustering analysis, it is necessary to conduct 3-dimensional time warping where "X, Y" coordinates and temporal sequence are incorporated together.

References

Aach, J., & Church, G. M. (2001). Aligning gene expression time series with time warping algorithms. *Bioinformatics*, *17*(6), 495-508.

Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering–A decade review. *Information Systems*, *53*, 16-38.

Alcock, R. J., & Manolopoulos, Y. (1999, August). Time-series similarity queries employing a feature-based approach. In *7th Hellenic conference on informatics* (pp. 27-29).

Bar-Joseph, Z., Gerber, G., Gifford, D. K., Jaakkola, T. S., & Simon, I. (2002, April). A new approach to analyzing gene expression time series data. In *Proceedings of the sixth annual international conference on Computational biology* (pp. 39-48). ACM.

Berndt, D. J., & Clifford, J. (1994, July). Using dynamic time warping to find patterns in time series. In *KDD workshop* (Vol. 10, No. 16, pp. 359-370).

Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2017). Quantifying situation awareness of control room operators using eye-gaze behavior. *Computers & Chemical Engineering*, *106*, 191-201.

Brown, M. B., & Forsythe, A. B. (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association*, *69*(346), 364-367.

Caiani, E. G., Porta, A., Baselli, G., Turiel, M., Muzzupappa, S., Pieruzzi, F., Crema, C., Malliani, A., & Cerutti, S. (1998, September). Warped-average template technique to track on a cycle-by-cycle basis the cardiac filling phases on left ventricular volume. In *Computers in Cardiology 1998* (pp. 73-76). IEEE.

Chu, S., Keogh, E., Hart, D., & Pazzani, M. (2002, April). Iterative deepening dynamic time warping for time series. In *Proceedings of the 2002 SIAM International Conference on Data Mining* (pp. 195-212). Society for Industrial and Applied Mathematics.

Cokely, E. T., Feltz, A., Ghazal, S., Allan, J. N., Petrova, D., & Garcia-Retamero, R. (2017). Decision making skill: From intelligence to numeracy and expertise. *Cambridge handbook of expertise and expert performance*.

Dadmohammadi, Y., Salehi, S., Kiran, R., Jeon, J., Kang, Z., Cokely, E. T., & Ybarra, V. (2017, October). Integrating Human Factors into Petroleum Engineering's Curriculum: Essential Training for Students. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Deacon, T., Amyotte, P. R., & Khan, F. I. (2010). Human error risk analysis in offshore emergencies. *Safety science*, *48*(6), 803-818.

Debrégeas, A., & Hébrail, G. (1998, August). Interactive Interpretation of Kohonen Maps Applied to Curves. In *KDD* (Vol. 1998, pp. 179-183).

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors*, *37*(1), 32-64.

Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998, October). A comparative analysis of SAGAT and SART for evaluations of situation awareness. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 42, No. 1, pp. 82-86). Sage CA: Los Angeles, CA: SAGE Publications.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd* (Vol. 96, No. 34, pp. 226-231).

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: the dtw package. *Journal of statistical Software*, *31*(7), 1-24.

Giorgino, T., & Giorgino, M. T. (2012). Package 'dtw'. *R package version,* 1.18-1

Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006, April). The validity of the stimulated retrospective think-aloud method as measured by eye tracking. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1253-1262). ACM.

Haass, M. J., Matzen, L. E., Butler, K. M., & Armenta, M. (2016, March). A new method for categorizing scanpaths from eye tracking data. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (pp. 35-38). ACM.

Hahsler, M., & Piekenbrock, M. (2017). dbscan: Density Based Clustering of Applications with Noise (DBSCAN) and Related Algorithms. *R package version*, 1.1-1.

Härtel, C. E. J., Smith, K. A., & Prince, C. (1989). Defining aircrew coordination: Searching mishaps for meaning. In *Fifth International Symposium on Aviation Psychology*. Ohio State University.

Helmreich, R. L. (2000). On error management: lessons from aviation. *Bmj*, *320*(7237), 781-785.

Howell, D. C. (2012). *Statistical methods for psychology*. Cengage Learning. Duxbury. (pp. 324–325).

Kadous, M. W. (1999, June). Learning Comprehensible Descriptions of Multivariate Time Series. In *ICML* (pp. 454-463).

Keogh, E. and Kasetty, S., 2003. On the need for time series data mining benchmarks: a survey and empirical demonstration. *Data Mining and knowledge discovery*, *7*(4), 349-371.

Kiware, S. (2010). *Detection of outliers in time series data*. Marquette University.

Kriegel, H. P., Kröger, P., Sander, J., & Zimek, A. (2011). Density-based clustering. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *1*(3), 231-240.

Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern recognition*, *38*(11), 1857-1874.

Lin, J., Vlachos, M., Keogh, E., & Gunopulos, D. (2004, March). Iterative incremental clustering of time series. In *International Conference on Extending Database Technology* (pp. 106-122). Springer, Berlin, Heidelberg.

Ma, R., & Angryk, R. (2017, November). Distance and Density Clustering for Time Series Data. In *Data Mining Workshops (ICDMW), 2017 IEEE International Conference on* (pp. 25-32). IEEE.

Minitab 17 Support. (2015). *One-Way ANOVA*. Retrieved from http://support.minitab.com/en-us/minitab/17/Assistant_One_Way_ANOVA.pdf

Norwegian Petroleum Directorate Guidance for Drilling and Well Activity. (2004). Well integrity in drilling and well operations. Norsok D-010 standard Rev3.

Rai, P., & Singh, S. (2010). A survey of clustering techniques. *International Journal of Computer Applications*, *7*(12), 1-5.

Rani, S., & Sikka, G. (2012). Recent techniques of clustering of time series data: a survey. *International Journal of Computer Applications*, *52*(15).

Ratanamahatana, C. A., & Keogh, E. (2004, August). Everything you know about dynamic time warping is wrong. In *Third workshop on mining temporal and sequential data*. Citeseer.

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package 'mass'. *R package version*, 7.3-47

Roberts, R., Flin, R., & Cleland, J. (2015). Staying in the zone: Offshore drillers' situation awareness. *Human factors*, *57*(4), 573-590.

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing*, *26*(1), 43-49.

Sanfilippo, F. (2016, June). A multi-sensor system for enhancing situational awareness in offshore training. In *Cyber Situational Awareness, Data Analytics And Assessment (CyberSA), 2016 International Conference On* (pp. 1-6). IEEE.

Selcon, S. J., & Taylor, R. M. (1990). Evaluation of the Situational Awareness Rating Technique(SART) as a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 8p (SEE N 90-28972 23-53)*.

Sharma, C., Bhavsar, P., Srinivasan, B., & Srinivasan, R. (2016). Eye gaze movement studies of control room operators: A novel approach to improve process safety. *Computers & Chemical Engineering*, *85*, 43-57.

Sneddon, A., Mearns, K., & Flin, R. (2006). Situation awareness and safety in offshore drill crews. *Cognition, Technology & Work*, *8*(4), 255-267.

Sneddon, A., Mearns, K., & Flin, R. (2013). Stress, fatigue, situation awareness and safety in offshore drilling crews. *Safety Science*, *56*, 80-88.

The Minitab Blog. (2015, February 19). *Choosing Between a Nonparametric Test and a Parametric Test [Web log post]*. Retrieved from http://blog.minitab.com/blog/adventures-in-statistics-2/choosing-between-a-nonparametric-test-and-a-parametric-test

Tobii Technology AB. (2010, November). Tobii Pro TX300 Eye Tracker Product Description. Retrieved from http://www.tobiipro.com/siteassets/tobii-pro/product-descriptions/tobii-pro-tx300-product-description.pdf/?v=1.0

Tobii Technology AB. (2016, January). Tobii Studio User's Manual. Ver. 3.4.5. Retrieved from https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-studio-user-manual.pdf/?v=3.4.5

US Chemical Safety and Hazard Investigation Board. (2010). Investigation Report Drilling Rig Explosion and Fire at the Macondo well. Government Printer, 2010-10-I-OS, Vol 3.

Van Gog, T., Paas, F., & Van Merriënboer, J. J. (2005). Uncovering expertise-related differences in troubleshooting performance: combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, *19*(2), 205-221.

Venables, W. N., & Smith, D. M. (2004). An Introduction to R: Notes on R, A Programming Environment for Data Analysis and Graphics, v. 2.0. 1. *Network Theory, Bristol, UK*.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose.

Williams, B., Quested, A., & Cooper, S. (2013). Can eye-tracking technology improve situational awareness in paramedic clinical education?. *Open access emergency medicine: OAEM*, *5*, 23.

Xue, Q., Wang, Y., Zhai, H., & Chang, X. (2018). Automatic Identification of Fractures Using a Density-Based Clustering Algorithm with Time-Spatial Constraints. *Energies*, *11*(3), 563.