INFLUENCES OF DISPLAY DESIGN AND TASK MANAGEMENT STRATEGY ON SITUATION AWARENESS, PERFORMANCE, AND WORKLOAD IN PROCESS CONTROL ENVIRONMENTS

THESIS

James R. Bowden

AFIT-ENV-MS-14-D-29

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-MS-14-D-29

INFLUENCES OF DISPLAY DESIGN AND TASK MANAGEMENT STRATEGY ON SITUATION AWARENESS, PERFORMANCE, AND WORKLOAD IN PROCESS CONTROL ENVIRONMENTS

THESIS

Presented to the Faculty

Department of Systems Engineering and Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Systems Engineering

James R. Bowden, BS

December 2014

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENV-MS-14-D-29

INFLUENCES OF DISPLAY DESIGN AND TASK MANAGEMENT STRATEGY ON SITUATION AWARENESS, PERFORMANCE, AND WORKLOAD IN PROCESS CONTROL ENVIRONMENTS

James R. Bowden, BS

Approved:

//SIGNED//	5 Dec 2014
Christina Rusnock, Maj, USAF, PhD (Chair)	Date
//SIGNED//	5 Dec 2014
Michael Haas, PhD, PE (Member)	Date
//SIGNED//	5 Dec 2014
Brent Langhals, Lt Col, USAF, PhD (Member)	Date

AFIT-ENV-MS-14-D-29

Abstract

Process control environments demand well informed high performing human monitors to maintain effectual control of multiple processes. Most research aims to satisfy this requirement through the evaluation of competing heuristic-based display design constructs. Contrary to that method, this study takes a novel approach by examining both factors internal and external to the human observer to identify where beneficial outcomes actually reside. External factors explore the underlying design construct attributes, while internal factors focus on the effect of operator task management strategy, age, and experience. Results from this study present several key findings relative to operator situation awareness, performance, and workload. Findings suggest the specific manner in which external information is presented and oriented on a process control room display is inconsequential toward situation awareness and performance. Further, operator preferred task management strategy has a profound effect on their performance and experienced workload, while exhibiting only a mild effect on situation awareness. In most cases, an Adaptive Attack strategy produces desirable results, while an Adaptive Avoidance does not. Interleaving and Multitasking fall between these two extremes. Lastly, findings indicate subject variables, age and experience have negative effects on overall situation awareness and system deviation prediction times.

AFIT-ENV-MS-14-D-29

To my wife and three outstanding young men I am fortunate to call upon each as "son"

Acknowledgments

I would like to extend my sincere appreciation and gratitude to my faculty advisor, Maj Christina Rusnock for devoting countless hours toward this endeavor and facilitating the development of an outstanding product. I would like to thank Dr. Michael Haas and Lt Col Brent Langhals, both of whom imparted upon me the requisite foundational knowledge to begin and complete this pursuit. I would also like to express my sincere appreciation for the unwavering support and encouragement I received from my leadership, coworkers, and friends – past and present – at the Air Force Research Laboratory; Mr. John Fedon, Mrs. Cheryl Skipper, Dr. Richard Fingers, Mr. Curtis Kessler, Mrs. Lesley Murawski, Mr. David Dye, Mr. Jack Vondrell, all of whom serve as outstanding mentors and role models; Mr. Mike Youngker for his expertise and assistance in making the experimental designs a reality; Mr. William "Pete" Buckley for his assistance with experiment volunteers and beyond; the Component Research Air Facility crew for being perpetually first class; the Instrumentation & Controls technicians who have borne the brunt of the workload in my stead and each of whom I am honored to interact with on a daily basis.

James R. Bowden

Table of Contents

Page

List of Figures

List of Tables

Page

INFLUENCES OF DISPLAY DESIGN AND TASK MANAGEMENT STRATEGY ON SITUATION AWARENESS, PERFORMANCE, AND WORKLOAD IN PROCESS CONTROL ENVIRONMENTS

I. Introduction

Background

Two recent industrial process control tragedies in the chemical and oil refinement field's, respectively, are the Bayer CropScience pressure vessel explosion in Institute, WV (U.S. Chemical Safety and Hazard Investigation Board, 2011) and the BP oil refinery explosion in Texas City, TX (U.S. Chemical Safety and Hazard Investigation Board, 2007). Combined, these two incidents resulted in the loss of 17 lives and inflicted injuries upon an additional 188 individuals. These figures are quite sobering and point to exactly how dangerous the process control industry can be when operator situation awareness (SA) is incomplete, especially when put into the context these were only two incidents that resulted in such a high number of casualties. In both cases, the processes under human control were not directly observable by the equipment operators; thus, the operators had to rely heavily upon information transmitted back to them in a central control room by a host of automation mechanisms to include panel board indicators and user interface displays. The BP centralized process control room is shown in Figure I-1. Looking at this arrangement consisting of no less than 11 displays, it is easy to see how SA is lost.

Figure I-1 BP Texas City Control Room Layout
Representative of a typical centralized process control room layout consisting of multiple displays to monitor remote processes. (U.S. Chemical Safety and Hazard Investigation Board, 2007)

Naturally, maintaining appropriate operator SA is more difficult with remotely controlled equipment than when an operator is fully immersed within the affected environment, therefore remote operations present many difficulties and challenges that can have a negative impact on a human monitor. Among those in the visual field are limitations to the operator's view of the system, latency of information presented at the remote location, and a limited depth or richness in context to the information provided through the user interface (Chen, Haas, & Barnes, 2007). Further, the very manner in which information is displayed to an operator can seriously degrade SA and allow a dangerous situation to unfold. This is noted in these excerpts taken from the incident reports of the BP and Bayer catastrophes:

On the day of the incident, however, the computerized control system display provided neither flow data in and out of the raffinate unit on the same display screen, nor a material balance calculation, hindering the Board Operator's ability to recognize the need to send liquid raffinate to storage.

The detailed process equipment displays in the DCS were difficult to navigate. Routine activities like starting a reaction or troubleshooting alarms would require operators to move between multiple screens to complete a task, which degraded operator awareness and response times.

The old control system used "percent full" to indicate the level in a vessel, but the new control system listed the level in total gallons inside the vessel.

These key insights reinforce the position that operator SA immediately preceding each disastrous event was negatively impacted simply by the control room interface display design. The manner in which crucial process control information was being cognitively managed by the operators who made these statements points to an inability to maintain an accurate model of system status due in large part to insufficient methods of information presentation. Compounding matters is the fact numerous processes are in need of monitoring, meaning if any observer is to remain on top of the system pictured in Figure I-1 interface design needs to support rather than inhibit SA.

Technological advances in automation have ushered about the integration of instrumentation and controls capable of collecting and disseminating massive amounts of data over virtually limitless distances. The impact this has had on human-automation interaction in process control environments has led to higher degrees of automated processing with the human serving primarily as system observer. Reduced human interaction as the result of increased automated processing of information creates challenges for maintaining operator SA via the control room interface. Thus, uncovering those informative and underlying display design attributes that allow an observer to intuitively identify systematic failures would be beneficial in the development of future display designs (Hancock, Jagacinski, Parasuraman, Wickens, Wilson, & Kaber, 2013). The control room interface must be more effective at clearly communicating information to keep pace with fewer operators who interact with the system less frequently than ever before.

The requisite human oversight necessary to effect control has been reduced by great economies of scale: it simply requires less human capital to oversee more systems when information is consolidated into a centralized point of control. For this reason, the relentless migration away from decentralized control philosophies toward more centralized oversight of multiple operations has become the new norm. Evidence of this exists through the adoption and implementation of centralized control schemas across a wide spectrum of industries – both private and public – from the manufacturing plant floor to the military's utilization of unmanned aircraft. The benefits of centralized control are agreeably many; however, they have not come without many tradeoff challenges in the need to intelligently represent the increased onslaught of information to fewer and fewer human monitors. Cummings, Bruni, and Mitchell (2010) reflect upon these challenges by considering incidents in both government and private industry involving network-centric operations. They identify ten specific challenges attributable to either the technology used or human performance characteristics, ranging from items such as information overload to multimodal technologies. To counter these effects, more often than not, heuristic guidelines are established to accommodate the translation of raw data in a central control room through the human machine interface (HMI). This is evidenced in several previous works where guidelines are established to aid in the development of an overall “best” display design construct (Norman, 1984; Gould, 1988; Vicente & Rasmussen, 1992; Ponsa, Vilanova, Perez, & Andonovski, 2010; Shneiderman, Plaisant, Cohen, & Jacobs, 2010). Problematic with such guidance is it is severely lacking in addressing factors internal to the human operator.

How operators internally manage the information presented to them could be equally as important as the external interface design construct. While technological progress has resulted in consolidation to centralized control room architectures, tasks previously handled by multiple operators, have been increasingly consolidated into the responsibility of fewer personnel. Not only are control room operators faced with juggling multiple processes under their purview of control, but they must also perform related yet dissimilar secondary tasks associated with routine facility management. The manner in which individuals go about handling more than a single task becomes highly relevant in centralized control operations. One individual’s preferred task management strategy to cope with multiple tasks may be more advantageous than another. Task execution when switching between tasks can have a profound impact on SA, task performance, and perceived workload. The work of Morgan, et al(2013) has revealed an individual’s ability to adapt the way they manage multiple tasks varies based in part on the chosen task management strategy. How this influences operator SA and task performance in a central control environment is worthy of further exploration as suggested in the Morgan, et al. research.

Both internal and external challenges exist in search of human-machine symbiosis. It is believed these challenges are not completely insurmountable. This research seeks to advance the theory that the manner of information presentation at the display attribute level – external to a human observer overseeing multiple processes – can be manipulated to produce positive outcomes toward operator SA and task performance. In addition, how internal factors such as operator task management strategy and demographics play a role in a central process control environment are also explored.

Problem Statement

The current body of process control interface design knowledge and research has not delved deeply enough into the industry's need for tangible evidence toward appropriate interface attributes that will improve operator SA and task performance. To address the need of information presentation for multiple process control, heuristics and best practices have often been applied, yet problems still exist and are evidenced whenever a catastrophic breakdown of SA contributes to or is directly attributed to a process control disaster. Visual information presentation to a human observer plays a crucial role in how modern day operators rely upon external factors to assess both acceptable and unacceptable system status relative to their mental model of processes under their direct control. It is therefore imperative that operators have the ability to recognize system changes immediately via the user interface and be able to perceive, comprehend, and project a system's current state in order to react appropriately. Development of designs that support this requires the identification of display attributes that externally enhance operator awareness. Because previous research has focused solely on external factors, a gap exists in how internal factors to the human monitor also play a part in maintenance of operator SA.

Research Objective

The primary focus of this research effort is to determine whether competing process control information display designs provide for beneficial outcomes toward operator SA and task performance when task management strategies are taken into consideration. To meet this objective the fundamental attributes of interface designs needed to be studied. Competing methods of information presentation (numeric vs. graphic) and how that information is oriented (functionally grouped vs. spatially mapped) on a process control display are identified and investigated as external factors. In addition to this, internal factors for task management strategy are identified and defined (Interleaving, Multitasking, Adaptive Attack, Adaptive Avoidance) along with subject demographic information to support analysis of human behavior to determine if either has an impact on SA, task performance, or workload. The five investigative questions and respective hypotheses for both internal and external factors are described below. The first two questions address factors external to the human through investigation of display design attributes and overall constructs, while the remaining questions address factors internal to the human such as preferred task management strategy employment and subject variables with respect to individual demographics.

Investigative Questions / Hypotheses

Hypothesis formulation for external factor investigative questions, 1 and 2 are based on findings in the literature suggesting the use of a functional grouping orientation and graphical means of information presentation for display design yield positive SA and performance outcomes (Handal & Ikuma, 2012; Tharanathan, Bullemer, Laberge, Reising, & McLain, 2012). The questions are tailored to address gaps in these previous works by examining the underlying design attributes used in a series of competing designs instead of the overall aggregate design construct as a whole. While both questions present an approach that differs from previous studies, they reflect the anticipation of an ability to duplicate previous findings, which are reflected in the hypotheses that follow.

1. How does the process control information display construct used during an interactive monitoring task impact levels 1, 2, and 3 SA?

   It is hypothesized a graphic means of information presentation and functionally grouped orientation will result in higher level 1, 2, and 3 SA.

2. How does the process control information display construct impact primary and secondary task measures of performance?

   It is hypothesized a graphic means of information presentation and functionally grouped orientation will result in higher primary and secondary task performance.

The remaining investigative questions 3 through 5 focus on internal factors that influence SA and performance. Hypothesis formulation for question 3 builds upon the finding of Tombu and Jolicoeur (2004) which indicates individuals who engage in multitasking activities experience negative outcomes. Question 4 is grounded in the work of Morgan, et al. (2013) which introduces an architecture toward identification of consistent performance through individual task adaptation. Using the Morgan, et al. architecture, this study anticipates an increase in time spent on a given task in a multiple task environment yields a more favorable outcome for the favored task. The final investigative question focuses on individual demographics and posits the inherent differences that exist in individuals will be reflected in both SA and performance outcomes.

3. In what way does the task management strategy utilized during a process control monitoring activity affect operator levels 1, 2, and 3 SA?

   It is hypothesized a multitasking task management strategy will result in lower level 1, 2, and 3 SA.

4. How does operator task management strategy impact primary and secondary task measures of performance?

   It is hypothesized both adaptive task management strategies (Adaptive Attack and Adaptive Avoidance) will have positive outcomes on the primary task and a negative effect on the secondary task.

5. How do subject variables affect overall SA and primary task performance?

   It is hypothesized individual demographic differences exist that will have a negative effect on overall SA and positive effect on primary task performance.

Methodology

This research methodology follows a multi-phased approach (Appendix A). Essential to this are the supporting objectives undertaken prior to execution of the formal research experiment. These include:

- Establishing the feasibility of a formal research study into display design and task management strategy by conducting a case study and cognitive task analysis (CTA) at a relevant process control facility.

- Developing a formal system description and underlying task networks in support of how information should be presented in a multi-task, multiple process centralized control environment.
- Identifying suitable competing design constructs and underlying attributes based on real world applications.
- Determining the appropriate metrics and generating the questions necessary to gauge SA, performance, and workload during a multi-task, multiple process activity.
- Executing a pilot study to validate the appropriate level of task load to produce results of relevancy for a multiple task simulated environment.

A case study of the Air Force Research Laboratory's (AFRL) Component Research Air Facility (Air Force Research Laboratory / Aerospace Systems Directorate, 2014) was conducted and a cognitive task analysis (CTA) completed during active facility operations in the fall of 2013 (Appendix B). The Component Research Air Facility provides an appropriate case example because of its large industrial complex layout and its use of a central control room to monitor multiple geographically dislocated processes. Observation and interview data from the CTA are used to establish a formal description of the system under investigation as well as generate a series of hierarchal task analysis (HTA) networks to aid in the development of four competing experimental interface designs.

An AFIT internal review board (IRB) exemption request was granted prior to commencement of any work involving human subjects (Appendix C). The next phase of the research methodology began with a pilot study to validate the experimental design and test apparatus through subject matter experts and active experimentation. The pilot study was followed by a formal 2x2 within subjects experiment using 24 participants following a Latin Square design. Data collection transpired over four 30-minute trials using each of the competing display designs (numeric or graphic; functionally grouped or spatially mapped) as part of a primary task executed simultaneously with a secondary reading comprehension task. A host of real time data was collected automatically by the experimental setup and through direct researcher observation. Ancillary informative data was also captured through demographic, pre-, and post-experimental questionnaires (Appendix D) completed by all participants.

Assumptions/Limitations

The experimental setup, display designs, and test location were heavily scrutinized for applicability. Feedback from the pilot study was integrated into the final experimental setup and assumed to have led to the most robust means of data collection possible; within the operational constraints of the available equipment and area housing the experiment. Attempts were made to make the process control simulation experience as consistent from participant to participant as possible. Realism was also a concern. It was assumed the findings from the laboratory setting translate to the real world with minimal of consequence, however a known limitation to laboratory research is it can only closely reflect research conducted *in situ*, or furthermore actions in the real world. Because the researcher was collocated within the context of the experimental environment there exist potential biases relative to how the participant interpreted the researcher's presence. Training attempted to mitigate the effects of researcher presence, but it is not possible to know to what extent this was successful. Furthering this translation to the real world, it is assumed the results from a 3-4 hour total experiment contact time for each participant produced results that are relevant to an industry standard 8-hour process monitoring shift. Lastly, the pre- and post- experiment questionnaires administered to all participants intended to capture an extremely broad combination of factors taking into consideration participant performance and other variables outside of the researcher's control. Total elimination of confounding behaviors such as errors of omission, failure to act, and consistently vigorous participation by each test subject was never guaranteed. However, it is assumed all participants took their participation seriously and gave the most honest and precise of answers possible at all times – to include the responses on the demographic and post experimental feedback questionnaires. Subject privacy and assurances of freedom from reprisal were well communicated to each participant, but it must be considered an implied limitation that not every subject was comfortable providing the most candid of answers to someone they did not know.

Implications

Results from this body of work seek to contribute to the field of process control by providing insightful perspectives toward end user SA, task performance, and workload in a centralized control environment. Future Component Research Air Facility interfaces will be constructed with the results of this research in mind. Other sectors of the broader process control industry may see benefits as well. Application examples include: the power industry regulating distribution of resources across a large grid network, the nuclear power industry monitoring complex large scale reactor processes, the oil and chemical industries monitoring refinement and chemical processing, the waste water treatment for many municipalities maintaining the hygienic integrity of processed water, and the mining industry monitoring subterranean hazardous vapor detection assets because they utilize central process control architectures similar to the Component Research Air Facility. Even military applications could see some degree of benefit relevant to unmanned flight. All of these remotely controlled process activities would benefit from the application of interface designs intent on improving operator SA and task performance while reducing experienced workload. But findings also have the potential to cross over into other fields and applications that do not involve a central control room at all, since many of the cognitive tasks performed by process control operators (e.g. use of external displays to communicate information, vigilance monitoring task, multiple task environment) are performed in kind beyond the process control industry. Examples of this include the transportation sector and TSA baggage screeners examining luggage at an airport terminal while also monitoring passenger behavior, the automotive industry line worker viewing a display to monitor productivity and quality control while executing an assembly task, and the agricultural industries implementation of autonomous farm monitoring where farmers track asset location and concurrently examine crop yield data.

Preview

This introductory chapter conveys the essence of the experimental research and detailed body of work that follows. Chapter 2 contains a review of the literature into situation awareness (SA) and task management strategies as both relate to process control environments. A conference paper and journal article address the investigative questions presented in the introduction and build upon a review of the literature. Both are presented in subsequent chapters where Chapter 3 contains a draft conference paper based upon investigation into the effects of display design outcomes toward SA and task performance. It also addresses the first two investigative questions with results from an analysis of the data reflecting little findings of significance toward the external factor, display design constructs and their underlying attributes. Chapter 4 addresses the three remaining investigative questions and presents findings in a draft journal article format. The journal article reports the effects of the internal factor, task management strategy on operator SA, task performance, and workload and shows high degrees of significance toward all three outcomes. It also details two subject variables, age and experience, to answer the fifth and final investigative question. Chapter 5 begins with a brief overview of this research effort and further explores the investigative questions. It concludes by offering suggestions for future work as they can be applied to both future experimental designs and the remaining data set yet to be investigated.

II. Literature Review

Chapter Overview

The purpose of this chapter is to cover those aspects of literature and previous research uncovered during a practical investigation into situation awareness (SA), as it relates to display design, and task management strategies as they are applied in a process control environment. The topics of SA and task management provide support for the formal research effort carried out in this body of work. Each topic is discussed in detail to formulate the relevancy to process control operations and establish the need for the research initiatives detailed in both the conference paper and journal article which are contained in chapters three and four, respectively.

Situation Awareness (SA)

Endsley's formal theory and definition of SA are well known and heavily cited as being *"the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"* (1995). Levels 1, 2, and 3 of SA as adapted here for subsequent ease of explanation are the perception, comprehension, and projection of contextual information, respectively, and can best be described in more general terms as simply having an astute understanding of one's surroundings. A common method to capture SA data from human subjects involves the Situation Awareness Global Assessment Technique (SAGAT), which has been used with a high degree of success in numerous studies to date (Endsley M. R., 2000). SAGAT involves random pauses administered during a human subject experience to effectively gauge all three levels of SA. Questions are to be terse and to the point so they detract minimally from the subjects experimental context. Some example questions are provided in Table II-1.

SA Level	Question	Response
1	Have any of the processes experienced a deviation?	Yes No
2	What two processes are running the Worst?	1 2 3 4
3	What process number will deviate next?	1 2 3 4

Table II-1 Example SAGAT Questions

SAGAT pauses for questions and answers are executed during a research experiment while a participant is actively immersed and engaged in a context specific task under formal investigation. This provides for immediacy in responses by not requiring a participant to recall later what their evaluation of then current conditions were at a later point in time.

SA and the Role of the Human Machine Interface (HMI)

The modern user interface display has become the de facto standard for monitoring remote process control operations. Often times, it is the sole means of providing a dislocated operator intelligible information as to how the process is actually running in a remote location. When this is the case, information presentation through the user interface is of paramount importance to maintaining overall operator SA. But the ability to exact a positive influence over operator SA during a central control room monitoring task has been difficult for the domain of human-automation interaction (Cummings, Bruni, & Mitchell, 2010; Li, Horberry, & Powell, 2010; Moyle, 2005). Despite the many advances in HMI technologies, and several well thought out heuristic guidelines for designs over the years (Norman, 1984; Gould, 1988; Vicente &

Rasmussen, 1992; Ponsa, Vilanova, Perez, & Andonovski, 2010; Shneiderman, Plaisant, Cohen, & Jacobs, 2010; U.S. Nuclear Regulatory Commission, 2002), process control disasters still pose a danger to many different industries (U.S. Chemical Safety and Hazard Investigation Board, 2007; U.S. Chemical Safety and Hazard Investigation Board, 2011).

User Centered Design Only Notionally Supports SA

Information presentation plays a crucial role in how modern day operators visually assess acceptable and unacceptable system status relative to their mental model of the process under their control. It is imperative the operator have the ability to recognize changes to process control data at the user interface and be able to react accordingly when necessary. The idea of a user centered design process dating back to the 1980's (Norman, 1984) has made a positive impact on interface display designers by directing them to involve the user up-front and early in the design process. It is thereby implied that a user interface adhering to a user centered design process will result in a best fit for the end user. While this has added significantly to the application of user interface design and given designers a solid starting point, user centered design does not serve as an actual metric or body of evidence to quantitatively evaluate one means of information presentation – and its impact on SA – against another. Much of the focus of user centered design revolves around ways of engaging the user in a design effort as opposed to the manner in which designers should actually answer the question of what design integration results are most effective (Carr-Chellman & Savoy, 2004). Even much of the input from the field of human factors engineering has involved qualitative best practices and usability criteria on how to incorporate the user's desires into the system design. One formal study even resulted in "*heuristic evaluation*" being deemed the best method for assessing one interface design better than another without addressing the need for an overtly quantifiable metric (Jeffries, Miller, Wharton, & Uyeda, 1991). This has become evident in multiple recent works involving user-centered design theories too, all of which culminated into the identification of myriad heuristic based best practices (Endsley & Jones, 2003; Landry & Jacko, 2004; Moyle, 2005; Panteli, Kirschen, Crossley, & Sobajic, 2013). Moreover, these works continue to heavily focus upon heuristics as a means to garner increased SA during design integration (Endsley & Jones, 2003; Panteli, Kirschen, Crossley, & Sobajic, 2013). These examples of heuristic based inputs are extremely valuable and are not without merit. However, they should instead be forming the foundations for further research into the quantitative mechanisms that constitute an effective means of appropriate interface design – a design best suited to maintain operator SA and improve task performance. Such metrics, or at a minimum, research into what manner of information display attributes are better than another are sorely needed. Investigative findings could build upon a ground swell of component level research leading to a defined set of empirically justified design criteria.

Research Trends – Interface Designs that Improve SA

An investigation of SA literature suggests future works should focus on application of existing theory in efforts to elicit the "*optimal*" state of human and machine cooperative relationships (Hancock, Jagacinski, Parasuraman, Wickens, Wilson, & Kaber, 2013). Unfortunately defining what exactly constitutes an "*optimum*" state can be elusive when no agreed upon quantifiable means to rank order competing design constructs and attributes exists. Some dialogue in the literature leads to the idea that perhaps only the operator – not the designer – truly knows the optimal state, because, given the opportunity, operators often manipulate the interfaces at their disposal to provide a more suitable environment for their own personal monitoring needs. This leads to a concept of knowledge-driven monitoring, whereby operators are constantly monitoring a plant's state limited to only those parameters they have deliberately selected to monitor (Mumaw, Roth, Vicente, & Burns, 2000). This presumption, that the user knows best, is thus subject to the inherent biases and experience of the user. While insightful to user behavior, the knowledge that users desire to customize their interfaces falls short of providing a true solution for optimal user interface design because it does not account for whether or not customization actually improves SA and performance. This is not to say user centered design best practices and knowledge-driven monitoring are not important, rather it simply points again to the lack of widely accepted quantifiable means or body of research toward evaluating interface designs for process control.

Research into user interface design should start by exploring what means of displaying information are statistically more significant for increasing SA and performance during a given task than others. Several recent studies have attempted to do just this (Scholtz, Antonishek, & Young, 2005; Huibin & Wang, 2009; Wang, Zhuang, Wei, & Wanyan, 2012; Handal & Ikuma, 2012; Tharanathan, Bullemer, Laberge, Reising, & McLain, 2012). Of these studies, two focused entirely on process control applications. Tharanathan et al. (2012), sought to identify whether or not central control room operator SA was impacted by display design type. Two heuristically developed competing interface designs – functional versus schematic – ran in a simulated process control scenario whereby pre-recorded system information was played back to an experienced human monitor. The "functional" display type relied on the heuristic best practice of spatially mapping graphic information around a simulated process object, whereas the "schematic" display design was an actual process mimic representation of numeric variables in relative location to their placement within the system's context. Findings indicated levels 1 and 2 SA were higher when using the functional versus schematic display type. The study did not attempt to address level 3 SA but did involve an assessment of participant subjective views toward usability, which was also supportive of the functional displays. One potentially confounding factor toward the findings of Tharanathan et al. was in the display construct itself. The functional display utilized a suite of newly developed graphical color indicators that were not utilized on the schematic display. This is addressed by the researchers in their explanation of the competing designs, however left out of the formal results and analysis as to why one means of information presentation was originally selected over another. Another key limitation of this study, as pointed out by the authors, was the inability for the test subjects to interact with the competing displays due to the pre-recording method. Participants monitored a video of the process they were charged to monitor and simply indicated when the process was behaving inappropriately. Unfortunately, this was done counter to how real world process control operators normally perform a monitoring task. Typically, operators will perceive (level 1 SA), comprehend (level 2 SA), and then act upon the system's controls at will to return a process to its steady state. Any future works should remedy this limitation in interaction between test subject and system control before proceeding.

A separate yet similar research effort to the Tharanathan, et al study utilized two competing process control designs to investigate the impact the interface had on SA and performance outcomes (Handal & Ikuma, 2012). This study also involved a process control specific set of competing interface designs. Design constructs for the research effort utilized heuristic evaluation prior to commencement as a means to distinguish “good” versus “poor” designs in terms of information presentation for a between subjects study. Both interface designs revolved primarily around numeric information presentation and color usage along with indicator proximity to a spatially mapped system object depiction. The “good” design utilized functional grouping and information proximity to each associate control, whereas the “poor” design exhibited high color contrast and numerous intersecting process flow lines intent to confuse the order of operations. Findings from the Handal & Ikuma experiment failed to yield significant outcomes toward interface design directly relevant to performance, but did reveal much about the effect of interface design on SA and workload: SA scores were higher when using the “good” interface, likewise workload was lower with the “good” interface inferring the “good” design was better. The researchers did note an absence of correlation directly between their performance measures and the interface design types utilized, but inferred high SA was indicative of high performance despite the lack of support from statistical significance. All of the potentially confounding factors in the Handal & Ikuma study imply the heuristically developed design principles may not have been a major factor toward their research findings. Future research should attempt to negate the effect of heuristic development as much as possible to ensure the underlying design attributes are the true driver of SA and performance findings.

Task Management Strategies

Task switching is a broad topic discussed most frequently in the realm of the cognitive sciences (Hirst & Spelke, 1980; Brown, 1998; Monsell, 2003; Yeung, Nystrom, Aronson, & Cohen, 2006; Squire, Trafton, & Parasuraman, 2006). Research into task switching focuses heavily on the effects of forced task switching, whereby the researcher determines when a participant is allowed to work on a primary or secondary task. Notable here is that frequently both tasks are given equal weight and this information is communicated directly to the participant. The secondary effects from moving between tasks are then monitored and results correlated to performance outcomes. Because process control operations require system monitors to conduct activities in addition to a primary monitoring task there can be benefits to exploring the relationships between how operators manage multiple tasks and the effects task switching. Devising research efforts using the paradigm of an operator-driven task management strategy based on task switching would be a novel approach to investigating systematic problems in the process control industry. Specifically, how a human observer's preferred task management strategies interact with display design constructs and the impact these have on SA, performance, and workload outcomes are worthy of investigation.

Process Control Task Management

The manner in which individuals manage more than a single task varies. This is directly observable through the differences in which process control operators go about tackling their primary system monitoring and additional secondary tasks. Some operators are inclined to interleave between primary and secondary tasks in a serial fashion, devoting full attention to one task while disregarding another for extended periods of time. What triggers an individual to interleave can be attributed to either internal (e.g. unspecified, self-induced) or external (e.g. distractions, alerts) mechanisms, with self-induced, internal interruptions being the most problematic once the decision is made to transition back to the departed task (Duggan, Johnson, & Sorli, 2013). Still other process control operators appear to attempt to apply all of their cognitive resources to both tasks simultaneously, referred to as multitasking. Evidence of this behavior has been documented in research into equal task timing and dividing attention by Hirst & Spelke (1980) and Schumacher et al. (2001). And much of the literature in the field of cognition refers to multitasking as a term used to describe transitioning between more than one task expeditiously, so as to give the appearance of simultaneity. And as such, the simultaneous execution of multiple tasks means the actual term multitasking could be a bit of a misnomer as evidenced in the work of Dismukes, Loukopoulos, & Barshi (2009): an entire book entitled, "*The Multitasking Myth: Handling Complexity in Real-World Operations*".

The root problem with multitasking is that a consequence of attempts by an individual to multitask often result in degraded performance outcomes (Tombu & Jolicoeur, 2004). Thus, multitasking – whether or not it is truly a different activity for the purposes of task management from interleaving – could be thought of as rapid interleaving. Notable is that this time factor between the two is what makes either strategy recognizable to an outside observer when an individual is performing a dual task effort, because they execute the time on tasks in a distinctly different way. Toward investigation into process control display designs the manner in which an individual manages both a primary monitoring and secondary task can be cataloged as either interleaving or multitasking based on direct researcher observation. Support for making this determination will largely involve an operators time on each individual task.

Adaptation

Absent a unifying theory of cognitive control, researchers may be left to segregate the discrete nature of task management to simply interleaving or multitasking. Especially problematic is when multitasking is thought of as rapid interleaving. This leaves only a single task management strategy, which would mean task management is really one singular activity of infinitely varying degrees. In reality, the particular way an individual manages multiple tasks may point to the existence of a subgroup of one or the other. Direct observation of individuals engaged in multiple task efforts indicates there are indeed differences in how task load and task switching are handled. This is where the idea of adaptability in individuals as a task management strategy is put forth by Morgan, et al. (2013). Adaptability is a progressive approach to identifying how individuals cope with tasks, although Morgan et al. concede that adaptability may not actually be an absolute and different strategy mutually exclusive of multitasking. The idea of adaptation is a resultant outcome of an individual’s response to the quantity and complexity for a set of given tasks. Adaptive Attacking is described by Morgan, et al. as when an individual aggressively pursues a more difficult task in a multiple task scenario by diverting attention from another less demanding task. In direct opposition to this is the notion of Adaptive Avoidance wherein an individual has made a concerted effort to avoid the more complicated task altogether. The work of Morgan, et al. is in search of a balanced stratagem depicting the proper mix of tasks and management strategy for sustained performance. But the value of exploration into the individual differences with the strategies covered by Morgan, et al. is where future work may capitalize on task management tenets that can impact SA, task performance, and workload.

Need for Task Management Strategy Evaluation

Despite individual differences in how task load is managed, the process control industry and even user centered design principles have not taken the internal factor, the task management strategy, into consideration for an operator in a central control room. Nor have other research efforts involving process control attempted to correlate task management strategies to SA, performance, or workload outcomes. Heuristic based design principles have only covered the external factors of physical appearance and usability of the user interface external to the user and do not address the user's internal task management interaction with the final design construct. Based on a review of the literature four distinguishably different types of task management strategies have been identified for process control research consideration: Interleaving, Multitasking, Adaptive Attack, and Adaptive Avoidance. Definitions of each strategy for the purposes of this research effort are presented in Table II-2.

Strategy	Definition
Interleaving	Switching back and forth between a primary and secondary task, applying full attention to only one task at a time
Multitasking	Dividing and balancing attention equally between both a primary and secondary task
Adaptive Attack	Aggressively pursuing a secondary task in attempts to complete it as quickly as possible so as to devote full attention to a primary task when done
Adaptive Avoidance	Purposefully focusing attention on a primary task in efforts to disengage from a secondary task

Table II-2 Task Management Strategy Definitions

Conclusion

Endsley's definition of SA (1995) and SAGAT metric (2000) have relevancy to the evaluation of process control interfaces. Both can be applied to a multiple process monitoring, multiple task environment to research competing interface design constructs and attributes. The human machine interface display plays a key role in the maintenance of a process control room operators SA and is typically developed upon a host of underlying display attributes, making each individual one worthy of further exploration. Therefore, research into interface design development should seek to determine those underlying attributes that influence SA and task performance the most. However, evidence in the literature shows interface design evaluations have been conducted with a heuristic-based approach, rather than an empirical approach, to not only develop but also identify one design construct as better than another.

Previous works measuring SA using SAGAT have included the field of chemical process control to simulate a monitoring task. Results from these studies have shown significance relative to SA, by studying the effects of each design under differing task loads (Handal & Ikuma, 2012; Tharanathan, Bullemer, Laberge, Reising, & McLain, 2012). Unfortunately, this may be problematic, because high task load is not reflective of the process control – and many other industries – making it feasible the results from both works are not truly indicative of a real world process control activity. Most of an operator's time spent monitoring an automated system is spent performing routine monotonous supervisory control (Sheridan & Parasuraman, 2005), which falls under the paradigm of a vigilance monitoring task. Contrary to previous works the focus of this research effort is to maintain a relatively low task load to elicit a more appropriate evaluation of the display design characteristics on the user interface. Another research factor to take into account is how a process control operator manages concurrent tasks during a routine monitoring activity. This has not been considered before in the realm of a process control application to evaluate SA, task performance, and workload. But differing strategies have the potential to result in different outcomes and should be investigated for statistical relevancy.

Current research trends typically involve some form of heuristic development of a series of competing overall interface designs benchmarked by a formal study involving human subjects. These attempts to yield a "best" design are often met with mixed results, because they fail to focus on the underlying design attributes that apply to a more generalized audience. This research seeks to fill a gap by focusing on those underlying process control display attributes external to the human monitor and the preferred manner of task management internal to the human monitor to determine if either influences SA, task performance, or workload outcomes.

III. Evaluation of Human Machine Interface Design Factors on Situation Awareness and Task Performance

Abstract

In centralized process control facilities system performance likely hinges on effective interface design, because these interfaces are typically the only connection operators have with the systems they are managing. Decisions regarding interface design can be influenced by a variety of factors from user centered design principles to regulatory guidelines. While such guidance adds value to interface design, it does not reveal the underlying attributes that result in increased operator situation awareness and task performance. Current research focuses on design heuristics, neglecting empirical evaluations of interface design construct attributes. The purpose of this research effort was to explore the effects on situation awareness and task performance for four competing display design constructs: numeric versus graphic and functionally grouped versus spatially mapped. Findings show negligible differences amongst these design constructs for a conventional multi-process monitoring task. However, data trends toward graphic depictions arranged in a functionally grouped manner cannot be discounted as potentially being beneficial toward SA and task performance.

Introduction

Technological advances in the field of automation have produced a favorable return on investment for those industries and end users that have embraced the idea of integrating contemporary control methodologies into existing and newly devised process control applications. The human machine interface (HMI) provides a prime example.

HMI's are typically found where multiple remote processes are likely to be overseen and directly controlled from a single remote location outfitted with numerous HMI's or other visual interfaces and staffed by a small team or even a single individual. In effect, the automation evolution has implicitly displaced operators working side by side with the system under their direct control and migrated toward a central control room with total system oversight. In the central control room, a deluge of information about system status for multiple processes is passed in real time back to a small contingent of individuals responsible for maintaining systems health. The operator interface has played a key role in this evolutionary shift toward increased automation and reduced manpower. This is because a requisite part of the HMI's integration in a centralized control room has been to consolidate the comprehensive list of critical system data once observed by many into a concise, meaningful representation of the system for a smaller team.

The HMI's greatest challenge has been in aiding the human monitor to maintain an optimum level of situation awareness (SA) through a highly effective interface design construct, even though the operator is no longer interacting continuously with the system. More often than not, a single individual is charged to monitor multiple processes simultaneously. It is of paramount importance that the individual operator has a clear mental model and understanding of what is going on in the field with the system under his/her direct control.

Current methods for determining HMI design center on design heuristics, subjective best practices, and user-centered design principles. While addressing the needs of the interface design community, these practices fail to determine if competing design constructs differ in SA and performance outcomes. The purpose of this research is to conduct a controlled experiment to identify the impact on SA and task performance from two means of information presentation (numeric, graphic) at two different levels of arrangement (functionally grouped, spatially mapped) for a remotely monitored series of processes.

Background

Two recent industrial process control tragedies in the chemical and oil refinement field's, the Bayer CropScience pressure vessel explosion in Institute, WV (U.S. Chemical Safety and Hazard Investigation Board, 2011) and the BP oil refinery explosion in Texas City, TX (U.S. Chemical Safety and Hazard Investigation Board, 2007) resulted in the combined loss of 17 lives and inflicted injuries upon an additional 188 individuals. These figures are quite sobering and point to exactly how dangerous the process control industry can be when operator SA is incomplete. Maintaining appropriate operator SA is more difficult with remotely controlled equipment than when an operator is fully immersed within the affected environment. Therefore, remote operations present many challenges toward maintenance of operator SA including limitations to the operator's view of the system, latency of information presented at the remote location, and a limited depth or richness in context to the information provided through the user interface (Chen, Haas, & Barnes, 2007).

The ability to exact a positive influence over operator SA and task performance during a process control monitoring operation has been difficult (Mumaw, Roth, Vicente, & Burns, 2000; Sheridan & Parasuraman, 2005; Miller & Parasuraman, 2007; Li, Horberry, & Powell, 2010; Cummings, Bruni, & Mitchell, 2010). Despite many advances in automation technologies involving interface development, and several well thought out heuristic guidelines for display designs over the past 30 years (Norman, 1984; Gould, 1988; Vicente & Rasmussen, 1992; Ponsa, Vilanova, Perez, & Andonovski, 2010; Shneiderman, Plaisant, Cohen, & Jacobs, 2010; U.S. Nuclear Regulatory Commission, 2002), process control disasters like those mentioned persist, and still pose a danger to many different industries. Contributing to the reasons behind this are that the current body of interface design knowledge and research has not delved deeply enough into the process control industry's need for tangible evidence toward appropriate interface design constructs determined to improve operator SA. As evidenced in two recent studies into the effects of display design on SA, current trends continue to lean toward the formal evaluation of heuristic based designs (Handal & Ikuma, 2012; Tharanathan, Bullemer, Laberge, Reising, & McLain, 2012). Handal & Ikuma employed design constructs that were qualitatively defined as either "good" or "poor" with each revolving primarily around the use of color, contrast, and indicator proximity to a spatially mapped depiction of the system. In a similar manner Tharanathan, et al. focused on competing designs, but more along the lines of object layout: one being defined as "functional" and featuring grouped dynamic graphical indicators, the other as "schematic", having spatially mapped static indicators relative to their physical relevancy to the underlying process. Study findings in both efforts were mixed. Handal & Ikuma found no significant outcomes toward SA based on display type, whereas Tharanathan, et al. found higher levels 1 and 2 SA when using a "functional" type display. These findings are not without merit, but problematic for both research efforts is that it is difficult to determine the fundamental design attribute that contributed to the results most. Rather, it can only be inferred the combination of heuristic design principles each study selected to create their designs contributed – positively and/or negatively – to the overall outcomes.

Using Endsley's formal theory and definition of SA: "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*" (1995), this research departs from the norm of using a heuristic based design evaluation approach to explore the fundamental manner in which information is presented and how this influences operator SA and task performance.

Experiment

A case study for this experiment utilized the Air Force Research Laboratory, Component Research Air Facility (AFRL/CRAF) (Air Force Research Laboratory / Aerospace Systems Directorate, 2014). The Component Research Air Facility is an appropriate case example because of its large industrial complex layout and its use of a central control room to monitor multiple geographically dislocated processes. A task analysis performed at the facility offered numerous insights into existing display design usage *in situ* and led to the development of four competing experimental designs as shown in Figure III-1. The designs differed primarily in the means of information presentation, being either numeric or graphic representations of underlying process variables and arranged through either functional grouping by variable type or spatial mapping about a fictitious piece of equipment.

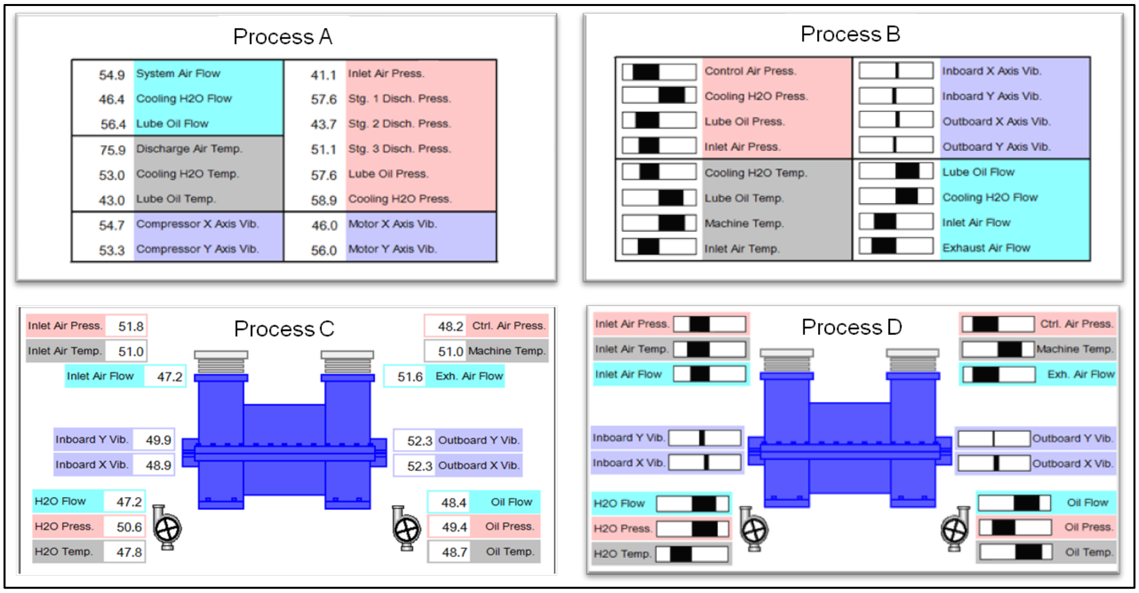


Figure III-1 Competing Display Designs
Processes A and C - Numeric; B and D - Graphic;
A and B - Functionally Grouped; C and D - Spatially Mapped

A human subjects research experiment was conducted at the Component Research Air Facility with the primary focus to determine if the manner of information presentation from each competing design influenced operator SA. Subordinate to this was the effect of the display design on operator performance of the primary monitoring and an additional secondary task. The experiment consisted of a 2x2 factorial design with information presentation (numeric, graphic) and information arrangement (functionally grouped, spatially mapped) serving as the two factors. The dependent variables were the three levels of SA: perception, comprehension, prediction as well as primary and secondary task performance outcomes. Participant SA was evaluated using the Situation Awareness Global Assessment Technique (SAGAT) (Endsley M. R., 2000). Pauses for SAGAT polling were preset at varying intervals across the four trials to give the appearance of true randomness to the participant. Performance measures were evaluated as a combination of correct SAGAT responses and system data relative to deviation acknowledgement time and prediction accuracy.

Twenty-four human subjects participated in the experiment, 19 male and 5 female ranging in age from 21-58 years (M = 42, SD = 13). Total contact time for the experiment was approximately 4 hours with each participant completing a 1-hour training and practice session followed by a series of 4, 30-minute trials for data collection. In addition to the training session, counterbalancing measures to preclude learning and order effects due to presentation order were managed by use of a Latin Square Design. During each trial a series of 8 simulated processes using one of the design constructs were manipulated directly by the participant via a standard computer mouse. Each participant was charged to predict and react to deviations that occurred on the displayed processes every 2-minutes on average. These deviations occurred most frequently on the processes exhibiting erratic operational characteristics. In parallel with the primary monitoring task, the participant also executed a secondary reading comprehension test, designed to mimic the cognitive load of reading and responding to written communications (e.g. email). The test setup with a participant carrying out both primary and secondary tasks is shown in Figure III-2.

Figure III-2 Experimental Test Setup

Analysis and Results

Situation Awareness

SAGAT queries administered during trial pauses were balanced across Levels 1 and 2 SA, each having 6 questions asked for the duration of each trial. Level 3 SAGAT injects were also included to ensure the SAGAT polls covered all three levels of SA and to keep participants from anticipating questions relevant to only perception and comprehension. Level 3 SAGAT injects were not used in the final assessment to eliminate any possible confounding of the data due to participant guessing. Rather, to negate these effects, Level 3 SA was assessed based on the participant's ability to accurately predict which processes would deviate next, which was part of the primary task. Predictions made for one of the two worst running processes out of eight were considered successful predictions. Because the simulation was dynamic with process variables traversing normal and erratic states at different points in time this measure was less likely to result in false positives than the Level 3 SAGAT injects.

No statistically significant correlations were found between SA outcomes for the four competing display designs relative to each other. All instances produced $p > .05$. However, ANOVA results at 95% CI did show mild trends toward the positive effects of graphic information presentation for levels 1 ($F_{3,92} = 0.75$, $p > .05$), 3 ($F_{3,92} = 1.90$, $p > .05$), and overall ($F_{3,92} = 0.90$, $p > .05$) SA and a functional grouping orientation for level 2 ($F_{3,92} = 0.63$, $p > .05$) SA. These results are summarized in Table III-1. Given the lack of statistical significance, they are considered notional and not meant to infer a true difference in each design. The data reflect the manner of combined information presentation has resulted in negligible differences affecting SA.

	Display Type	Mean	Std Dev	95% CI	
Level 1 SA %	**Graphic, Grouped**	**79.51**	**15.73**	**72.87**	**86.16**
	Graphic, Spatial	**79.03**	**15.94**	**72.3**	**85.76**
	Numeric, Grouped	77.43	18.79	69.50	85.37
	Numeric, Spatial	73.33	12.39	68.10	78.57
	p = .525				
Level 2 SA %	**Graphic, Grouped**	**71.18**	**15.49**	**64.31**	**78.05**
	Graphic, Spatial	67.88	17.05	61.01	74.75
	Numeric, Grouped	**73.61**	**17.53**	**66.74**	**80.48**
	Numeric, Spatial	68.06	17.62	61.19	74.92
	p = .598				
Level 3 SA %	**Graphic, Grouped**	**88.33**	**12.74**	**80.42**	**96.24**
	Graphic, Spatial	**90.00**	**12.16**	**82.09**	**97.91**
	Numeric, Grouped	77.92	25.19	70.01	85.83
	Numeric, Spatial	82.92	24.04	75.01	90.83
	p = .135				
Overall SA %	**Graphic, Grouped**	**79.68**	**11.07**	**74.88**	**84.48**
	Graphic, Spatial	**78.97**	**8.09**	**74.17**	**83.77**
	Numeric, Grouped	76.32	14.99	71.52	81.12
	Numeric, Spatial	74.77	12.17	69.97	79.57
	p = .445				

Table III-1 Summary Results for Display Type Effect on SA.
(Highlighted areas reflect trends toward construct resulting in higher mean SA)

Using individual standard deviations to calculate the intervals and 95% confidence interval bars, ANOVA results are shown visually in Figure III-3 through Figure III-6. Each graph reveals a high degree of variability in participant responses across all of the experimental display types, with the only exception being level 3 SA (prediction accuracy) where graphic designs show variability roughly half that of numeric (Figure III-5). Overall, these results fail to reject the null hypothesis of no difference between display design types.

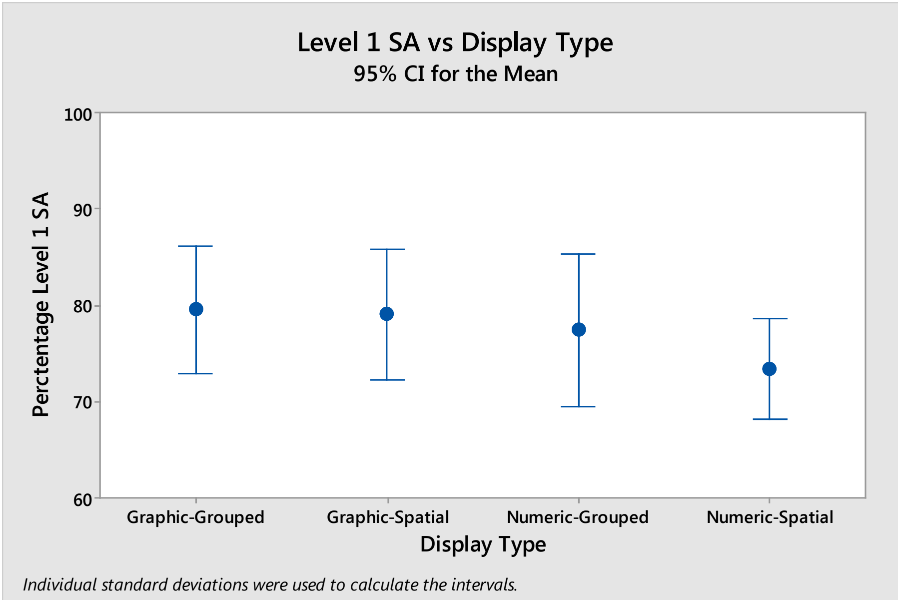


Figure III-3 Level 1 SA versus Display Type
Results indicate percentage of correct responses to Level 1 SAGAT queries

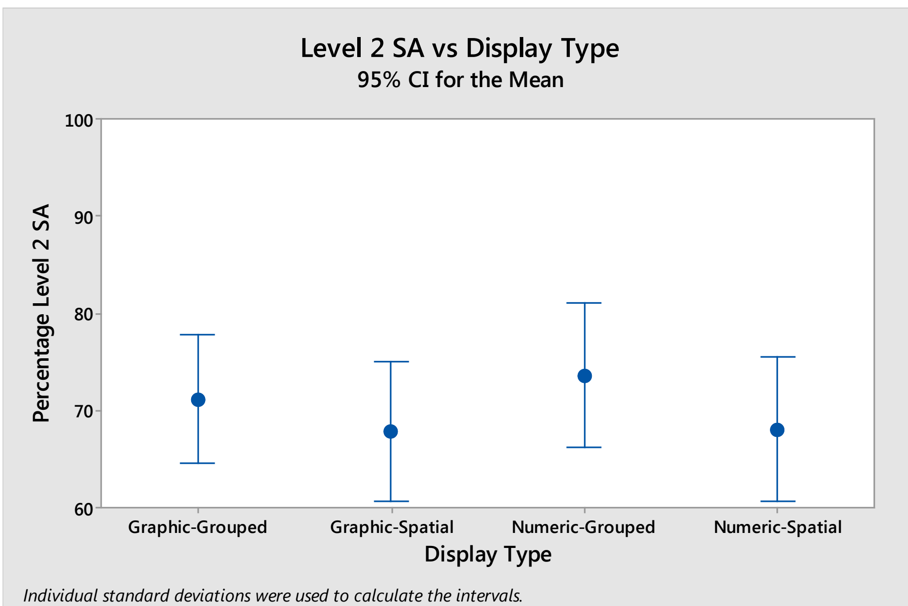


Figure III-4 Level 2 SA versus Display Type
Results indicate percentage of correct responses to level 2 SAGAT queries

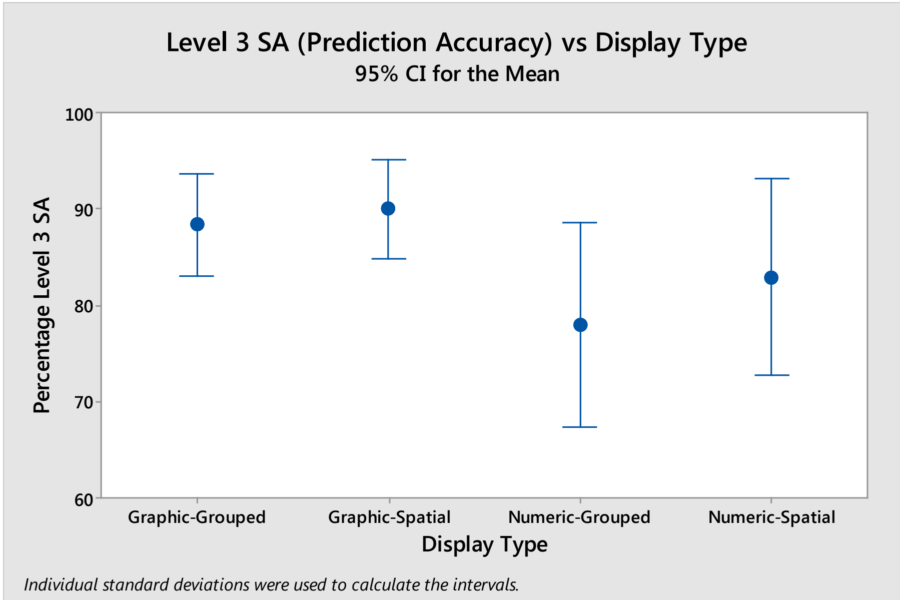


Figure III-5 Level 3 SA versus Display Type
Results indicate percentage of predictions where one of the two most erratic processes was predicted to deviate

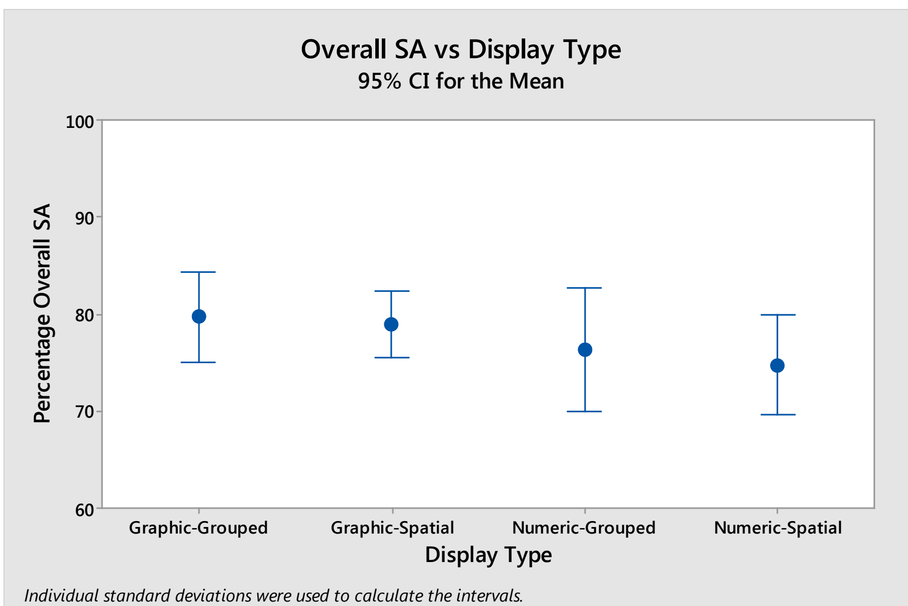


Figure III-6 Overall SA versus Display Type
Results indicate the aggregate combination of levels 1, 2, and 3 SA

Table III-1 shows results of the display constructs as a combination of information presentation and orientation. Breaking each display's construct down even further into individual attributes (graphic and numeric), an examination for effect on SA again found no significance, with one exception: graphic information presentation had a statistically significant advantage over numeric presentation ($F_{1,94} = 4.88$, $p = .030$) for Level 3 SA only – measured as the participant's ability to predict where deviations would occur. Individual attribute data is shown in Table III-2. In sum, for the underlying display attributes, no other direct inferences of SA advantage for one over another could be made.

	Display Attribute	Mean	Std Dev	95% CI	
Level 1 SA %	Graphic	79.27	15.67	74.75	83.79
	Numeric	75.38	15.88	70.86	79.90
	p = .230				
	Grouped	78.47	17.18	73.93	83.02
	Spatial	76.18	14.42	71.64	80.72
	p = .481				
Level 2 SA %	Graphic	69.53	16.20	64.68	74.38
	Numeric	70.83	17.61	65.98	75.68
	p = .707				
	Grouped	72.40	16.41	67.59	77.21
	Spatial	67.97	17.15	63.16	72.78
	p = .200				
Level 3 SA %	**Graphic**	**89.17**	**12.35**	**83.61**	**94.72**
	Numeric	80.42	24.49	74.86	85.97
	p = .030				
	Grouped	83.13	20.44	77.44	88.81
	Spatial	86.46	19.18	80.78	92.14
	p = .412				
Overall SA %	Graphic	79.32	9.59	75.96	82.68
	Numeric	75.54	13.53	72.18	78.90
	p = .118				
	Grouped	78.00	13.14	74.60	81.40
	Spatial	76.87	10.44	73.47	80.27
	p = .642				

Table III-2 Analysis of Display Attribute Effect on SA
Results indicate the only statistically significant finding (highlighted in bold) to be graphic displays resulted in higher level 3 SA, prediction accuracy

Performance

Time based performance metrics for the primary monitoring task were broke out into two components of the participant's deviation management capabilities: prediction and response times. Deviation prediction times were measured as the amount of time in seconds it took a participant to assess all of the displayed processes and make a prediction about where the next deviation would occur. Deviation response times represented how long it took to perceive, then acknowledge a deviation on a display using a computer mouse. ANOVA results for both metrics did not show statistical significance between the four competing display types. Results are depicted visually in Figure III-7 and Figure III-8. Nor did any of the results show significance when individual display attributes were examined either, therefore no summary data is shown for those.

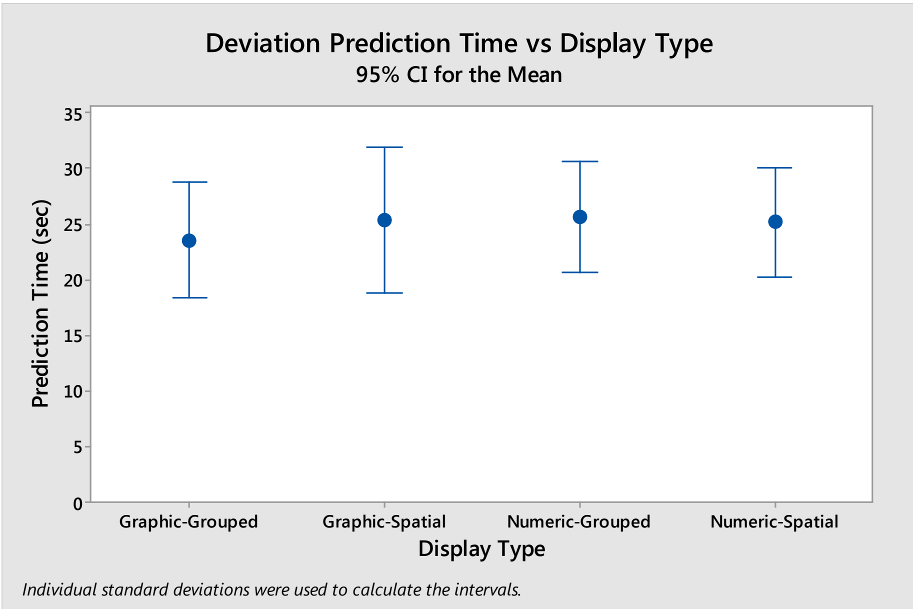


Figure III-7 Deviation Prediction Time versus Display Type
Results indicate the average time in seconds a participant was able to assess all monitored processes then make a prediction. A faster prediction time indicates better performance

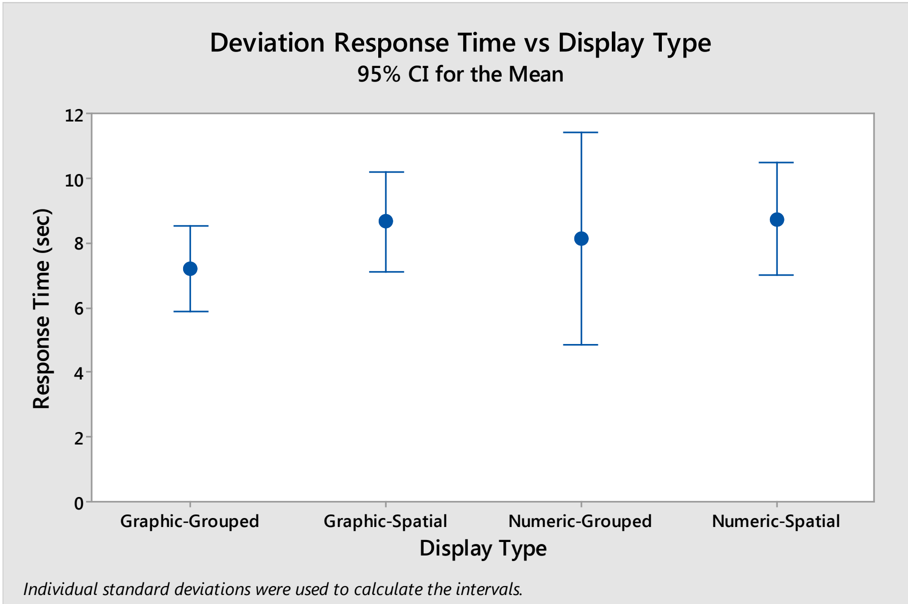


Figure III-8 Deviation Response Time versus Display Type
Graph shows the average time factor in seconds it took a participant to perceive then acknowledge a deviation on the display. A faster response time indicates better performance

Discussion and Conclusion

Outcomes of this research effort produced SA results that ran counter to the Handal & Ikuma (2012) and Tharanathan, et al. (2012) findings. The manner of information presentation as numeric or graphic, functionally grouped, or spatially mapped was found to be largely inconsequential with the exception of graphical presentation and the ability of an individual to predict future deviations (level 3 SA). Perhaps using an alternative measure of SA in lieu of SAGAT would have produced different results, but this is left for future endeavors to explore. This study differed from previous works primarily by focusing on individual display design attributes as opposed to carrying out an evaluation of a heuristically motivated design. Another difference was with task load, which remained consistent for this study, but was alternated between low, medium, and high in efforts to show differences between the competing designs in the other studies. Because most process control operations involve a high degree of vigilance monitoring over long periods of time with virtually no interaction between human and machine, this study focused on low task load to be consistent with the intended environmental context. While task load remained relatively low, each trial was relatively short (20 min). Future work could extend this evaluation period to determine the impacts of display designs on SA and performance in a vigilance setting. Future works should also explore the myriad of other display attributes within the visual spectrum (e.g. effects of color, global alarm/alert indication, flashing indication) and beyond. The effects of audible context could add yet another dimension to future efforts through the use of audible cueing or multi-dimensional sound directing an observer to a particular display or area within a display. Given the high variability across participant responses observed in this study there may be other factors driving the outcomes of this and previous research that should be investigated. Future works should seek to determine how factors such as operator task management strategy and individual personality characteristics potentially impact outcomes toward SA and task performance.

Summing up the results, findings were the fundamental design attributes and manner of information presentation play little to no role in influencing SA and task performance in a process control environment. It may be beneficial based upon these findings to consider the alternative of allowing maximum user preference for process monitoring tasks, making rigid, heuristically developed constructs a thing of the past.

Acknowledgements

This work was supported in part by the Air Force Research Laboratory's (AFRL) Aerospace Systems Directorate (RQ) and its Component Research Air Facility (CRAF). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of AFRL, RQ, or the U.S. Government.

References

Air Force Research Laboratory / Aerospace Systems Directorate. (2014). *Facility Factsheet Component Research Air Facility (CRAF).* (AFRL/RQOEE, Ed.) Retrieved January 1, 2014, from www.wpafb..af.mil: http://www.wpafb.af.mil/shared/media/document/AFD-130410-048.pdf

Chen, J. Y., Haas, E. C., & Barnes, M. J. (2007, November). Human Performance Issues and User Interface Design for Teleoperatored Robots. *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews, 37* (6), pp. 1231-1245.

Cummings, M. L., Bruni, S., & Mitchell, P. J. (2010). Human supervisory control challenges in network-centric operations. *Reviews of Human Factors and Ergonomics, 6* (1), 34-78.

Endsley, M. R. (2000). Direct Measurement of Situation Awareness: Validity and Use of SAGAT. In M. R. Endsley, & D. J. Garland (Eds.), *Situation Awareness Analysis and Measurement.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37* (1), 32-64.

Gould, J. (1988). How to design usable systems (excerpt). IBM Research Center Hawthorne. In M. Helander (Ed.), *Handbook of Human-Computer Interaction* (pp. 757-789). Yorktown Heights, NY, 10598, 757-789.

Handal, C., & Ikuma, L. H. (2012). Good Interface Design Improves Situation Awareness in Control Room Operators. In G. Lim, & J. W. Herrman (Ed.), *Proceedings of the 2012 Industrial and Systems Engineering Conference.* Orlando.

Li, X., Horberry, T., & Powell, M. (2010). Human Control in Mineral Processing Plants: An Operator Centered Investigation. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 284-288). Human Factors and Ergonomics Society.

Miller, C. A., & Parasuraman, R. (2007). Designing for Flexible Interaction Between Humans and Automation: Delegation Interfaces for Supervisory Control. *The Journal of the Human Factors and Ergonomics Society, 49* (1), 57-75.

Mumaw, R. J., Roth, E. M., Vicente, K. J., & Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 42* (1), 36-55.

Norman, D. A. (1984). *Cognitive engineering principles in the design of human-computer interfaces.* Human Computer Interaction Amsterdam: Elsevier Science.

Ponsa, P., Vilanova, R., Perez, A., & Andonovski, B. (2010). SCADA Design in Automation Systems. *3rd Conference on Human System Interactions (HSI)* (pp. 695-700). IEEE.

Sheridan, T. B., & Parasuraman, R. (2005). Human-Automation Interaction. *Reviews of Human Factors and Ergonomics, 1* (1), 89-129.

Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. (2010). *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Reading, MA: Addison-Wesley Publishing Co.

Tharanathan, A., Bullemer, P., Laberge, J., Reising, D. V., & McLain, R. (2012). Impact of Functional versus Schematic Overview Displays on Console Operators' Situation Awareness. *Journal of Cognitivce Engineering and Decision Making, 6* (2), 141-164.

U.S. Chemical Safety and Hazard Investigation Board. (2011). *Pesticide Chemical Runaway Reaction - Pressure Vessel Explosion.* Investigation Report, No. 2008-08-I-WV, Bayer CropScience, LP, Institute, WV, August 28, 2008.

U.S. Chemical Safety and Hazard Investigation Board. (2007). *Refinery Explosion and Fire.* Investigation Report, No. 2005-04-I-TX, British Petrolium (BP), Texas City, TX, March 23, 2005.

U.S. Nuclear Regulatory Commission. (2002). *Human-system interface design review guidelines.* Brookhaven National Laboratory, Energy Science & Technology Department.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface Design: Theoretical Foundations. *IEEE Transactions on Systems, Man, and Cybernetics, 22* (4), 589-606.

IV. Influences of Task Management Strategy on Situation Awareness, Performance, and Workload in a Process Control Environment

Abstract

Objective: The purpose of this study was to identify individual task management strategies utilized by process control operators to determine the effect each had on situation awareness (SA), performance, and workload outcomes.

Background: Process control operations have suffered catastrophic failures when operator awareness of the underlying system was incomplete. The individual differences between operator task management strategies utilized in a multiple task environment – as a possible contributing factor – has not been heavily researched.

Method: A case study of an operational facility led to the development of a fully interactive process control simulator, whereby participants were charged to monitor 8 processes and simultaneously execute a demanding secondary task. Task load remained consistent across four separate trials utilizing differing interface design schemas. Direct researcher observation and self-report metrics were used to identify how tasks were managed during each trail. Measures of SA, performance, and workload followed the Situation Awareness Global Assessment Technique (SAGAT), time based responses by each participant, and the NASA-TLX, respectively.

Results: Four competing strategies were identified -- Interleaving, Multitasking, Adaptive Attack, Adaptive Avoidance -- and showed significance towards several factor responses. Adaptive Attack and Multitasking strategies demonstrated more advantageous outcomes toward SA maintenance and performance while also resulting in lower participant experienced workload. Interleaving revealed marginally higher SA and performance results, but high workload. Adaptive Avoidance resulted in the worst outcomes for all cases with the exception of experienced workload measures of effort.

Conclusion: Task management strategies influenced SA, performance, and workload in a process control environment. Individual differences in task management had both positive and negative ramifications toward system oversight.

Application: Identification of task management strategies relevant to process control can lead to identification of personnel characteristics appropriate for process monitoring tasks and aid in the development of training methods for more effective control.

Keywords: task management, task switching, task performance, process control, process monitoring, situation awareness, workload, human computer interaction, human automation interaction

Introduction

Background

Process control operations of today are demanding environments that require facility personnel to function in a multiple process, multi-task atmosphere. These personnel are routinely monitoring more than one system or series of systems as their primary task. Simultaneously, they also perform a host of unrelated secondary tasks associated with routine facility operations. This managing of competing primary and secondary tasks is problematic, because it is difficult to maintain appropriate operator situation awareness (SA) when switching between tasks. Making this even more difficult is with the operation of remote equipment where an operator is located in a central control room away from the context of the machinery they oversee. Two unfortunate and tragic examples of this problem are in the chemical and oil refinement fields: the Bayer CropScience pressure vessel explosion in Institute, WV (U.S. Chemical Safety and Hazard Investigation Board, 2011) and the BP oil refinery explosion in Texas City, TX (U.S. Chemical Safety and Hazard Investigation Board, 2007). Just these two disasters resulted in the combined loss of 17 lives and inflicted injuries upon an additional 188 individuals. Averting these type of loss of life and injury statistics is not just a top priority for Bayer and BP, but the entire process control industry as a whole.

While remote operations have presented many challenges toward maintenance of operator SA, much of the human factors debate has sought to address the problem of maintaining appropriate operator SA with remotely operated systems through the investigation of external factors outside of the human monitor. Examples of this include: Chen, Haas, and Barnes (2007) who identified limitations to the operator's view of the system, latency of information presented at the remote location, and a limited depth or richness in context to the information provided through the user interface as impediments to effective process control; Pantelli et al.'s (2013) suggestion the main sources of SA degradation for power system control centers lies within six factors, only one of which was identified as pertinent to the individual alone, and this only relative to operator training. Other works focus primarily on the means of developing heuristic guidance to combat process control failures through an improved operator interaction experience (Norman, 1984; Gould, 1988; Vicente & Rasmussen, 1992; Ponsa, Vilanova, Perez, & Andonovski, 2010; Shneiderman, Plaisant, Cohen, & Jacobs, 2010; U.S. Nuclear

Regulatory Commission, 2002). Despite these efforts to improve external factors process control disasters like the Bayer and BP incidents persist, and pose a significant danger spanning many different industries.

Identification of external elements and heuristic design improvements have provided a solid step toward maximizing the return on tackling factors inhibiting process control, yet identification and investigation of internal factors specific to the human monitor has remained largely unaddressed. This research espoused the identification of four individual preferred task management strategies internal to the human operator functioning in a multiple task environment. The identification and development of these strategies was based on the cognitive sciences idea of task switching and is further detailed in the following sections.

Task Management

Task switching is a broad topic discussed most frequently in the realm of the cognitive sciences(Hirst & Spelke, 1980; Brown, 1998; Monsell, 2003; Yeung, Nystrom, Aronson, & Cohen, 2006; Squire, Trafton, & Parasuraman, 2006). Research into task switching has focused much on the effects of forced task switching, whereby an experimental design is configured to predetermine when a participant is permitted to work on either a primary or secondary task. The secondary effects of moving between tasks are then monitored and results correlated to performance outcomes. Because process control operations require system monitors to conduct activities in addition to a primary monitoring task at their own discretion, there exist parallel benefits to exploration of the relationships between an operator’s chosen method to manage multiple tasks – defined here as an operator's task management strategy – and subsequent outcomes based upon the foundation of task switching principles.

Interleaving and Multitasking

The manner in which individuals manage more than a single task varies. This is directly observable through the differences in which process control operators go about tackling their primary system monitoring and additional secondary tasks. Some operators are inclined to interleave between a primary and secondary task in a serial fashion, devoting full attention to one task while disregarding another for extended periods of time. What triggers an individual to interleave can be attributed to either internal or external mechanisms, with self-induced interruptions being the most problematic once the decision is made to transition back to the departed task according to research by Duggan, Johnson, & Sorli (2013). They found a degradation in performance associated with internal decisions to interleave versus external triggers. Still other process control operators appear to attempt to apply all of their cognitive resources to both tasks simultaneously, referred to as multitasking. Evidence of this behavior has been documented in research into equal task timing and dividing attention by Hirst & Spelke (1980) and Schumacher et al. (2001). Much of the field of cognition refers to multitasking as a term used to describe transitioning between more than one task expeditiously, so as to give the appearance of simultaneity. And as such, the simultaneous execution of multiple tasks means the actual term multitasking could be a bit of a misnomer as evidenced in the work of Dismukes, Loukopoulos, & Barshi(2009), an entire book entitled, "*The Multitasking Myth: Handling Complexity in Real-World*

Operations." Most problematic with multitasking is that a consequence of attempts by individuals to engage in it often exhausts their mental resources and results in degraded performance (Tombu & Jolicoeur, 2004). Thus, multitasking – whether or not it is truly a different activity from interleaving – could be thought of as rapid interleaving. Fortuitously, this makes it recognizable to an outside observer, since an individual performing a dual task effort will spend distinctly different amounts of time on tasks before switching from one to another. Toward investigation into process control display designs, the manner in which an individual manages both a primary monitoring and secondary task can be cataloged as either interleaving or multitasking based on researcher observation compared to participant self-reported behavior.

Adaptation

Absent a unifying theory of cognitive control, researchers are left to segregate the discrete nature of task management to either interleaving or multitasking. Especially problematic is when multitasking is thought of as rapid interleaving. This leaves only a single task management strategy, which would mean task management is really one singular activity of infinitely varying degrees. In reality, the particular way an individual manages multiple tasks may point to the existence of a subgroup of one or the other. Direct observation of individuals engaged in multiple task efforts indicates there are indeed differences in how task load and task switching are internally managed. This is where the idea of adaptability in individuals as a task management strategy is put forth by Morgan, et al. (2013). The proposal of adaptability is an original approach to how individuals cope with tasks, although Morgan et al. concede it may not actually be an absolute and different strategy from multitasking. However, the overarching idea of adaptation is a resultant outcome of an individual's response to the quantity and complexity for a set of given tasks according to Morgan, et al. They put forth two different types of adaptation: Adaptive Attacking is when an individual aggressively pursues a more difficult task in a multiple task scenario by diverting attention from another less demanding task. In direct opposition to this is the notion of Adaptive Avoidance wherein an individual has made a concerted effort to avoid the more complicated task altogether. Although the work of Morgan, et al. was in search of a balanced stratagem depicting the proper mix of tasks and management strategy for sustained performance, this research examined the impact of utilizing either Adaptive Attack or Adaptive Avoidance strategies to identify the relative impact each had on SA, task performance, and workload.

Task Management Strategies Defined

Despite individual differences in how task load is managed, the process control industry has not taken task management strategies fully into consideration for an operator in a centralized process control environment. Nor have other research efforts outside of process control attempted to correlate task management strategies to SA, performance, and workload outcomes. Heuristic based design principles have only sought to address the factors external to the human inhibiting effective process control. Therefore, solely for the purpose of exploring internal factors inhibiting operator SA, four distinguishably different task management strategies were identified and defined to satisfy the objectives of this research initiative: Interleaving, Multitasking, Adaptive Attack, and Adaptive Avoidance. Definitions of each strategy are presented in Table IV-1.

Strategy	Definition
Interleaving	Switching back and forth between a primary and secondary task, applying full attention to only one task at a time
Multitasking	Dividing and balancing attention equally between both a primary and secondary task
Adaptive Attack	Aggressively pursuing a secondary task in attempts to complete it as quickly as possible so as to devote full attention to a primary task when done
Adaptive Avoidance	Purposefully focusing attention on a primary task in efforts to disengage from a secondary task

Table IV-1 Task Management Strategy Definitions

Research Focus

Devising a research effort using the paradigm of an operator-driven task management strategy based on task switching took a novel approach to investigating a systemic problem in the process control industry. Using Endsley's (2000) SAGAT method, time based response metrics, and the NASA-TLX (Hart & Staveland, 1988) this effort departed from the norm of investigating factors external to the human monitor in a process control environment. Identification of operator task management strategies and how they influenced SA, task performance, and workload was the focus of this research initiative.

Method

Case Study

A case study was carried out at the Air Force Research Laboratory's, Component Research Air Facility (AFRL/CRAF) (Air Force Research Laboratory / Aerospace

Systems Directorate, 2014). The Component Research Air Facility provided an appropriate case example because of its large industrial complex layout and its use of a central control room to monitor multiple geographically dislocated processes. A task analysis was also performed at the facility and offered numerous insights into existing process control operations, to include operator task management behavior *in situ.* This led to the development of four competing experimental display designs as shown in Figure IV-1. The designs differed primarily in the means of information presentation, being either numeric or graphic representations of underlying process variables and arranged through either functional grouping by variable type or spatial mapping about a fictitious piece of equipment. Formal evaluation of the competing designs was conducted by Bowden & Rusnock(2014). For the purposes of this research effort, the competing designs were used to vary the participant's simulated environmental context between trials. This not only provided greater opportunity to evaluate participant reactions to differing process control displays, but also meant the experiment proceeded with mitigating effects in place to ensure one particular design could not have potentially confounded subsequent findings of significance.

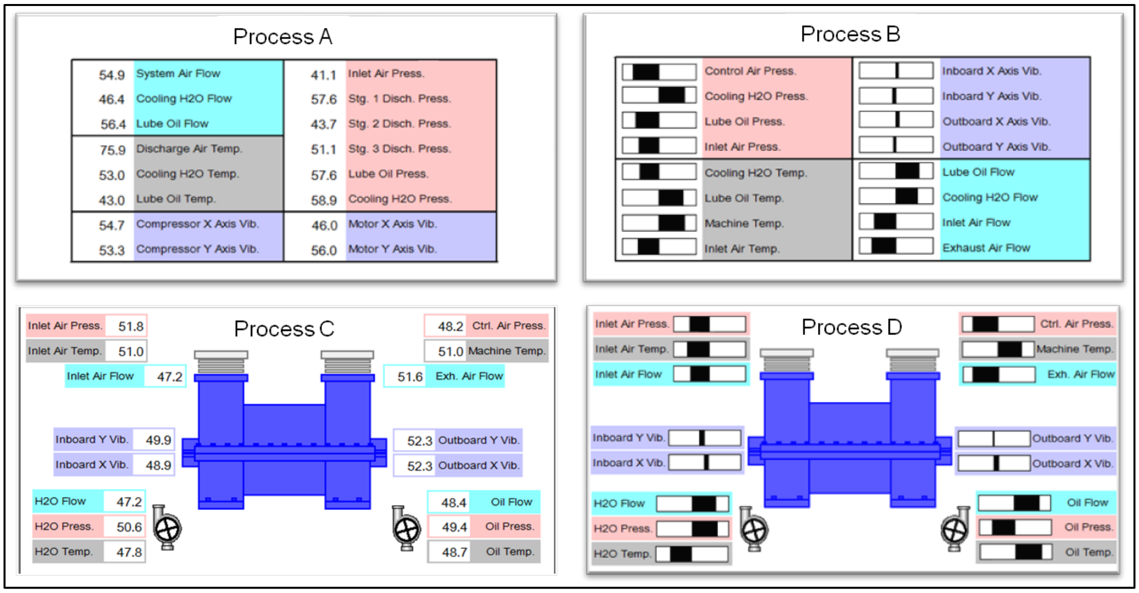


Figure IV-1 Trial Display Designs
Processes A and C - Numeric; B and D - Graphic;
A and B - Functionally Grouped; C and D - Spatially Mapped

Participants

Twenty-four participants, both military and civilian from the Air Force Research Laboratory and Air Force Institute of Technology volunteered for the experiment. Age ranged from 21-58 with a mean of 42 (SD = 13) and consisted of 19 males, 5 females. All participants had normal or corrected to normal vision with only one exhibiting color blindness per Ishihara's (2012) Colour Deficiency standard. Previous process control experience across participants was not required to participate in the study and resulted in seventeen participants having no experience with process control at all, three participants with between 2-6 years of experience, and four participants having more than 15 years of experience.

Apparatus and Equipment

Two Panasonic TH-42PF20U, 42-inch class high definition plasma displays featuring a 16:9 aspect ratio and 1920x1080 pixel resolution were mounted in a modular work center inside a soundproof room at the Component Research Air Facility. Two standard Lenovo ThinkCenter M77 workstations each with 3.2GHz AMD Athlon II B26 CPU, 4 GB of RAM, and one outfitted with a Sapphire AMD Radeon R7 240 GPU graphics card were used. Only one of the workstations had a standard keyboard and computer mouse connected, because the second Lenovo unit was only present as a redundancy measure should the first fail. The active workstation served 8 simulated processes to both displays using an Iconics Genesis GraphWorx32 software application. The Lenovo workstation also automatically captured NASA-TLX data at the end of each trial. Figure IV-2 shows a participant executing the primary and secondary tasks at the experimental work center.

Figure IV-2 Experimental Test Setup

Housed outside of the immediate experimental work center was a Schneider Modicon 140CPU43412A, Quantum series industrial programmable logic controller (PLC) with 80486 processor and math coprocessor clocked at 66MHz. The PLC ran the entire simulation and was programmed using the IEC 61131-3 function block diagram (FBD) programming language in Modicon's Concept V2.6 application software. The PLC controlled the trial master timer, start/stop, and all SAGAT pauses. It also stored event triggered data in internal registers until the end of each session. Event data was passed between the Lenovo workstation to the PLC by way of an OPC sever running KepServer software.

Procedure

The full experiment procedural checklist as well as all other experiment documentation can be found in Appendix D. Informed consent was discussed in detail and obtained from each participant prior to their inclusion in the study. Then a pre-experiment questionnaire was administered to capture participant demographic information. Total contact time for the duration of the experiment with each participant was approximately 4-hours to complete the training / practice and data collection. The 1-hour training / practice session consisted of a PowerPoint presentation and four 2-minute trials involving just the primary monitoring task with each interface type. At the end of the four practice trials, one 5-minute trial consisting of both the primary and secondary tasks together was executed and followed immediately by a NASA-TLX. The training presentation was administered by the researcher and the practice sessions were facilitated to answer any questions and familiarize each participant with trial flow. Additional training was offered, but not requested or deemed necessary for any of the participants. Upon completion of the training / practice session a series of four 20-minute trials for the data collection session were completed in succession. Each of the four trials ran approximately 30-minutes, including SAGAT pauses and NASA-TLX administration.

The experiment consisted of a 2x2 factorial design with information presentation (numeric, graphic) and information arrangement (functionally grouped, spatially mapped) comprising the different methods of presentation. Counterbalancing measures to preclude learning and order effects due to presentation order during the data collection session were managed by use of a Latin Square Design. A short break (5-10 minutes) was taken between the training and data collection sessions, with each participant executed all four trials in a single sitting over a 2-hour period. The remainder of the total contact time outside either session was consumed by the participant completing the pre- and post-experimental questionnaires as well as answering any questions by the researcher the participant had about the experimental design and their participation.

During each trial a series of 8 simulated processes was manipulated directly by the participant via the computer mouse. Each participant was charged to predict and react to deviations that occurred on the displayed processes, appearing every 2-minutes on average. Deviations were preprogrammed to occur most frequently on the processes exhibiting the worst operational characteristics. In parallel with the primary monitoring task, the participant also executed a secondary reading comprehension test, designed to mimic the cognitive load experienced by a process control room operator (e.g. reading and responding to written communications such as email).

Data Analysis

Dependent variables were identified as the three levels of SA (perception, comprehension, and prediction) per Endsley's formal definition (1995), primary and secondary task performance, and subjective workload. Independent variables consisted of the four competing task management strategies: Interleaving, Multitasking, Adaptive Attack, Adaptive Avoidance. Subject variables, age and experience were also targeted for further investigation into the effects of each on performance outcomes.

SAGAT queries administered during trial pauses were balanced across Levels 1 and 2 SA, each having six questions asked for the duration of each trial. Level 3 SAGAT injects were also included to ensure the SAGAT polls covered all three levels of SA and to keep participants from anticipating questions relevant to only perception and comprehension. During each trial pause the researcher handed the participant a parcel of paper containing the SAGAT questions. This was done to mitigate variance in the manner of question presentation across all participants. A sample of the questions asked during each SAGAT pause is provided in Figure IV-3. All SAGAT questions broken out by trial are contained in Appendix D.

Specifically which process number(s) experienced a deviation since the last pause? N 1 2 3 4 5 6 7 8
How many deviations total have there been since last pause? _____
Will you have the reading comprehension task completed by the end of the trial? Yes No
Indicate a process number that has **never** experienced a deviation (List no more than 2): _____ , _____

Figure IV-3 Sample SAGAT Questions

Level 3 SAGAT injects were not used in the final scoring assessment to eliminate any possible confounding of the data due to participant guessing. Rather, to negate these effects, Level 3 SA was assessed based on the participant's ability to accurately predict which processes would deviate next, which was part of the primary task. The method of prediction as seen in Figure IV-4 was when the participant selected the "Predict" text on the process they thought would deviate next. Once the prediction option was selected the ability to predict any other process went away and the "Predict" text on the selected process was highlighted in blue. Prediction capability returned after a deviation occurred.

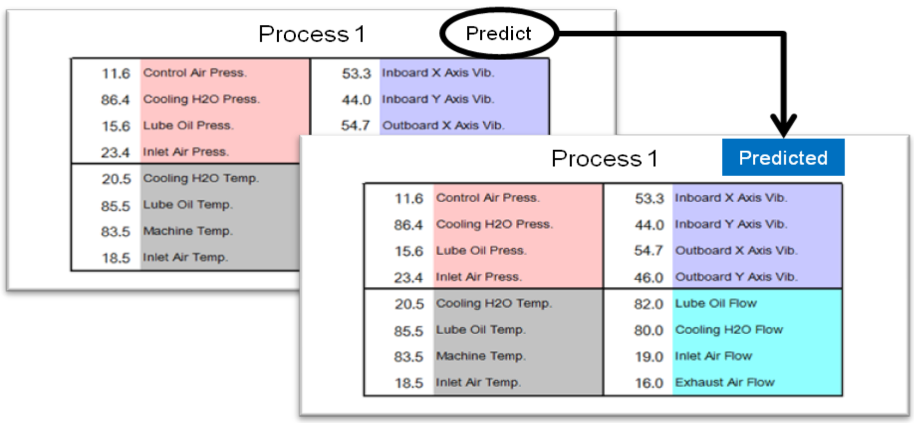


Figure IV-4 Experiment Simulator Prediction Functionality

Predictions made for one of the two worst running processes out of eight were considered successful predictions. Because the simulation was dynamic with process variables traversing normal and erratic states at different points in time, this measure was less likely to result in false positives than the Level 3 SAGAT injects.

Guidelines from SAGAT were given full consideration in the development of the experimental design (e.g. no pauses earlier than 3-5 min into a trial; no two pauses within 1-minute of each other). Pauses for SAGAT polling timers were generated using a random number generator (RANDOM.ORG) based on the constraints of the SAGAT method, then preset at varying intervals across the four trials to give the appearance of true randomness to the participant. SAGAT pause times broken out by trial are shown in Table IV-2.

Trial:	1	2	3	4
Predefined SAGAT Pauses (in minutes)	7	6	4	8
	12	9	10	13
	16	15	18	16
	20	20	20	20

Table IV-2 SAGAT Pause Times Per Trial (in minutes)

SAGAT responses were tallied for all three levels and rated by percentage of correct responses given at each of the three levels.

Performance measures were evaluated as to primary and secondary task scores associated with participant responses and interaction with the simulation. The primary task utilized a time based scoring mechanism that penalized participants the longer a deviation remained without acknowledgement, having a linear rate of decay that iterated point losses every 5-seconds; 20 seconds or longer resulting in zero points. Likewise, the secondary task utilized standard SAT scoring to penalize for incorrect answers, yet result in zero point value for questions left unanswered. Trial Scoring was an aggregate of the primary and secondary tasks having an 80/20 split, 800 points total for the primary task and 200 points total for the secondary. In both cases, primary and secondary task results are reported as percentages to ease interpretation without having to remember actual point values associated with actual scoring values. Performance was also measured relative to the simulator's trial timer in the form of deviation acknowledgement (response time) and the amount of time it took the participant to make a prediction (prediction time).

Direct observation of participant behavior and task management strategy were obtained by a researcher who was collocated in the room for the duration of each participant contact period. It was not communicated directly to the participant they were specifically being monitored to preclude any biases they might exhibit if such knowledge was known prior to the data collection session. Because post-experimental questionnaires were administered to capture an array of participant subjective feedback, responses were compared to observations made by the researcher and discussed with each participant to either confirm or clarify their experience and actions undertaken during the experiment.

Repeated measures ANOVA were used to identify statistical difference and analyze all categorical data, and regression analysis was used to analyze numerical data. An *a priori* probability level of significance was established at .05 and all analysis and calculations were completed using Minitab 17 software.

Results

For observed task management strategies utilized during 96 trials across 24 participants, Adaptive Avoidance was observed the least at 10 times, comprising 10.42% of all trials. Observation of the remaining three strategies follows: Interleaving 25, 26.04% of all trials; Adaptive Attack 30, 31.25% of all trials; and Multitasking 31, 32.29% of all trials. For all instances, participant task management strategy was not dictated to the participant, rather the participant managed tasks how they saw fit during each separate trial. Several participants changed strategies between trials resulting in unequal distributions of the four identified strategies.

Situation Awareness

A series of one way ANOVA results showed task management strategy responses approaching significance for level 1 SA - perception ($F_{3,92} = 2.58$, $p = .058$) and overall SA ($F_{3,92} = 2.28$, $p = .084$); no significance for level 2 SA - comprehension ($F_{3,92} = 1.84$, $p > .05$); and significance for level 3 SA – prediction ($F_{3,92} = 2.84$, $p < .05$). These results are shown graphically using pooled standard deviations for interval calculation and 95% confidence interval bars in Figure IV-5.

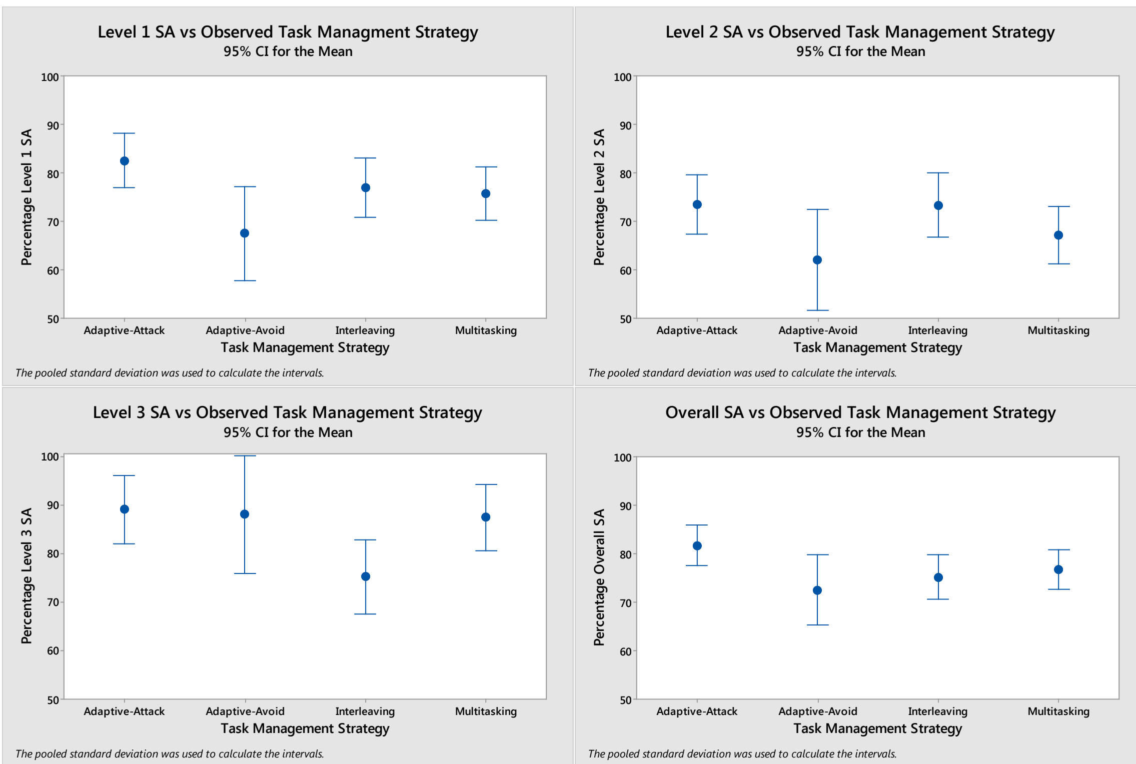


Figure IV-5 Observed Task Management Strategy Influence on SA

Mean values reflect the Adaptive Attack strategy resulting in the highest outcome (M = 82.50, SD = 12.25) followed by Interleaving (M = 76.93, SD = 16.60) and Multitasking (M = 75.81, SD = 16.58) for level 1 SA. Adaptive Avoidance ranked lowest (M = 67.50, SD = 17.32). This trend continued on to level 2 SA, however results did not approach significance with p = .145. Level 3 SA measured as prediction accuracy did show significance toward Adaptive Attack (M = 89.00, SD = 12.42), Adaptive Avoidance (M = 88.00, SD = 19.32), and Multitasking (M = 87.42, SD = 22.36) resulting in more favorable outcomes over Interleaving (M = 75.20, SD = 21.63). Lastly, Overall SA approached significance at p = .084 and showed Adaptive Attack (M = 81.66, SD = 8.32) results higher than Multitasking (M = 76.77, SD = 13.55), Interleaving (M = 75.16, SD = 12.89), and Adaptive Avoidance (M = 72.53, SD = 9.72).

To interpret meaning from these findings it is necessary to recall the definitions of each task management strategy as defined in Table IV-1. For levels 1 and 2 SA, Adaptive Attack's aggressive pursuit of the secondary task and Interleaving's serial task switching method resulted in a better ability to perceive data on both the primary and secondary tasks. No inference can be derived for level 2 SA, because findings were not significant. Considering the differences between the higher performing Adaptive Attack and Interleaving strategies the implication is that dedicating time toward one task for a longer duration of time has a positive effect. Multitasking and Adaptive Avoidance both involved some degree of constantly shifting between tasks implying the less time dedicated to each task before a switch decision was made to go back to the other resulted in negative outcomes toward perception. For level 3 SA, Adaptive Attack showed the best ability to predict problematic behavior among the processes with Adaptive Avoidance ranking second. The second place rating for level 3 SA and Adaptive Avoidance can be attributed to the purposeful disengagement from the secondary task. It was expected that the Adaptive Avoidance strategy would have the participant more attune to how the process deviations were trending over time, because Adaptive Avoidance is defined as disengagement from the secondary task with sole focus on the primary. This ran counter to the findings for Adaptive Avoidance on levels 1 and 2 SA where it ranked last in both cases. However it was noted for level 3 SA – the only data where significance was found – the top three strategies, Adaptive Attack, Adaptive Avoidance, and Multitasking (in that order) produced means that differed by only 1.58 on a scale of 100. Interleaving is not included in the previous list because it was by far the worst strategy associated with level 3 SA. For level 3 its mean fell well below the next ranked strategy (Multitasking was third) by a mean delta of 12.22. Overall SA results were a combination of all three levels of SA and reflect the aggregate influence toward for each task management strategy. Overall SA results favored Adaptive Attack first, followed by Multitasking, yet were not quite statistically significant, $p = .084$. Overall SA was an aggregate of the three underlying levels. Since Adaptive Attack ranked highest in each of the individual levels of SA, it was highest in overall SA as well.

Performance

Measures of performance for the purposes of scoring encompassed a total point value of 1000 points per trial. This was split between the primary and secondary tasks using an 80:20 ratio weighted in favor of the primary process monitoring task. Scoring methods for both tasks individually are discussed below to aid in the interpretation of the results that follow.

Primary Task Scoring: Specific performance measures for the primary task were evaluated in three ways to calculate the overall aggregate primary task score. First, the experimental simulation captured response times to deviations and awarded point values based upon a time weighted rate of decay algorithm [Equation: *Score = 20 – 5t*, where *t* represents an integer value from 0 to 4 iterated once by the simulator every 5 seconds after the appearance of a deviation]. Thus the algorithm for deviation acknowledgement awarded 20 points each event, but the participant netted 0 points after 20-seconds of time had elapsed without an acknowledgement. Given there were 10 deviations per trial, the highest possible deviation acknowledgement score per trial was worth 200 points. The next measure of performance toward the overall primary task score was the participant's ability to predict where deviations were going to occur. Each successful deviation prediction by the participant awarded 20 points, with successful predictions for all 10 deviations resulting in 200 points. There was no time factor associated with the deviation prediction scoring method, because deviations occurred at preprogrammed random intervals making this a discrete, all or nothing metric (It was noted none of the study participants were able to successfully predict all 10 deviations for any trial). The last contributing measure of performance for the primary task was participant SAGAT responses. There were 16 total SAGAT questions per trial (6x level 1 SA; 6x level 2 SA; 4x level 3 SA) worth 25 points each, for a 400 point potential award value. The three measures of performance for the primary task: deviation acknowledgement, deviation prediction, and SAGAT responses were aggregated together to produce an overall primary task score worth as high as 800 points.

Secondary Task Scoring:

The secondary task was a 12^{th} grade level reading comprehension exam based upon standard SAT questions and scoring methods. The secondary task consisted of eight questions, each worth 25 points for a grand total of 200 points. Correct responses were awarded the full 25 points per question, whereas incorrect responses reduced the score by 5 points; no response neither added to nor subtracted from the score, yielding 0 points.

Scoring Based Performance Results: Deviation management score was a combination of deviation acknowledgement and prediction scoring. This measure was not found to be statistically significant ($F_{3,92} = 1.35$, $p > .05$), thus task management strategy did not have a significant effect on participant ability to acknowledge and predict deviations as part of the primary task. Adding in SAGAT responses to generate the overall primary task score found results approaching significance ($F_{3,92} = 2.40$, $p = .073$) with a trend toward higher performance by use of either the Adaptive Attack (M = 68.16, SD = 6.01) or Multitasking (M = 65.12, SD = 8.62) strategy. Interleaving (M = 64.29, SD = 11.90) and Adaptive Avoidance (M = 59.94, SD = 7.32) produced less favorable results on the overall primary task score. Secondary task performance scores did show significance ($F_{3,92} = 5.27$, $p = .002$) as did overall trial scores ($F_{3,92} = 6.00$, $p = .001$), which were an aggregate of both the primary and secondary task scores combined. Findings revealed both secondary task and overall trial scores reflected higher performance outcomes for the task management strategies Adaptive Attack and Multitasking in that order, with Adaptive Avoidance faring worst each case. Secondary task score results were ranked as follows: Adaptive Attack (M = 45.50, SD = 24.05), Multitasking (M = 31.21, SD = 20.90), Interleaving (M = 26.70, SD = 27.88), Adaptive Avoidance (M = 15.25, SD = 19.45) to reflect a distinct advantage for the Adaptive Attack strategy. Overall trial scores followed suit and produced the same outcome ordering: Adaptive Attack (M = 636.3, SD = 71.7), Multitasking (M = 583.4, SD = 92.3), Interleaving (M = 567.7, SD = 108.9), Adaptive Avoidance (M = 510.0, SD = 62.1). Results from repeated measures one way ANOVA are shown in Figure IV-6.

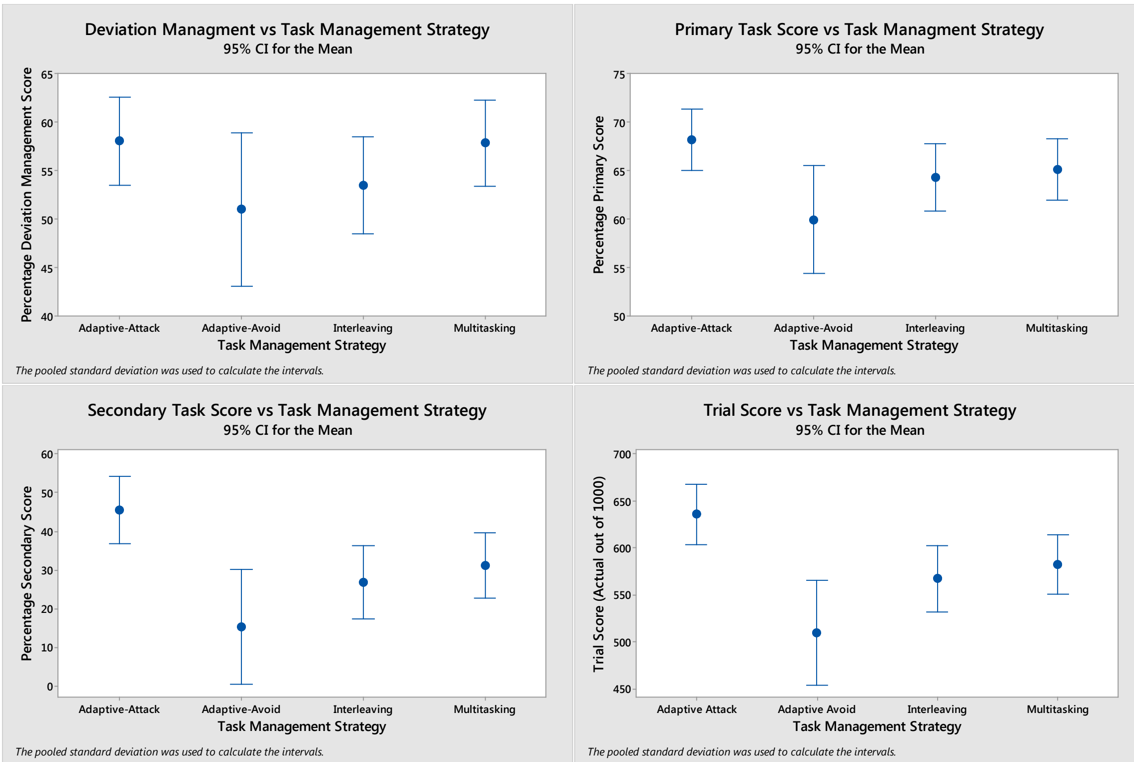


Figure IV-6 Performance Outcomes by Task Management Strategy

Interpreting only results where significance was discovered for both the secondary task and overall trial scores, Adaptive Attack and Multitasking provided the most favorable outcomes. As defined, this reveals the Adaptive Attack strategy of completing the secondary task quickly produced a favorable overall result, with Multitasking – the effects of splitting attention – not yielding as high a result, but having a similar effect on performance outcomes. For the overall primary task score, which approached significance, and deviation acknowledgement time, where no significance was found, the trend held in favor of Adaptive Attack and Multitasking for improved performance based on these particular task management strategies.

***Non-Scoring Measures of Performance / Results*:** In addition to the scoring methods listed above two alternative measures of performance were captured for each trial using time based events. These were associated with when deviations actually occurred and were identified as response times (the actual time it took a participant to perceive and acknowledge a deviation via mouse click) and prediction times (measured as the time it took to assess all 8 processes being monitored and commit to a deviation prediction via mouse click). Both time based event measures produced results of significance for task management strategy using one way repeated measures ANOVA. Response times ($F_{3,92} = 2.96$, $p < .05$) were nearly 3.5 seconds faster on average between first and last position, Adaptive Attack (M = 6.80, SD = 2.49) and Interleaving (M = 10.29, SD = 8.13), respectively with Multitasking (M = 7.39, SD = 3.16) and Adaptive Avoidance (M = 9.54, SD = 2.93) falling in between the two extremes. Prediction times ($F_{3,92} = 3.30$, $p < .05$) produced similar results where 12.8 seconds separated the fastest mean times to predict, Multitasking (M = 22.39, SD = 8.68) and slowest, Adaptive Avoidance (M = 35.23, SD = 15.46). In this case, Adaptive Attack (M = 22.62, SD = 13.29) and Interleaving (M = 26.81, SD = 13.69) were between the two. In both cases of time-based measurement, ANOVA results were significant and favored the Adaptive Attack and Multitasking strategies. These results are presented graphically in Figure IV-7.

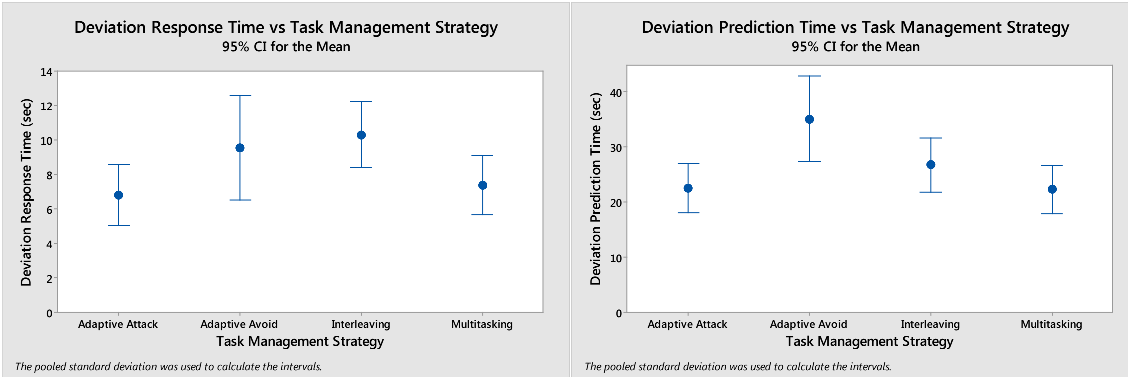


Figure IV-7 Performance Outcomes from Actual Event Time Measurement

Performance Results Summary: Notable are all measures of performance, those that produced scoring results and those based on event time supported the Adaptive Attack and Multitasking strategies for improved outcomes. Reasons for this are twofold: first, Adaptive Attack had a distinct advantage in performance on the primary task activities, once the participant had completed the secondary task. This benefitted overall primary task scores and produced higher aggregate outcomes as well due to relegation of the trial to a single task after dispatching the secondary. Next, Multitasking participants exhibited some of the same characteristics as Adaptive Attack, however they did not tend to complete the secondary task, rather remained engaged in both tasks fully until the end of each trial. Therefore, Multitasking and Adaptive Attack may have had similar performance results up until the point a conscious decision was made to completely set aside the secondary task. Notable differences in field of view strategy may have provided for a positive influence for Adaptive Attack and Multitasking results as well. Participants who engaged in both strategies were inclined to lift the secondary task from the table in attempts to enhance their field of view toward the displays. This behavior was not as noticeable or common in the Adaptive Avoidance and Interleaving strategies.

And since it was hypothesized Multitasking would result in degradations in performance it was an unexpected finding Multitasking fared very well, perhaps due in part to associate field of view strategy, not part of the focus of this research.

Workload

A NASA-TLX was automatically administered at the end of each trial by the simulator. Repeated measures one way ANOVA results were found to be significant across all six TLX factors, except the “physical” factor. This was unexpected, because the experiment was not a physically demanding task, rather participants remained seated in an ergonomically adjustable chair for the duration of all trials. The assertion the task was not physically exerting or strenuous was reinforced by researcher observations whereby none of the participants appeared to be under undue physical duress. Despite this assessment of the experimental design and observation data, the TLX reflected several high ratings from participant’s for the physical factor. These ratings were analyzed and found to be accompanied by high variability across all responses to the TLX factor, physical. The high degree of variability across participant responses for this factor were a contributing reason why statistical significance was not found.

Overall TLX results were tabulated as both raw and weighted. Both produced similar results. Raw TLX findings were significant ($F_{3,92} = 5.96$, $p = .001$) and showed Adaptive Attack (M = 45.83, SD = 18.04) resulting in the lowest rating of workload followed by Multitasking (M = 44.76, SD = 13.75), Adaptive Avoidance (M = 55.33, SD = 17.25), and Interleaving (M = 59.77, SD = 11.18). Weighted TLX results were also significant ($F_{3,92} = 6.36$, $p = .001$) and very similar in kind, revealing again Adaptive Attack (M = 53.29, SD = 18.88) and Multitasking (M = 51.95, SD = 17.36) on top with

Adaptive Avoidance (M = 64.97, SD = 23.57) and Interleaving (M = 70.37, SD = 14.43) fairing worse. Of the individual TLX factors showing significance, all indicated Adaptive Attack and Multitasking strategies resulted in lower experienced workload with one exception being Effort where Adaptive Avoidance resulted in the lowest. Individual TLX factor data are summarized in Table IV-3 and ANOVA graphs are depicted visually in Figure IV-8 where a pattern emerges showing reduced workload experienced using the Adaptive Attack and Multitasking task management strategies, and Interleaving resulting in the heaviest user experienced workload.

	F – Statistic / p-value	Strategy	N	Mean	SD
Mental	($F_{3,92}$ = 2.72, p < .05) **p = .049**	AAt	30	59.33	23.44
		AAv	10	64.00	28.94
		I	25	72.60	15.08
		M	31	55.65	25.65
Physical	($F_{3,92}$ = 0.66, p >.05) p = .578	AAt	30	19.50	16.10
		AAv	10	21.50	18.27
		I	25	17.00	12.58
		M	31	15.32	13.66
Temporal	($F_{3,92}$ = 4.32, p = .007) **p = .007**	AAt	30	51.83	26.80
		AAv	10	64.00	29.04
		I	25	69.40	19.60
		M	31	48.23	22.82
Performance	($F_{3,92}$ = 8.70, p < .001) **p = .000**	AAt	30	40.50	15.11
		AAv	10	69.50	16.91
		I	25	58.00	21.31
		M	31	48.55	16.89
Effort	($F_{3,92}$ = 4.95, p = .003) **p = .003**	AAt	30	54.50	26.98
		AAv	10	48.50	12.70
		I	25	72.40	13.70
		M	31	56.13	20.72
Frustration	($F_{3,92}$ = 7.64, p < .001) **p = .000**	AAt	30	47.67	18.60
		AAv	10	64.50	28.52
		I	25	69.20	25.07
		M	31	44.68	18.97
Key:	AAt – Adaptive Attack I – Interleaving	AAv – Adaptive Avoidance M - Multitasking			

Table IV-3 TLX Individual Factor Summary Data

Bolded p-values indicate significance, strategies highlighted in gray represent the lowest experienced workload for the given factor, only for factors of significance.

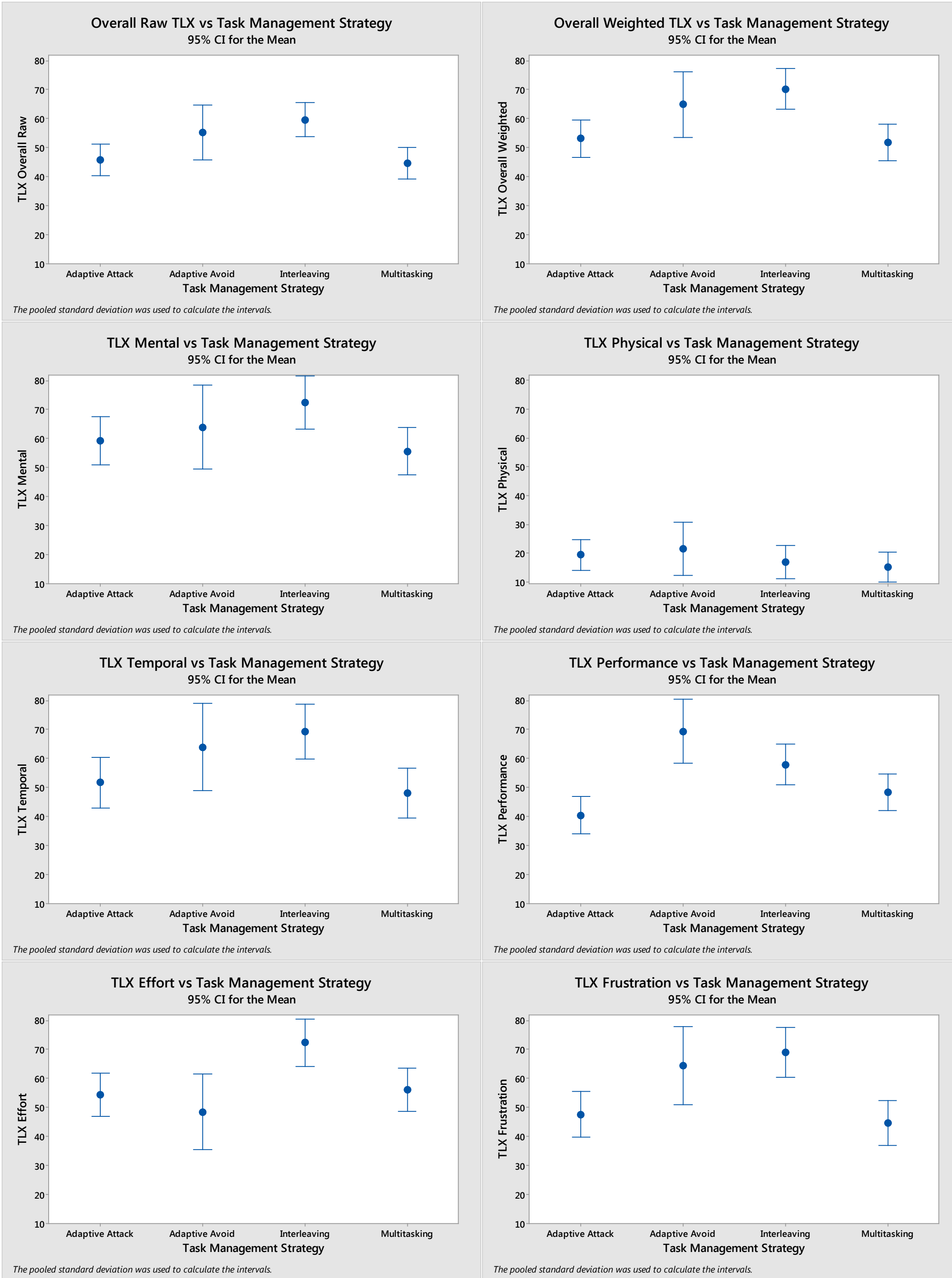


Figure IV-8 Task Management Summary NASA-TLX Workload ANOVA Results

Subject Variables

Age and experience were selected as two items of interest from the demographic data obtained from each participant. Regression analysis was chosen for both factors to derive equations for identified areas of significance.

Age – SA

Regression analysis showed significance for level 1 ($F_{1,94} = 12.80$, $p = .001$), level 3 ($F_{1,94} = 4.90$, $p < .05$), and overall ($F_{1,94} = 12.90$, $p = .001$) SA, but not for level 2 ($F_{1,94} = 1.96$, $p > .05$). Regression Equations are shown by percentage SA in Table IV-4.

Regression Equations for SA (where significance was found)
Level 1 SA % = 95.03 - 0.425 (Age)
Level 3 SA % = 99.03 - 0.342 (Age)
Overall SA % = 90.67 - 0.318 (Age)

Table IV-4 Regression Equations of Subject Variable Age Relative to SA

Equating regression results to the experimental investigation it was found there existed a decay in SA for a 40 year span of age at levels 1, 3, and overall. Evaluation at both age extremes, 20 and 60, resulted in the following percentage losses of SA: level 1 – 17%; level 3 – 13.68%; overall – 12.72%. Therefore a decrement in perception, prediction, and overall SA was noticed for the given subject pool as age increased. These losses were driven by participant responses to SAGAT polls and prediction capabilities, which indicated a reduced correct response rate as age increased.

Age – Performance

Regression analysis showed significance for a decrement in performance metrics due to age as well. This data is summarized and equations provided for both percentage and actual values in Table IV-5.

Deviation Management Score	($F_{1,94} = 4.82$, $p < .001$) $p = .000$	Dev Mgt Score % = 69.35 - 0.3196 (Age)			
		Dev Mgt Score = 277.4 - 1.278 (Age)			
		Mean	**SD**	**Min**	**Max**
		224.18	50.42	94.00	360.00
Primary Task Score	($F_{1,94} = 5.30$, $p < .001$) $p = .000$	Primary Score % = 77.53 - 0.2933 (Age)			
		Primary Score = 620.2 - 2.347 (Age)			
		Mean	**SD**	**Min**	**Max**
		522.52	72.10	325.25	698.75
Secondary Task Score	($F_{1,94} = 5.95$, $p < .001$) $p = .000$	Secondary Score % = 58.07 - 0.606 (Age)			
		Secondary Score = 116.1 - 1.212 (Age)			
		Mean	**SD**	**Min**	**Max**
		65.68	50.63	-40.00	170.00

Table IV-5 Regression Equations of Subject Variable Age Relative to Performance

These findings suggest there is a decrement associated with an increase in age. Putting this into context, the mean score for deviation management was 224.18 and the regression equation given as [Dev Mgt Score = 277.4 – 1.278 (Age)] results in a net loss on this performance metric of 12.78 percentage points for every 10 years of age. This equates to a loss of 51.12 points across the entire age range encompassing all participants in the study, ~20-60 years of age; roughly a 12.5% loss in point value for a 400 total point scored task over 40 years. The overall primary task score yielded similar results, with a 93.88 point loss across the 20-60 year age span for an 800 total point task; roughly a 11.7% loss in performance over 40 years. Completing analysis of the subject variable age, a 40 year span results in a 48.48 point loss for the secondary task. The decrease here is higher than the primary given that the secondary task was only a 200 total point task, meaning a 24% loss in performance over a 40 year span.

Age – Workload

Using Levene's test for equal variances only raw TLX data was found to have significant ($p < .05$) differences for the effect of age on workload. Contrary to this, weighted TLX data showed no significance for variance equality, thus the raw overall workload results are reported herein. Note that individual factors associated with the weighted workload shall not be considered due as well to the findings for homogeneity not being satisfied for overall weighted results ($p = .285$). Results from Levene's test and overall raw workload ANOVA results are shown in Figure IV-9.

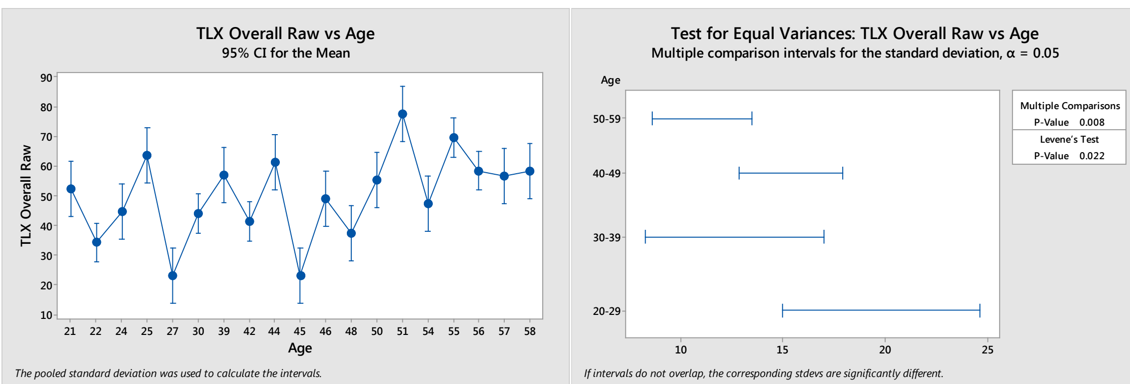


Figure IV-9 Age vs. Overall Raw NASA-TLX and Levene's Test Results

From one way ANOVA results ($F_{18,77} = 11.55$, $p < .001$) it is inferred increases in age result in an increased participant experienced workload for the purposes of a process control monitoring task environment. This is especially noticeable between the age groups 20-29 and 50-59 where corresponding standard deviations are significantly different.

Age – Additional Analysis

The subject variable age poses the potential risk for confounding factors toward the resultant data. For the population size sampled age ranged from 21-58 years and spanned multiple generations. Although results of significance were found showing decrements associated with age toward SA, performance, and workload outcomes it must be noted age correlated directly with task management strategy utilization as well. A correlation table is provided in Table IV-6, where it is revealed the task management strategy selected by each age group was significant.

Task Management Strategy	Age: 20-29	Age: 30-39	Age: 40-49	Age: 50-59
Adaptive Attack	**0.234** **0.022**	0.017 0.870	**0.234** **0.022**	**-0.429** **0.000**
Adaptive Avoidance	**-0.197** **0.055**	-0.129 0.211	0.118 0.252	0.158 0.123
Interleaving	**-0.233** **0.022**	-0.153 0.138	-0.123 0.231	**0.423** **0.000**
Multitasking	0.116 0.261	**0.211** **0.040**	**-0.193** **0.060**	-0.075 0.469
	Cell Contents: Pearson correlation P-Value			

Table IV-6 Correlation: Age and Task Management Strategy
(values in Blue reflect significance; values in Orange approach significance)

The age group 20-29 reflected a correlation toward the Adaptive Attack strategy and away from Adaptive Avoidance and Interleaving. On the other end of the spectrum the 50-59 age group had a strong tendency to use the Interleaving strategy and not the Adaptive Attack strategy. Since favorable outcomes were associated with the strategies utilized more often by the younger group and there exists a correlation between age and strategy the findings in the data might reflect more than simply the subject variable age alone. Other factors must be taken into consideration when spanning generational divides. For instance, age may be influenced by video gaming experience, education level, or occupational experiences. A brief review of these factors showed for the subjects participating in the experiment in the age group 40-59, 53% of them had no gaming experience at all. This was a contrast to the age group 20-39 where this number was only 33% of participants lacking gaming experience. Education level and occupation were also examined, but found virtually equal or negligible division for those subjects who had higher education (20-39, 78% vs. 40-59, 80%) and / or held professional positions (20-39, 60% vs. 40-59, 78%). Self-report data from the post-experimental questionnaire revealed differences across age groups as well. Questions pertaining to how long an individual could reasonably spend engaged in an activity were asked. The self-reported maximums shown in Table IV-7 reflect further comparisons and contrasts between the subject population age groups with respect to how long each felt they could perform a suggested activity.

	Max Hours Watching TV		
		Mean	Std Dev
Age	20-29	3.58	1.77
	30-39	3.67	2.31
	40-49	3.83	2.23
	50-59	2.67	1.32
	Max Hours Video Gaming		
		Mean	Std Dev
Age	20-29	3.25	1.89
	30-39	1.67	0.58
	40-49	1.83	3.06
	50-59	0.50	0.50
	Max Hours Monitoring a Display		
		Mean	Std Dev
Age	20-29	4.50	4.36
	30-39	4.00	2.00
	40-49	2.67	2.66
	50-59	2.17	1.54

Table IV-7 Self-reported Maximums by Age

Notable in Table IV-7 are the maximum number of hours for self-report watching a television do not differ that greatly. However, looking at the data for mean number of maximum hours subjects felt they could reasonably spend playing video games and monitoring a computer display vary quite heavily. These last two may or may not have contributed to the decrements noted with the factor age previously covered, but further investigation beyond the scope of this effort is necessary to properly vet these concerns.

Experience – SA and Performance

Participants were not required to have previous experience with the process control industry, however due to the population sampled it was inevitable some would. For those that had experience it was expected an increase in age would result in an increase in the number of years associated with process control familiarity. Both cases were found to be true among the participant pool where 17 out of 24 total participants had no experience whatsoever and the remaining 7 had a mean age of 48 (SD = 10.6) and mean number of years experience with the process control industry of 12.5 (SD = 8.8). Statistical significance between age and process control experience was highly expected and found across the participant pool to be ($F_{18,5}$ = 10.85, p = .008).

One way repeated measures ANOVA were used to analyze the process control experience data across factors of all three levels of SA, task performance, and workload. No significance was found between all levels of SA, thus process control experience had no effect on SA. ANOVA results for performance, however, did show the negative effects of process control experience on the secondary task ($F_{7,88}$ = 3.55, p = .002) and prediction times ($F_{7,88}$ = 10.98, p < .001), but nothing of significance for all other measures of performance. These results are shown in Figure IV-10.

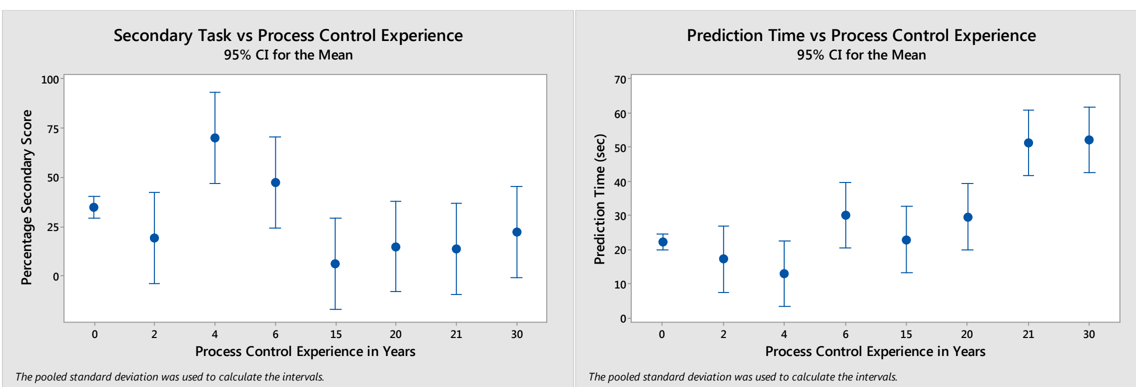


Figure IV-10 Effects of Process Control Experience on Performance Measures

Findings indicate a strong possibility that individuals with higher levels of experience with process control are more inclined to place priority, and thus focus attention, on the primary task. This was supported by direct researcher observation and could mean the secondary task was viewed as less important. Level of importance was communicated to all participants during the training and with the 80/20 scoring split. Therefore it is not unexpected nor is it considered a negative outcome that individuals with more experience were more adept at realizing lack of value in the secondary task relative to the primary. Thus the marked decline in secondary task performance by those with experience is reasonable, but did not explain why performance on the primary task responses to deviations was also degraded. Findings with response times revealed longer prediction times for those with experience. This was likely attributable to experienced individuals spending more time studying the processes in order to interpret and predict problematic process behavior. Qualitative feedback from individuals with experience confirmed this as several reported the primary task was the most important task, therefore they downplayed – exercised Adaptive Avoidance tactics on – the secondary task.

Experience – Workload

ANOVA analysis of workload found significance in all factors (all cases $p < .05$) with the exception of effort, which only approached significance with a $p = .080$. For both raw and weighted TLX, Levene's test for variance was satisfied ($p < .05$), thus results for both are presented in Figure IV-11. Overall raw ($F_{7,88} = 5.11$, $p < .001$) and weighted ($F_{7,88} = 6.01$, $p < .001$) TLX both showed significance as well.

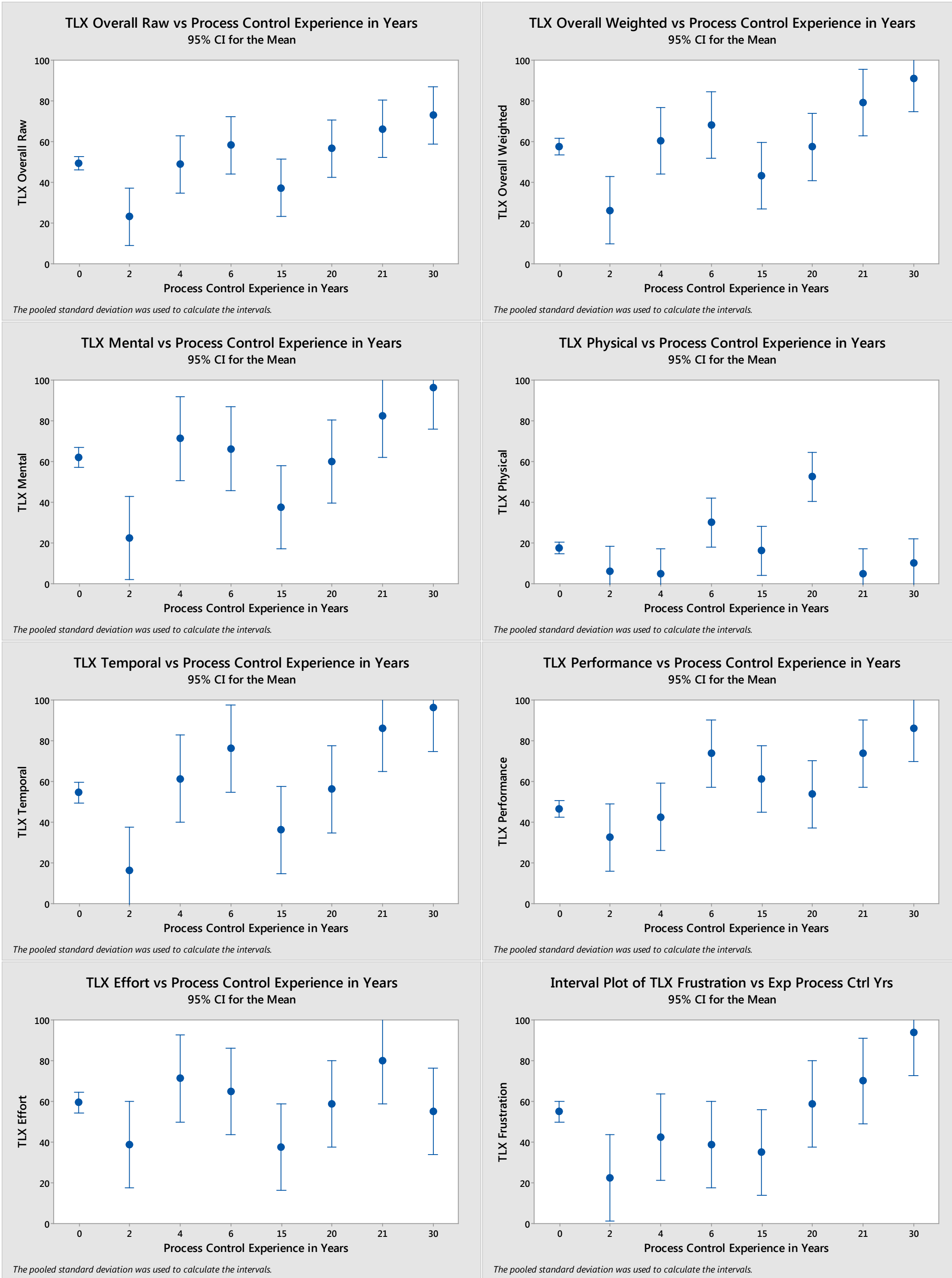


Figure IV-11 Process Control Experience NASA-TLX Workload Results
Confidence Interval bars restricted to TLX full scale, 0-100

For all cases except effort ($p = .080$) significance was noted as an increase in participant experienced workload as experience with process control systems increased. Results from physical are discounted as they were previously, because the task was not physical in nature and the outlier – having 20 years experience – biases the response across such a small group ($n = 7$) of individuals. For all other ANOVA graphs, trends support an increase in workload for higher levels of experience in process control. It is put forth here the reasoning for this may have to do with the highly experienced personnel attempting to over analyze the simulated experience in efforts to find meaning and patterns in the simulated data.

Discussion

The process control industry poses grave danger to individuals and surrounding communities when system operators have an incomplete mental model of the system under their control. Most attempts to alleviate this from happening have followed a heuristic-based or user-centered design approach to factors external from the human monitor. This is evidenced in the works of many (Chen, Haas, & Barnes, 2007; Panteli, Kirschen, Crossley, & Sobajic, 2013; Norman, 1984; Gould, 1988; Vicente & Rasmussen, 1992; Ponsa, Vilanova, Perez, & Andonovski, 2010; Shneiderman, Plaisant, Cohen, & Jacobs, 2010; U.S. Nuclear Regulatory Commission, 2002) and lent to the idea of investigating the internal factors associated with the human in control. The work of Morgan et al. (2013) put forth a framework to identify multitasking adaptability relative to varying degrees of task difficulty in search of consistent performance. Taking this a step further the idea of analyzing human-centric strategies outside of the realm of the cognitive sciences led to the theory that individual differences existed in task management strategies for process control that would produce favorable outcomes toward SA, performance, and workload.

Situation Awareness

SA outcomes identified the Adaptive Attack strategy as consistently yielding the most favorable outcomes for perception, comprehension, prediction, and overall SA. Despite only level 3 SA showing significance relative to task management, all factors yielded the highest means when participants employed the Adaptive Attack strategy. For level 3 SA specifically, only a single point difference in mean values (89.00 vs. 88.00) separated Adaptive Attack from Adaptive Avoidance. Giving critical thought to this, it is not unexpected Adaptive Avoiders would perform well on prediction, because as defined the strategy involved disengagement in the secondary task lending more time to the participant to concentrate on the activities encompassing the primary task. Also, the Adaptive Attack strategy might have achieved higher SA results due to the observed verve in which participants who employed this strategy went about the secondary task. Upon completion of the secondary task, each trial effectively became a single task experience after the Adaptive Attacker dispatched the secondary task in its entirety not to return to it later. On the other end of the spectrum, Adaptive Avoidance produced the poorest results for all cases of SA except level 3. Reasons behind negative results in all areas of SA with one exception are not readily obvious. But they might be attributable to the adaptation mechanism itself, driven by self-induced internal cueing whereby time on task for Adaptive Avoidance was insufficient during periods of non-avoidance – meaning despite a desire to avoid the secondary task it served as enough of a distractor to preclude consistent awareness about the primary task.

Performance and Workload

Both metrics supported Adaptive Attack and Multitasking having statistically significant and positive outcomes influencing process control monitoring environments. The trends for both strategies all indicated the employment of either would yield the most favorable outcomes toward performance with the least amount of operator experienced workload. This is important, because operator experienced workload is typically high in a vigilance task (Warm, Parasuraman, & Matthews, 2008) and most process monitoring tasks run for a duration of 8-hours or more. This research effort targeted a data collection session of approximately 2-hours – approximated, because of variance in SAGAT pauses across all participants existed. For this duration of time both Adaptive Attack and Multitasking strategies were sustainable. Future work is required to determine whether both are sustainable over longer durations of time, because each had a profound positive effect toward resultant outcomes making them desirable for process control operations. On the other hand, undesirable effects were associated with both the Interleaving and Adaptive Avoidance strategies. For each, performance was low while experienced workload was high, which may be attributable to the task switching mechanics involved in these strategies. Both utilized a tactic of purposeful task switching to move between tasks, but for different reasons. Interleaving task switch triggers were associated with the conscious decision for purposed engagement in competing tasks with the time spent engaged driving the trigger. Similar yet counter to this was Adaptive Avoidance exhibited the same purposed task switch trigger, however the reasoning behind the switch was associated with attempts to disengage from either task. Note by definition, Adaptive Avoidance was disengagement from the secondary task only, however direct observations uncovered participants adhering strictly to the Adaptive Avoidance strategy attempted to disengage from both tasks equally. For Interleaving and Adaptive Avoidance task switching led to higher perceived workload. This knowledge adds value to process control operators who engage in either strategy, because self-awareness of factors that increase workload can be mitigated through training methods intent on identification of task switching stressors. Further, process control room designs could take these findings into consideration and facilitate the layout of a process control room environment that intuitively fostered separation for competing tasks, thereby reducing the tendency to engage in Interleaving or Adaptive Avoidance.

Conclusion

This research took a novel approach toward examination into the influences of individual task management strategies for a process control monitoring application. Four strategies were defined and identified: Adaptive Attack, Adaptive Avoidance, Interleaving, and Multitasking. All were found to have significance for three key areas of process control: SA, performance, and workload. Adaptive Attack and Multitasking trended well in all areas of investigation while Interleaving and Adaptive Avoidance did not. Applications of process control utilizing human observers should seek to capitalize on the knowledge an individual's preferred task management strategy plays an important role in their ability to remain aware, perform well, and experience reduced workload. Process control room environments should seek to intuitively educe Adaptive Attack strategies from the human monitor to achieve the highest probability of maximizing the relationship between human and machine for a given system. It is important to mention this research did not explore personality traits, but there may exist a correlation between personal behavior and the task management strategy outcomes discussed. Future work should seek to either identify further differences between competing task management strategies or mark where each possibly converges. Future work should also take other internal factors into consideration beyond just task management strategies. These include additional individual demographic and personal factors. In sum, resolution of factors internal to the human engaged in a process control environment breaks from the norm of investigating external factors to develop a more comprehensive body of knowledge for the process control industry to assuage the negative effects of potential disastrous outcomes. This research sought to fill that gap and produced several findings of significance worthy of further investigation.

Acknowledgements

This work was supported in part by the Air Force Research Laboratory's (AFRL) Aerospace Systems Directorate (RQ) and its Component Research Air Facility (CRAF). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of AFRL, RQ, or the U.S. Government.

References

Air Force Research Laboratory / Aerospace Systems Directorate. (2014). *Facility Factsheet Component Research Air Facility (CRAF).* (AFRL/RQOEE, Ed.) Retrieved January 1, 2014, from www.wpafb..af.mil: http://www.wpafb.af.mil/shared/media/document/AFD-130410-048.pdf

Bowden, J. R., & Rusnock, C. F. (2014). Evaluation of Human Machine Interface Design Factors on Situation Awareness and Task Performance. *TS* . Wright-Patterson AFB, OH: Air Force Institute of Technology.

Brown, S. W. (1998). Automaticity versus timesharing in timing and tracking dual-task performance. *Psychological Research, 61* (1), 71-81.

Chen, J. Y., Haas, E. C., & Barnes, M. J. (2007, November). Human Performance Issues and User Interface Design for Teleoperatored Robots. *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews, 37* (6), pp. 1231-1245.

Dismukes, R. K., Loukopoulos, L. D., & Barshi, I. (2009). *The multitasking myth: Handling complexity in real-world operations.* Burlington, VT: Ashgate Publishing Company.

Duggan, G. B., Johnson, H., & Sorli, P. (2013). Interleaving tasks to improve performance: users maximse the marginal rate of return. *International Journal of Human-Computer Studies , 71* (5), 533-550.

Endsley, M. R. (2000). Direct Measurement of Situation Awareness: Validity and Use of SAGAT. In M. R. Endsley, & D. J. Garland (Eds.), *Situation Awareness Analysis and Measurement.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37* (1), 32-64.

Gould, J. (1988). How to design usable systems (excerpt). IBM Research Center Hawthorne. In M. Helander (Ed.), *Handbook of Human-Computer Interaction* (pp. 757-789). Yorktown Heights, NY, 10598, 757-789.

Haahr, M. (1998-2014). (Randomness and Integrity Services Ltd.) Retrieved 2013, from RANDOM.ORG: http://www.random.org/

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. (P. A. Hancock, & N. Meshkati, Eds.) *Human Mental Workload*, 139-183.

Hirst, W., & Spelke, E. S. (1980). Dividing Attention Without Alternation or Automaticity . *Journal of Experimental Psychology: General, 109* (1), 98-117.

Ishihara, S. (2012). *Ishihara's Design Charts for Colour Deficiency of Unlettered Persons.* Tokyo, Japan: Kanehara Trading Inc.

Monsell, S. (2003). Task Switching. *TRENDS in Cognitive Sciences, 7* (3), 134-140.

Morgan, B., D'Mello, S., Abbott, R., Radvansky, G., Haass, M., & Tamplin, A. (2013). Individual Differences in Multitasking Ability and Adaptability. *The Journal of the Human Factors and Ergonomics Society, 55* (4), 776-788.

Norman, D. A. (1984). *Cognitive engineering principles in the design of human-computer interfaces.* Human Computer Interaction Amsterdam: Elsevier Science.

Panteli, M., Kirschen, D. S., Crossley, P. A., & Sobajic, D. J. (2013). Enhancing Situation Awareness in Power System Control Centers. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)* (pp. 254-261). San Diego: IEEE.

Ponsa, P., Vilanova, R., Perez, A., & Andonovski, B. (2010). SCADA Design in Automation Systems. *3rd Conference on Human System Interactions (HSI)* (pp. 695-700). IEEE.

Schumacher, E. H., Seymour, T. K., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., et al. (2001). Virtually Perfect Time Sharing in Dual-Task Performance: Uncorking the Central Cognitive Bottleneck. *Psychological Science, 12* (2), 101-108.

Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. (2010). *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Reading, MA: Addison-Wesley Publishing Co.

Squire, P., Trafton, G., & Parasuraman, R. (2006). *Human Control of Multiple Unmanned Vehicles: Effects of Interface Type on Execution and Task Switching Times.* Salt Lake City: Naval Research Laboratory, Navy Center for Applied Research in Artificial Intelligence.

Tombu, M., & Jolicoeur, P. (2004). Virtually No Evidence for Virtually Perfect Time-Sharing. *Journal of Experimental Psychology: Human Perception and Performance, 30* (5), 795-810.

U.S. Chemical Safety and Hazard Investigation Board. (2011). *Pesticide Chemical Runaway Reaction - Pressure Vessel Explosion.* Investigation Report, No. 2008-08-I-WV, Bayer CropScience, LP, Institute, WV, August 28, 2008.

U.S. Chemical Safety and Hazard Investigation Board. (2007). *Refinery Explosion and Fire.* Investigation Report, No. 2005-04-I-TX, British Petrolium (BP), Texas City, TX, March 23, 2005.

U.S. Nuclear Regulatory Commission. (2002). *Human-system interface design review guidelines.* Brookhaven National Laboratory, Energy Science & Technology Department.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface Design: Theoretical Foundations. *IEEE Transactions on Systems, Man, and Cybernetics, 22* (4), 589-606.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance Requires Hard Mental Work and is Stressful. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50* (3), 433-441.

Yeung, N., Nystrom, L. E., Aronson, J. A., & Cohen, J. D. (2006). Between-Task Competition and Cognitive Control in Task Switching. *The Journal of Neuroscience, 26* (5), 1429-1438.

V. Conclusion and Recommendations

Chapter Overview

This chapter begins with a synopsis of the investigative research effort undertaken to explore the manner in which a human monitor functions within a process control environment. Next, the investigative questions and hypotheses from the Introduction are revisited and discussed in greater detail. Finally, the chapter concludes by noting the significance of this effort relative to the existing process control industry's body of knowledge and offers suggestions to advance the work further.

Research Overview

This research investigated factors both external and internal to the human observer engaged in a monitoring activity within a process control environmental context. A suitable location for investigative research was identified at the Air Force Research Laboratory's Component Research Air Facility. A case study and cognitive task analysis were conducted at the facility during an actual operational test involving coordination between multiple actors and equipment through the use of a centralized control architecture. This led to the development of a formal system description of the Component Research Air Facility, which yielded four very specific task networks associated with facility operators engaged in an actual process control activity (Appendix B). Most importantly, these task networks narrowed the focus of specific factors to the manner in which information was communicated via display design (external factor) and how operator task management strategy (internal factor) influenced human-machine system operational goals.

A human subjects experimental design proceeded with a twofold purpose. First the experimental design intended to capture underlying attributes associated with positive influences toward display design constructs. This was viewed as investigation of factors external to the human. The second purpose was to identify and observe those relevant internal factors to the human engaged in multiple tasks. This was done through exploration of task management strategies in search of those that promoted beneficial outcomes toward process control in general. Participants completed four ~30-minute simulated scenarios where the manner of information presentation varied across eight unique processes from numeric to graphic and orientation changed from functional grouping to spatial mapping. During each trial, task loading remained consistent to mimic real world process control operations that typically encompass a long duration vigilance activity. Deviations appeared on one of two displays and the participant was charged to acknowledge –"fix" – them as quickly as possible. Direct researcher observation captured the manner in which the participant executed both the primary monitoring of multiple processes and demanding secondary reading comprehension tasks. Information gathered was catalogued according to a predefined set of task management strategies based on the foundational cognitive principle of task switching. Participant task switching between primary and secondary tasks was not dictated by the experimental design, rather this was left to the participant to utilize their mental resources how best they saw fit. Insights from this and all data captured during the entire experimental contact time of approximately 4-hours for each participant included a host of demographic and subjective responses. Each metric and means of information collection was tailored to gather details lending toward an appropriate assessment of participant SA, primary and secondary task performance, and subjective perception of workload.

Answers to Investigative Questions

External factor: Display Design Influence on Situation Awareness

1. How does the process control information display construct used during an interactive monitoring task impact levels 1, 2, and 3 SA?

It was hypothesized a graphic means of information presentation combined with a functionally grouped orientation would result in higher level 1 (perception), level 2 (comprehension), and level 3 (prediction) SA. Resultant data did not support this hypothesis. The data showed no significant effects for any combination of information presentation (numeric vs. graphic) and orientation (functionally grouped vs. spatially mapped) on any of the levels of SA. This finding ran counter to previous investigative works that have found varying degrees of significance in competing display constructs. The reason for such contrary findings is that previous studies evaluated heuristically developed designs that featured multiple design differences, rather than isolated specific design constructs. Essentially, no two designs are ever truly the same, thus it should be expected that differing designs shall produce vastly different results.

This research broke from the norm of evaluating heuristically developed competing design constructs by seeking out the underlying manner of display attributes used to build them. This was done in the early phases of the investigative research where a case study at an operational facility was used to identify four underlying display attributes. As discussed above, the use of these attributes in an aggregated form (as combinations of information display and orientation) into a display construct had inconsequential effects on SA. However, when analyzing these four attributes individually, one attribute – graphic information presentation – did have a positive outcome toward level 3 SA. Thus the concluding finding with respect to influence of display design on SA for process control is that the ability to predict where deviations are going to occur are positively impacted by displaying information to a human monitor in a graphical manner.

External Factor: Display Design influence on Performance

2. How does the process control information display construct impact primary and secondary task measures of performance?

It was hypothesized a combination of graphic means of information presentation and functionally grouped orientation would result in higher deviation prediction and response times toward the overall primary task score. In addition to this, the same display construct was hypothesized to result in higher reading comprehension test (secondary task) scores. Neither hypothesis was supported by the data, because findings of significance were not realized for any of the four display design constructs (numeric-functionally grouped; graphic-functionally grouped; numeric-spatially mapped; graphic-spatially mapped). The underlying individual display attributes did not produce any results of significance either. Thus it was concluded the manner of information presentation and orientation to include the underlying attributes utilized in a process control environment did not have any influence on a human monitor's ability to respond to system deviations or predict where those deviations were going to occur. Nor was performance impacted by either construct or attribute on an unrelated secondary task.

Internal Factor: Task Management Strategy influence on Situation Awareness

3. In what way does the task management strategy utilized during a process control monitoring activity affect operator levels 1, 2, and 3 SA?

It was hypothesized a Multitasking task management strategy would result in lower levels 1, 2, and 3 SA than the three others (Interleaving, Adaptive Attack, Adaptive Avoidance). Results did not entirely support this hypothesis, because significance was only found at level 3 SA (prediction). For level 3 SA, predicting where deviations were going to occur, Multitasking placed third out of the four strategies in the following order: Adaptive Attack, Adaptive Avoidance, Multitasking, Interleaving. Thus, Multitasking was not the absolute worst strategy for prediction, but it was not the best either. Multitasking again placed third in level 1 SA (perception) where the resultant data approached, but did not quite achieve significance ($p = .058$). For level 1 SA the ordering was: Adaptive Attack, Interleaving, Multitasking, Adaptive Avoidance). Lastly, level 2 (comprehension) and overall SA did not yield results of significance toward the effect of Multitasking, thereby it is inferred this particular task management strategy did not have a significant influence on comprehension or SA overall in a process control environment. The factor that did appear to influence SA, however, was Adaptive Attack and is reflected in the data above where it ended up ranked first for all levels of SA. Concluding the findings of task management strategy on SA it was found Multitasking did not rank the highest at any level, but Adaptive Attack did for all.

Internal Factor: Task Management Strategy influence on Performance

4. How does operator task management strategy impact primary and secondary task measures of performance?

It was hypothesized that both adaptive task management strategies would have positive outcomes toward the primary task and a negative effect on the secondary task. [For the purposes of clarity in the following discussion about adaptation's effect on performance, Adaptive Attack and Adaptive Avoidance will be covered in two separate paragraphs below. It is also noted for this investigative question significance was found in all measures of performance data with only one exception – the overall primary (process monitoring task in its entirety) score only approached significance ($p = .073$), but did not fully achieve it.]

Adaptive Attack: Resultant data supported the hypothesis Adaptive Attack had positive outcomes toward the primary task. Two key contributing measures of the primary task in its entirety were deviation response and prediction times. For both, Adaptive Attack ranked first and second out of the four strategies, respectively (note in both cases Multitasking was in the other top position). The Adaptive Attack strategy was also ranked first in overall primary score, which was comprised of the times mentioned above added to SAGAT scoring. These findings support the Adaptive Attack strategy as having positive effects on the primary task. The same could also be said for the secondary task, where Adaptive Attack again ranked in the first position. High performance in the secondary task does not support the original hypothesis that adaptation will degrade secondary performance. Nor does overall trial score (a combination of both primary and secondary tasks) results that showed Adaptive Attack having the best outcomes. From these findings in the data it was concluded the Adaptive Attack strategy yielded the most positive performance results in a process control environment for both primary and secondary tasks.

Adaptive Avoidance: Results did not support the hypothesis that the Adaptive Avoidance strategy would have a positive effect on the primary task. Recall that Adaptive Avoidance was defined as avoidance of the secondary task in favor of the primary. However, despite avoiding the secondary task, no beneficial outcomes toward the primary were realized for the Adaptive Avoidance strategy. For deviation responses and prediction accuracy – both factors contributing to the overall primary task score – Adaptive Avoidance was ranked third and fourth out of four, respectively. This then played a major part in why the strategy also ranked last for overall primary score, because these factors were aggregated into the overall primary task score by combining them with the SAGAT scores. Similarly, the Adaptive Avoidance strategy resulted in the lowest secondary task scores, which does support the original hypothesis but does not bode well for trial performance. Conclusions from these results indicate the Adaptive Avoidance strategy has a negative influence on performance in a process control environment.

Internal Factor: Subject Variable influence on SA and Performance

5. <u>How do subject variables affect overall SA and primary task performance?</u>

It was hypothesized that individual demographic differences existed having 1) a negative effect on overall SA and 2) a positive effect on primary task performance. Results of significance were found in the data and supported the first hypothesis with respect to age. The research effort data reflected that overall SA for process control monitoring decreased as age increased with findings that showed a 12.72% decrease in overall SA over a 40-year period. However, the second hypothesis that increased process control experience would result in increased overall primary task performance was unsupported. Of the three contributing factors toward overall primary task score (deviation response/acknowledge time, deviation prediction time, and SAGAT responses) only deviation prediction time showed significance. However, counter to the hypothesis, the finding was a degradation in performance as experience increased. For this metric results revealed an actual increase in the amount of time it took to predict where a deviation was going to occur as experience increased. While this does not support the hypothesis it does possibly equate to a positive reaction from individuals with experience who had a tendency to analyze the primary task and act in a more methodical manner than those who lacked experience. In conclusion, both internal factors found increases in age that resulted in a decrease in overall SA and increases in process control experience that resulted in longer times to predict deviations.

Recommendations for Future Research

This research effort was successful in answering the investigative questions put forth and also identified several key factors leading to beneficial outcomes toward SA, task performance, and workload in a process control environment. While this work yielded results of significance other areas of further investigation remain that fall into two categories: recommendations for experimental design and the analysis of the existing residual data set from this study. Both are detailed further below.

Recommendations for Future Experimental Design

Development of the experimental design relied on the methods and metrics necessary to answer the investigative questions in this body of work. Lessons learned along the way in the development process led to the identification of three areas to consider for future experiments: SA Metrics, Correlations to Vigilance, and Display Design Attributes.

SA Metrics: This experiment utilized Endsley's SAGAT method to measure SA. While SAGAT was selected as the best method for this experimental design, it is not the only method available to gauge SA. Other SA metrics in the field of human factors exist and should be explored to see if greater granularity in the SA data set can be realized. These other metrics may be less intrusive than the SAGAT method, which involves the use of experimental pauses to query participants with context specific questions. While the SAGAT method and its use in this experimental design were robust, there is no way to tell if a competing metric would have produced the same results. Future work should consider evaluation of alternatives to SAGAT to determine how each compares and contrasts in the capture and analysis of SA data.

Correlations to Vigilance: Another area to consider in future designs is with the composition of the primary task relative to a vigilance monitoring activity. For this experimental design, deviations were presented at a rate of 2-minutes on average for every trial to mimic the real world process control interaction observed by the facility operators during the case study of the Component Research Air Facility. This time was selected primarily to strike a balance between mimicking the vigilance tasks experienced by process control operators yet provide for sufficient opportunities of data acquisition triggered by deviation events. Because the trial time was limited to 30 minutes to be able to run a participant through four trials in one sitting and the availability of facility resources these deviation triggers may have been too frequent to adequately reflect a vigilance activity. This per trial time factor could be increased to be a more accurate reflection of the vigilance task most operators encounter in a typical monitoring application, at least 8-hours.

Display Design Attributes: The process control room is a dynamic environment filled with many distractions. This research attempted to resolve the investigation down to only those underlying display design attributes that were anticipated to provide the best contrast between means of information and orientation. To do this the experimental design eliminated as many potentially confounding factors as possible. These include the use of sound, global alarm indication, and varying levels of colored indicators (e.g. warning is yellow, alarm is red) and other distractions commonly found in a process control environment. Thus it is advisable toward future work to explore the many other underlying display construct attributes that remain.

Further Analysis of Residual Data

This study captured a large amount of data only a portion of which contributed to answering the investigative questions and hypotheses presented. Thus the remaining data set is ripe with information that could be further analyzed for areas of significance not covered herein. Two areas of the existing data set remain unexplored and could be refined further to either support a similar research effort or cover topics left unanswered. These areas are the subject variables collected in the demographics questionnaire and further investigations into the task management strategies.

Subject Variables: Remaining data captured as part of the formal experiment includes many factors associated with participant subjective feedback questionnaires (e.g. preferences for display construct and self reported task management strategy), researcher observations, and demographic information.   The only areas of the demographics that were heavily scrutinized were the factors age and experience, but several interesting ones remain. These include items from education and occupational data to how much sleep the participant got the night before the experiment. On the subjective feedback side, individual preferences were collected for competing designs and other factors of the experiment. The area of participant preference would be especially worthy of further investigation to see if participant preferred display design constructs actually resulted in better performance.

Task Management Strategies: The last area suggested for future analysis of the existing data lies with the task management strategies variable. Four strategies were identified and defined as part of this research effort, however this list is by no means considered definitive. Further refinement of the strategies and investigation of the effects of underlying behaviors may produce alternative toward SA, performance, and workload. The Adaptive Attack strategy is one that mandates further exploration, because it faired so well in so many areas of this study. Specifically how the participant performed upon completion of the secondary task when using the Adaptive Attack strategy would produce interesting results. There is sufficient data to explore this specific idea further: How did the participant perform on the primary task during execution of the secondary and how did they perform after the secondary was completed? A potential investigative question would be: How did the adaptation impact performance after the multi-task environment was relegated to a single task environment upon completion of the secondary task. Notable is some Adaptive Attack participants completed the secondary task as quickly as 4-minutes into the trial (only one participant accomplished it this quickly on one trial). Exploring the existing data may yield results showing Adaptive Attack is actually similar in kind to Multitasking until the secondary task is completed. If substantiated, this would suggest the substantial bump in performance realized by the Adaptive Attack strategy was a consequence of completing the secondary task early and reducing the multiple task environment into a single task.

Summary

By exploring factors both external and internal to the human observer in a process control environment this research identified areas for improvement in the evaluation of display designs and the influence of task management strategies on facility operators. External factor findings suggest investigation into underlying display construct attributes should be studied instead of simply performing evaluations on a set of competing, heuristically developed designs. Supporting this was evidence from a study into the effects of display design on SA and performance showing irrelevancy toward four competing designs, but significance in one of the underlying attributes: graphic information presentation. For the internal factors investigated, task management strategy, age, and experience, operator preferred strategy was found to be just as important to process control outcomes as the external factors most often investigated. Adaptive Attack and Multitasking were found to be the most effective for achieving desired SA, performance, and workload. Age and experience with process control, on the other hand resulted in decreased SA and performance results. These findings were uncovered by giving equal credence to both factors – external and internal to the human observer – in a combined approach for evaluation of process control environments. They have elicited areas for further consideration to improve SA and task performance while reducing operator workload and even suggested methods to garner better results from display designs for the process control industry.

References

Air Force Research Laboratory / Aerospace Systems Directorate. (2014). *Facility Factsheet Component Research Air Facility (CRAF).* (AFRL/RQOEE, Ed.) Retrieved January 1, 2014, from www.wpafb..af.mil: http://www.wpafb.af.mil/shared/media/document/AFD-130410-048.pdf

Brown, S. W. (1998). Automaticity versus timesharing in timing and tracking dual-task performance. *Psychological Research, 61* (1), 71-81.

Carr-Chellman, A., & Savoy, M. (2004). User-Design Research. In D. H. Jonassen, *Handbook of Research on Educational Communications and Technology. A Project of the Association for Educaitonal Communications and Technology* (2nd ed., pp. 701-716). Mahwah, NJ: Routledge.

Chen, J. Y., Haas, E. C., & Barnes, M. J. (2007, November). Human Performance Issues and User Interface Design for Teleoperatored Robots. *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews, 37* (6), pp. 1231-1245.

Cummings, M. L., Bruni, S., & Mitchell, P. J. (2010). Human supervisory control challenges in network-centric operations. *Reviews of Human Factors and Ergonomics, 6* (1), 34-78.

Dismukes, R. K., Loukopoulos, L. D., & Barshi, I. (2009). *The multitasking myth: Handling complexity in real-world operations.* Burlington, VT: Ashgate Publishing Company.

Duggan, G. B., Johnson, H., & Sorli, P. (2013). Interleaving tasks to improve performance: users maximse the marginal rate of return. *International Journal of Human-Computer Studies , 71* (5), 533-550.

Endsley, M. R. (2000). Direct Measurement of Situation Awareness: Validity and Use of SAGAT. In M. R. Endsley, & D. J. Garland (Eds.), *Situation Awareness Analysis and Measurement.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37* (1), 32-64.

Endsley, M. R., & Jones, D. G. (2003). *Designing for situation awareness: An approach to user-centered design.* US: Taylor & Francis.

Gould, J. (1988). How to design usable systems (excerpt). IBM Research Center Hawthorne. In M. Helander (Ed.), *Handbook of Human-Computer Interaction* (pp. 757-789). Yorktown Heights, NY, 10598, 757-789.

Hancock, P. A., Jagacinski, R. J., Parasuraman, R., Wickens, C. D., Wilson, G. F., & Kaber, D. B. (2013). Human-Automation Interaction Research Past, Present, and Future. *Ergonomics in Design: The Quarterly of Human Factors Applications , 21(2)*, pp. 9-14.

Handal, C., & Ikuma, L. H. (2012). Good Interface Design Improves Situation Awareness in Control Room Operators. In G. Lim, & J. W. Herrman (Ed.), *Proceedings of the 2012 Industrial and Systems Engineering Conference.* Orlando.

Hirst, W., & Spelke, E. S. (1980). Dividing Attention Without Alternation or Automaticity . *Journal of Experimental Psychology: General, 109* (1), 98-117.

Huibin, J., & Wang, L. (2009). Applying Situation Awareness to Human-Machine Interface Design of Aviation. *Second International Symposium on Knowledge Acquisition and Modeling. 1*, pp. 387-390. Wuhan: IEEE.

Jeffries, R., Miller, J. R., Wharton, C., & Uyeda, K. (1991). User interface evaluation in the real world: a comparison of four techniques. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, (pp. 119-124).

Landry, S. J., & Jacko, J. A. (2004). The use of enhanced information displays to support pilot procedure following. *Digital Avionics Systems Conference.*

Li, X., Horberry, T., & Powell, M. (2010). Human Control in Mineral Processing Plants: An Operator Centered Investigation. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 284-288). Human Factors and Ergonomics Society.

Monsell, S. (2003). Task Switching. *TRENDS in Cognitive Sciences, 7* (3), 134-140.

Morgan, B., D'Mello, S., Abbott, R., Radvansky, G., Haass, M., & Tamplin, A. (2013). Individual Differences in Multitasking Ability and Adaptability. *The Journal of the Human Factors and Ergonomics Society, 55* (4), 776-788.

Moyle, S. A. (2005). Process aggregation and the situation awareness display. *3rd IEEE International Conference on Industrial Informatics (INDIN)* (pp. 623-626). IEEE.

Mumaw, R. J., Roth, E. M., Vicente, K. J., & Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 42* (1), 36-55.

Norman, D. A. (1984). *Cognitive engineering principles in the design of human-computer interfaces.* Human Computer Interaction Amsterdam: Elsevier Science.

Panteli, M., Kirschen, D. S., Crossley, P. A., & Sobajic, D. J. (2013). Enhancing Situation Awareness in Power System Control Centers. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)* (pp. 254-261). San Diego: IEEE.

Ponsa, P., Vilanova, R., Perez, A., & Andonovski, B. (2010). SCADA Design in Automation Systems. *3rd Conference on Human System Interactions (HSI)* (pp. 695-700). IEEE.

Scholtz, J. C., Antonishek, B., & Young, J. D. (2005). Implementation of a Situation Awareness Assessment Tool for Evaluation of Human–Robot Interfaces. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 35* (4), 450-459.

Schumacher, E. H., Seymour, T. K., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., et al. (2001). Virtually Perfect Time Sharing in Dual-Task Performance: Uncorking the Central Cognitive Bottleneck. *Psychological Science, 12* (2), 101-108.

Sheridan, T. B., & Parasuraman, R. (2005). Human-Automation Interaction. *Reviews of Human Factors and Ergonomics, 1* (1), 89-129.

Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. (2010). *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Reading, MA: Addison-Wesley Publishing Co.

Squire, P., Trafton, G., & Parasuraman, R. (2006). *Human Control of Multiple Unmanned Vehicles: Effects of Interface Type on Execution and Task Switching Times.* Salt Lake City: Naval Research Laboratory, Navy Center for Applied Research in Artificial Intelligence.

Tharanathan, A., Bullemer, P., Laberge, J., Reising, D. V., & McLain, R. (2012). Impact of Functional versus Schematic Overview Displays on Console Operators' Situation Awareness. *Journal of Cognitivce Engineering and Decision Making, 6* (2), 141-164.

Tombu, M., & Jolicoeur, P. (2004). Virtually No Evidence for Virtually Perfect Time-Sharing. *Journal of Experimental Psychology: Human Perception and Performance, 30* (5), 795-810.

U.S. Chemical Safety and Hazard Investigation Board. (2011). *Pesticide Chemical Runaway Reaction - Pressure Vessel Explosion.* Investigation Report, No. 2008-08-I-WV, Bayer CropScience, LP, Institute, WV, August 28, 2008.

U.S. Chemical Safety and Hazard Investigation Board. (2007). *Refinery Explosion and Fire.* Investigation Report, No. 2005-04-I-TX, British Petrolium (BP), Texas City, TX, March 23, 2005.

U.S. Nuclear Regulatory Commission. (2002). *Human-system interface design review guidelines.* Brookhaven National Laboratory, Energy Science & Technology Department.

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface Design: Theoretical Foundations. *IEEE Transactions on Systems, Man, and Cybernetics, 22* (4), 589-606.

Wang, Q., Zhuang, D., Wei, H., & Wanyan, X. (2012). Evaluation Method of Cockpit Display Interface Based on Situation Awareness. *5th International Conference on BioMedical Engineering and Informatics* (pp. 528-531). Chongqing: IEEE.

Yeung, N., Nystrom, L. E., Aronson, J. A., & Cohen, J. D. (2006). Between-Task Competition and Cognitive Control in Task Switching. *The Journal of Neuroscience, 26* (5), 1429-1438.

Appendix A – Research Methodology: A Phased Approach

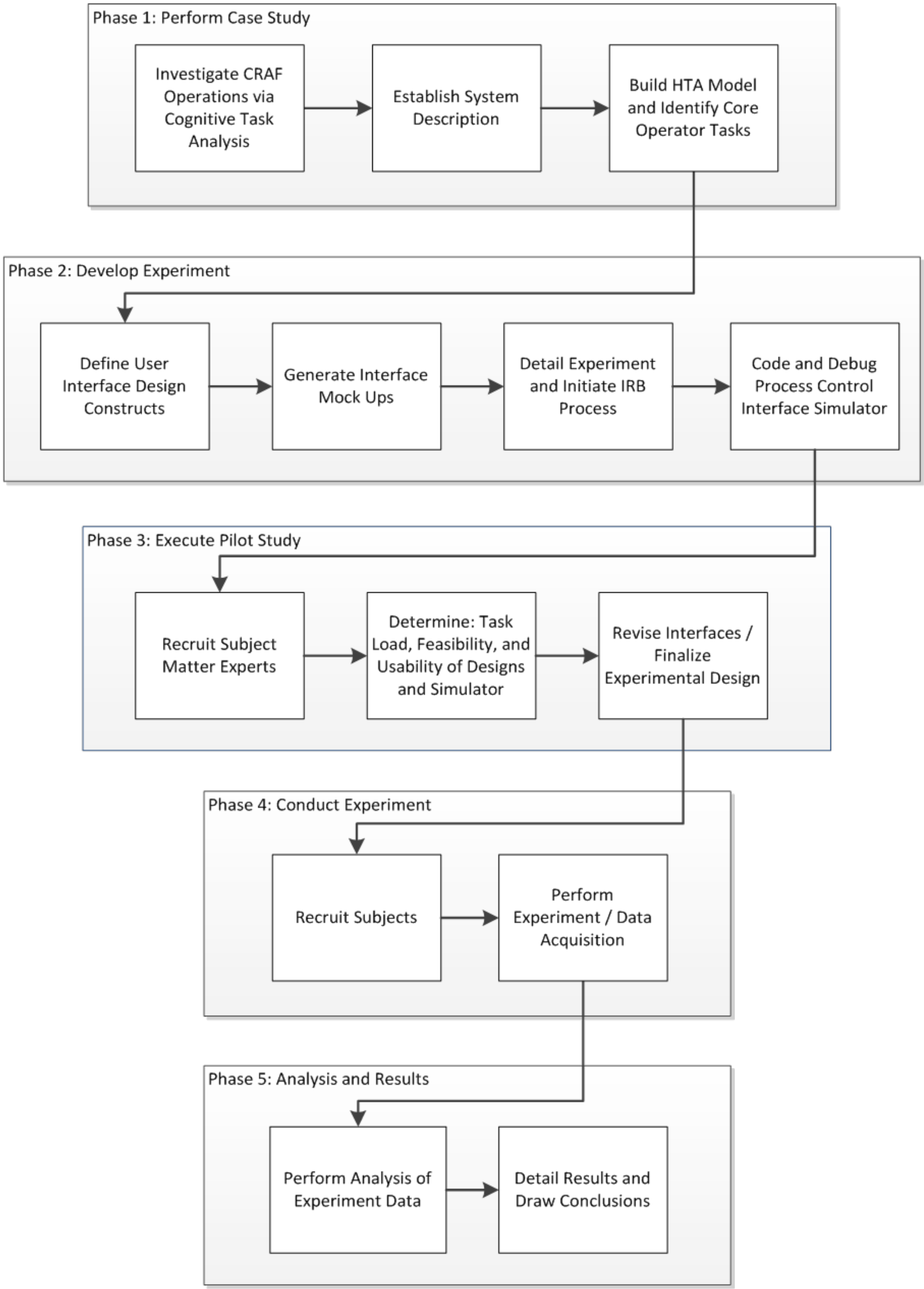


Figure A-1 Methodology Flow Chart

Introduction

This appendix covers the research methodology phases in their order of execution from start to finish. Beginning with the methodology flow chart shown above, efforts are made to provide a step-wise description of the research effort depicted throughout this document. How this research effort progressed is best understood by following the inputs and outputs of each phase used in the phased approach and further detailed below. Thus an underlying tenet of this appendix is this research effort is intended to be repeatable, either in-kind or through minor adaptation toward future efforts.

Phase 1: Perform Case Study

Overview: During phase one, a Component Research Air Facility case study and cognitive task analysis (CTA) were performed during a typical operation the evening of 09OCT13. This was done to identify existing user interface capabilities and note suitable areas for research and potential improvement. During the CTA, it became evident vigilance monitoring was a large portion of the facility operator's responsibility during continuous operations. Further, it appeared the human machine interface designs being used had a direct impact on the operators' ability to maintain SA and adequate levels of performance in a multiple process, multi-task environment. Immediately after operations that evening, crew members were polled for input about the pros and cons of the competing interface display types in formulating key decisions during the run. Operators and the facility manager were formally interviewed at a later date as well, 1-2 days later, where each was asked how they used the differing display types while performing their operational tasks. This information along with facility documentation was used to develop a formal description of the facility and operations. From this a series of task network diagrams were put together to model the operational facility and further refine operator core tasks.

Inputs: This phase required an operational facility, substantial prolonged periods of monitoring where the operators were unaware they were being directly monitored, detailed researcher observations documentation of crew interaction / equipment management, and crew member / Component Research Air Facility manager / subject matter expertise (SME) input *ex post facto*.

Outputs: CTA results to include unstructured interview responses from all operators, Hierarchal task analysis (HTA) diagrams depicting facility work flow and procedural protocol, Task network diagrams reflecting operator core tasks and key decision points, and a well documented system description were either obtained directly or generated shortly thereafter.

Phase 2: Develop Experiment

Overview: This phase represented the mechanics of actually defining and developing the formal research experiment. User interface designs were mocked up and the entire experimental design was submitted for an AFIT/AFRL internal review board (IRB) human subjects research exemption request approval.

Inputs: This phase required a process control system description for the Component Research Air Facility and conceptual interface design constructs intended to positively enhance operator SA, task performance, and workload.

Outputs: This phase was considered complete only upon successful reception of an IRB exemption approval and a fully operational interface simulation capability readied for subsequent phases.

Phase 3: Execute Pilot Study

Overview: This phase aimed to refine the experimental setup based on SME and a small contingent (n = 4) of preliminary participants. It incorporated the input of existing subject matter experts to determine the appropriate experimental task loading to mimic real world operations, feasibility of the design constructs, and captured overall interaction usability through a usability survey, Air Force Institute of Technology survey control number: 2014-04. Post pilot study, the interface designs were revised based on all forms of feedback prior to moving forward to the formal experiment.

Inputs: IRB exemption approval letter was required before this phase could commence. Participant inputs were as follows: demographics questionnaire; signed informed consent document (ICD); color blindness test results.

Outputs: Established task load and fully vetted simulator to be applied toward the formal research effort; usability survey results and suggested interface revisions, which were integrated into the simulator. A Post-experiment questionnaire that captured subject feedback.

Phase 4: Conduct Experiment

Overview: The purpose of this phase was to administer the formal experimental interface designs to as many voluntary participants as possible within a reasonable time period. A Latin Square design was utilized for counterbalancing purposes. To achieve perfect counterbalancing a 24-person participant pool was targeted.

Inputs: Test subjects and their input as follows: demographics questionnaire; signed informed consent document (ICD); color blindness test results. The Situation Awareness Global Assessment Technique (SAGAT) was utilized to measure levels 1, 2, and 3 SA;

NASA-Task Load Index (NASA-TLX) was administered to gauge user perceived workload for the competing constructs. Post-experiment questionnaire to capture subject qualitative feedback.

Outputs: Collected data relevant to SA, primary and secondary task performance, and workload to include any and all additionally captured documentation, which was used as inputs to the final analysis, and results phase that followed. Researcher observation's were recorded and catalogued to determine participant task management strategy.

Phase 5: Analysis and Results

Overview: This phase analyzed all data relevant to any and all investigative questions and contributed to the culmination of the final thesis body of work.

Inputs: Data captured from the formal experiment was necessary to complete this phase.

Outputs: Answers to the primary research questions and hypotheses. Statistics of significance for each competing design construct and task management strategy were identified.

Appendix B – Case Study: Component Research Air Facility

Overview

This appendix provides additional information about the Air Force Research Laboratory (AFRL) Component Research Air Facility (CRAF) operated by the Aerospace Systems Directorate.

Background

A suitable location representative of the process control industry had to be identified before research could commence. The successful candidate site needed to utilize a central control room operational philosophy, have sufficient means of automation and user interfaces available, have varying levels of experienced operators, and most importantly be representative of many other industrialized facilities using a similar construct. These criteria were deemed necessary to be able to generalize any findings of the research effort toward a larger subset of the process control industry. The Component Research Air Facility at Wright-Patterson AFB, OH (WPAFB, OH) was selected due to its scale and complexity corollaries to a wide array of industrial process control facilities worldwide. The facility is where an initial case study and cognitive task analysis (CTA) were executed. Insights gained from the facility while engaged in an operational test in the fall of 2013 formed the basis of the accompanying research effort.

Overview of System under Investigation[1]

The primary purpose of the Component Research Air Facility, in brief: is to provide AFRL programs of record throughout the laboratory's propulsion complex with resources necessary to conduct component level testing of turbine engines, general propulsion systems and subsystems, and fuels and combustion research. The facility is an integral part of the research and development (R&D) efforts being conducted at AFRL. Its primary use is to simulate flight conditions by providing process related resources to all facility interconnected research areas. The Component Research Air Facility gives researchers the ability to simulate actual airborne flight conditions without ever leaving the ground. This is done chiefly in direct support of U.S. Government and DoD contracted research efforts, but is additionally a dual use facility supporting both defense and private industry interests in advancing all forms of propulsion research relative to flight.

Despite the extremely unique nature of the Component Research Air Facility's primary purpose, it like many other facilities is heavily reliant upon user interfaces to aid operator SA during all aspects of operation and maintenance. There are a suite of automation controls to include programmable logic controllers (PLC), human-machine interfaces (HMI), data acquisition (DAQ) systems, and instrumentation distributed throughout the facility. The facility's control philosophy centers mainly on reliability and accuracy of resource delivery to the research test articles under the vigilance oversight of human operators located in a remote central control room. Operators are

[1] *All information relating to AFRL Aerospace Systems Directorate, CRAF Description and Purpose derived from the public domain. Information herein is Distribution A, Cleared for Public Release IAW 88ABW-2013-1629*

therefore required to perform monitoring tasks using their preference of a set of numerical and graphical user interfaces in an environment that also houses a traditional panel board display. Some information is redundant across the differing user interfaces and panel boards, but some only exists in one location or another. As unusual this may seem upon first encounter, it's typical of many process control environments that have evolved over a period of time relative to advances in automation. As new technology is integrated, often times existing technologies are never displaced. In this regard, the Component Research Air Facility is as close a representation to many industries utilizing a central control room construct that was readily accessible and had a suitable level of access, user interfaces, process control instrumentation, and automation available.

The Component Research Air Facility houses numerous pieces of large capital equipment that constitute a combined total in excess of 20,000 horsepower worth of machinery and associate subsystems. To provide simulated turbine engine inlet air to research areas there are several large air compressors, three reciprocating types provide a total of 7.5 pound-mass per second (lbm/sec) of air at 315 pounds per square inch absolute (psia) and two centrifugal type compressors providing for a total of 30 lbm/sec of air at 750 psia. An in-line, indirect fired process air heater gives the facility the capability to heat incoming inlet air to the test areas continuously from 250 to 1150 degrees Fahrenheit (degF). Turbine engine exhaust and inlet testing suction is provided by way of four turbo-exhausters each having an ability to pull 36,000 cubic feet per minute (cfm) at an absolute pressure of 11 inches of mercury (in/Hg), simulating 25,000 feet (ft) of altitude at near sea level. Process control changes to the configuration of all facility exhaust systems can provide flow rates and pressures from 36,000 cfm at 4 in/Hg to 75,000 cfm at 11 in/Hg, simulating any point from 25,000 to 46,000 ft of altitude, respectively. (Air Force Research Laboratory / Aerospace Systems Directorate, 2014)

Facility Architecture

Prior to the CTA a formal description of the Component Research Air Facility architecture was developed to determine where time would best be focused for the larger investigative work. The resultant architecture is represented graphically in Figure B-1 below. Note that each task has many associate subtasks necessary to achieve the overarching "Run Facility" goal and several are interrelated to one another.

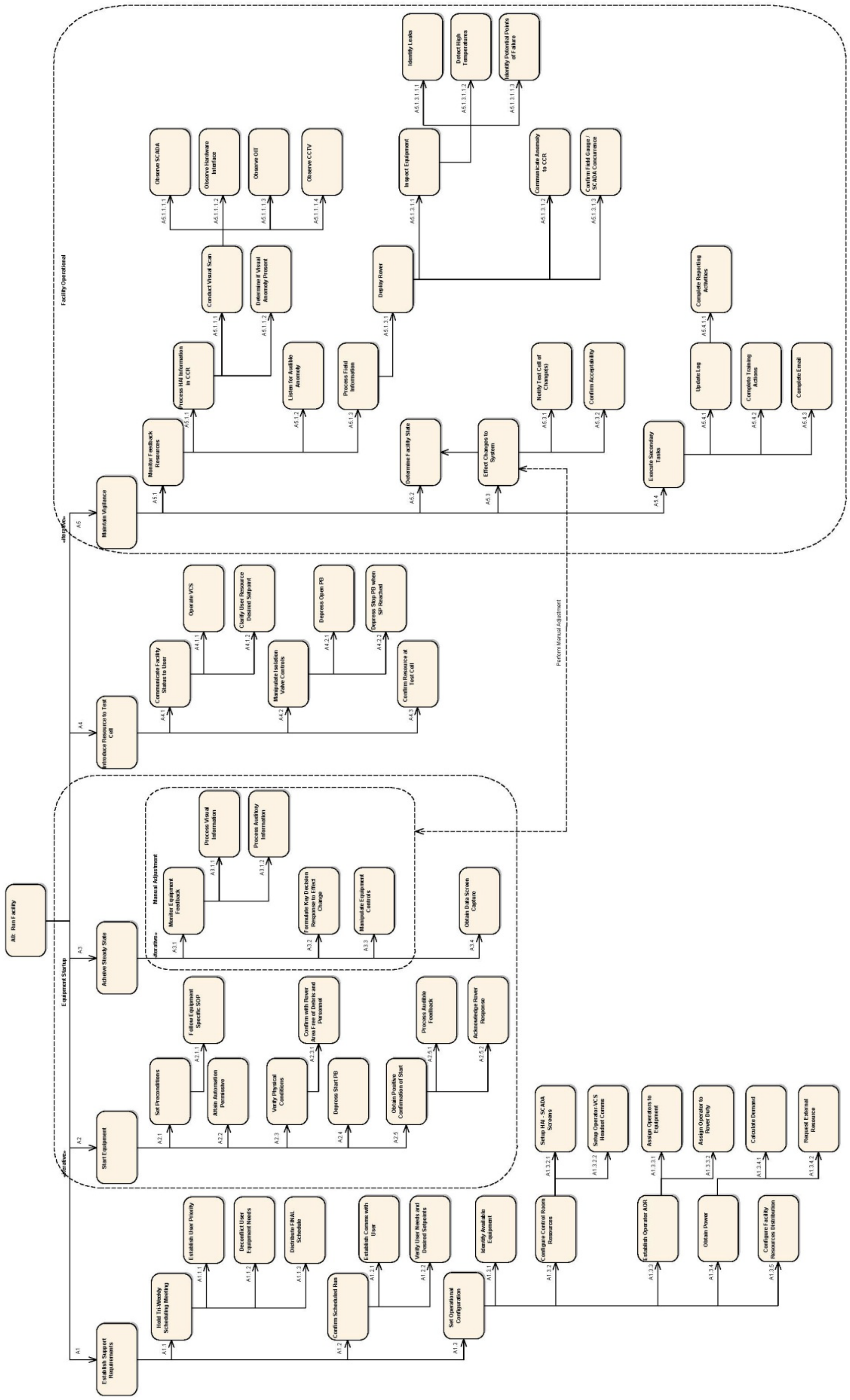


Figure B-1 CRAF Facility Architecture Diagram

Shown in Figure B-2 below A0, Run Facility there are five subordinate functions: A1, Establish Support Requirements; A2, Start Equipment; A3, Achieve Steady State; A4, Introduce Resource to Test Cell; A5, Maintain Vigilance. The facility is operational for the longest duration of time within A5, Maintain Vigilance, which is detailed in Figure B-3.

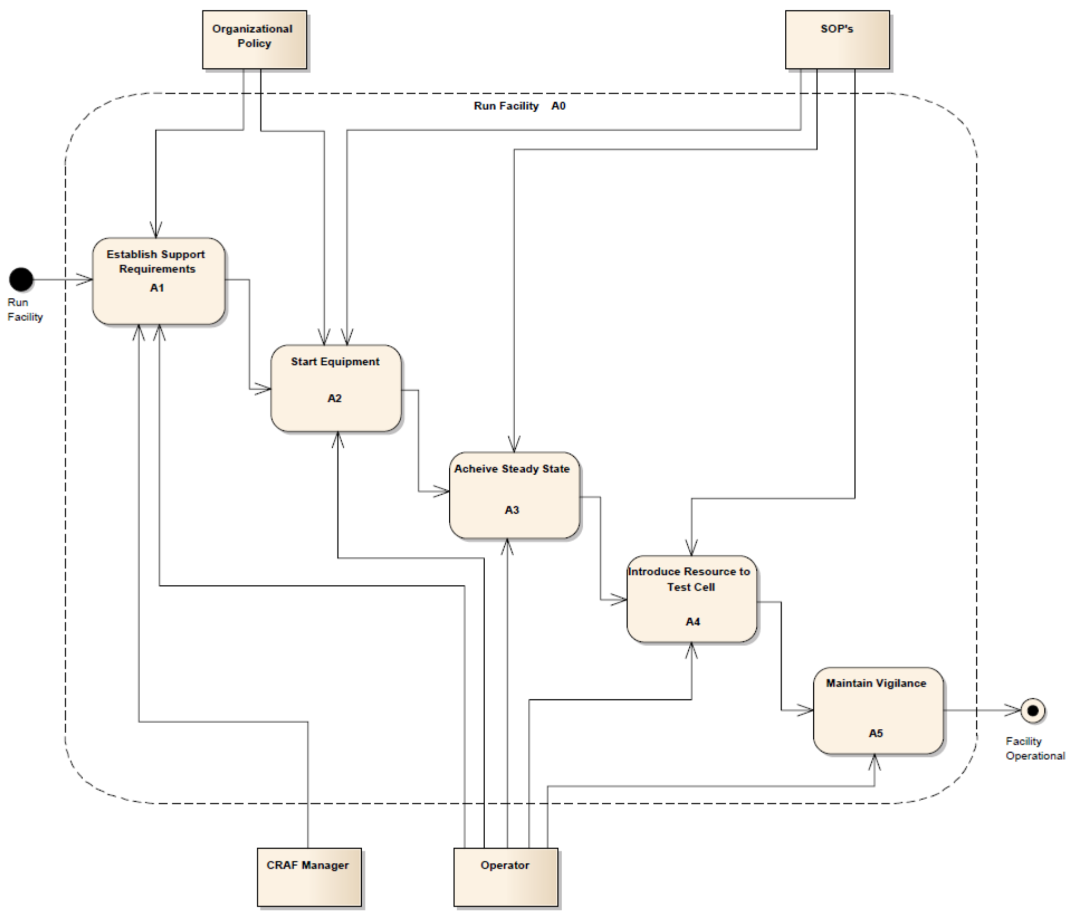


Figure B-2 CRAF A0, Run Facility Diagram

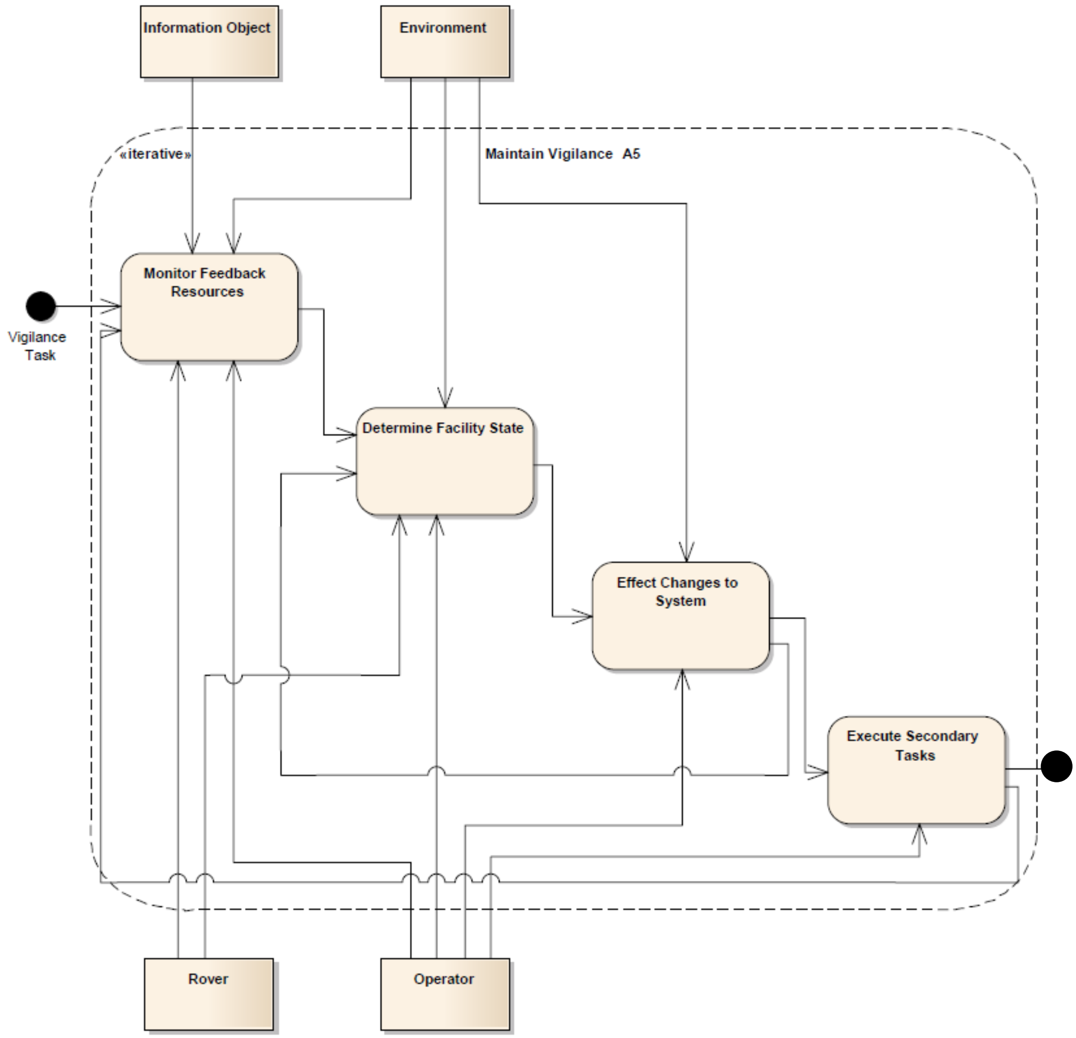


Figure B-3 CRAF A5, Maintain Vigilance Diagram

Cognitive Task Analysis

During the CTA, it became evident vigilance monitoring was a large portion of the operator's responsibility during continuous operations. The primary monitoring activity was directly attributable to the maintenance of situation awareness (SA) for the operators to stay abreast of system status throughout the facility. Two primary competing user interface display types were already present in the central control room, the first being described as a numerical display containing numbered data arranged in a functional grouping as shown in Figure B-4. The other display was described as a spatially mapped type containing a mimic representation of the process under control with numeric data depicted on the screen in its actual location relative to the equipment, as best as can be represented on a two-dimensional display. The spatially mapped type display is shown in Figure B-5. Notable for both figures was that they represent two alternative means of displaying the same information for the same piece of equipment. Some operators had a strong preference for one type over the other.

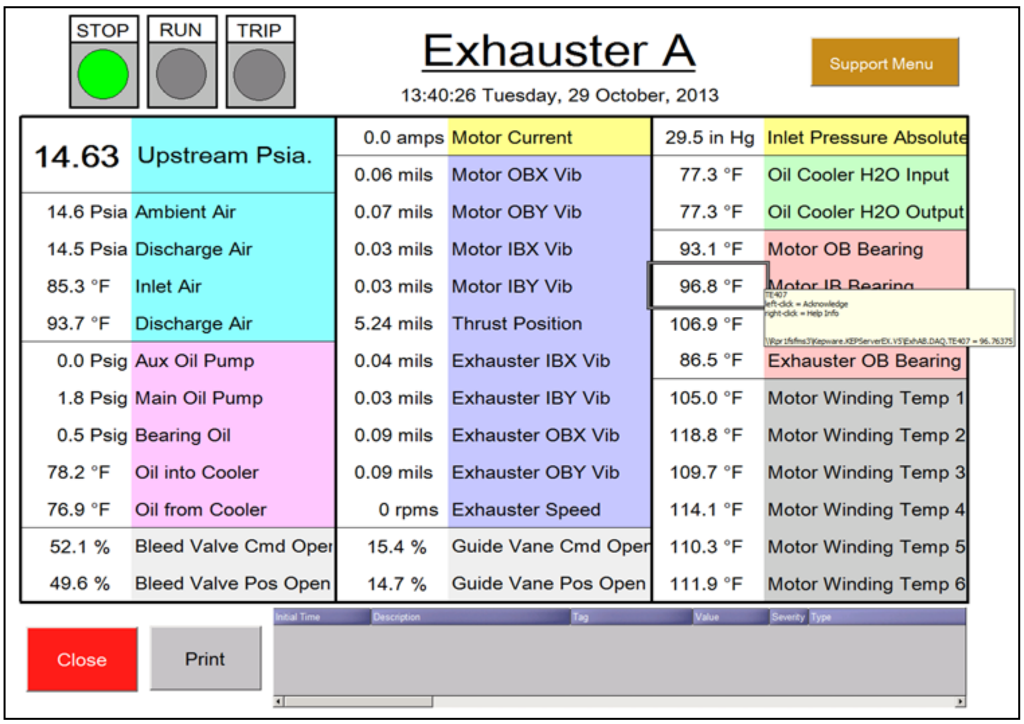


Figure B-4 CRAF Informative Display: Numeric and Functionally Grouped

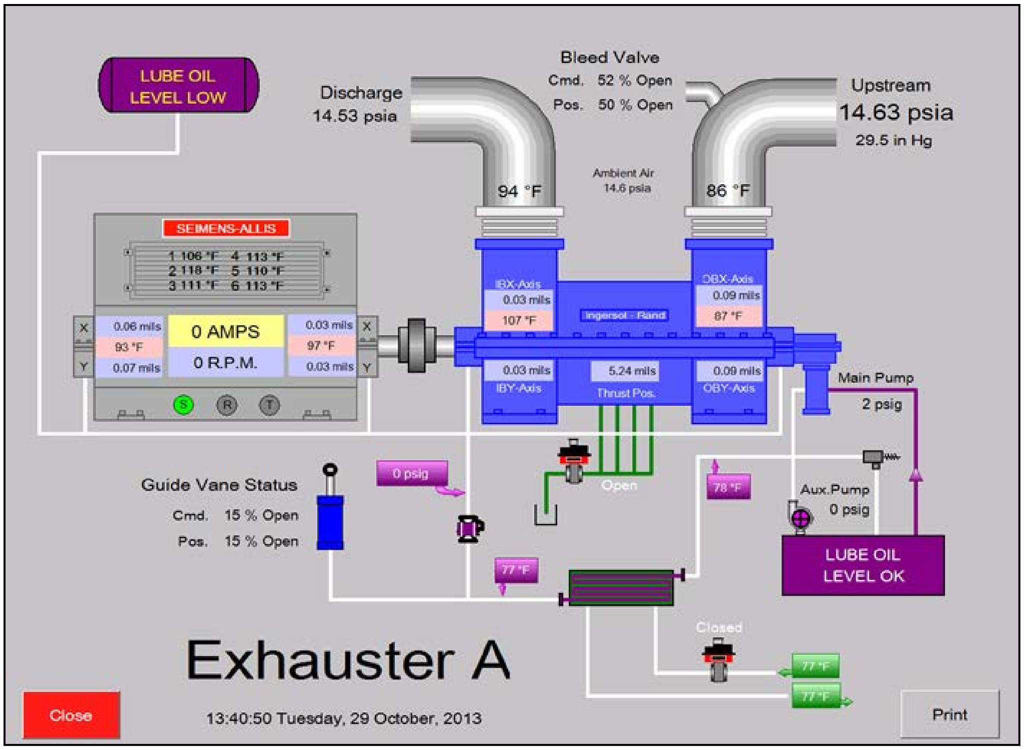


Figure B-5 CRAF Informative Display: Numeric and Spatially Mapped

The Component Research Air Facility manager and facility operations crew were observed as part of a CTA during an active operation supporting a real world test (research test article was external to the facility, but fed by facility equipment and resources). The observation period on this occasion lasted approximately 10-hours. The manager was aware of the ongoing observation and intent, however the operators were not informed they would be monitored to preclude any biases toward foreknowledge in having an observer present. All Component Research Air Facility parties involved were polled after the support effort about the pros and cons of the competing display types in formulating key decision points and maintenance of SA for process control. Operators and the manager were also formally interviewed individually at a later date, 1-2 days following the active operation, as to how each used the differing display types when the facility was operational. Over 16-pages of direct observation and 10-pages of interview information with the facility manager and operators was gathered. In addition to this, a host of original equipment manufacturers manuals, data files, and Component Research Air Facility standard operating procedures were reviewed to better understand overall system interaction and operation. The entire CTA archive shall not be presented here due to considerations of space constraints, however an excerpt from the CTA observations and subsequent responses from operator interviews to clarify a key decision point is provided in Figure B-6. For this figure: information in red reflects in-line questions noted by the researcher during the observation period to follow up on later. Italics text below each question is the detailed response from the operators involved.

Key decision point – problem at external facility has severely negative effect on CRAF equipment

- Opr2 notes two machines surging, moves rapidly to HMIs; concurrently Opr1 moves to open isolation valves
    - What cued Opr2 to act? Where was he looking? Did Opr1 react to Opr2 or otherwise?
        - *Opr2 tells me later it was the SCADA red blinking alarm at bottom of screen for high current on the machine that prompted him to act; he was monitoring this screen at the time the problem presented itself.*
        - *Opr1 states he reacted to Opr2's concern and went to hit the PB's for isolation after confirming Opr2's concern on the SCADA screens (it was at this point he noticed the alarms on the screen).*
    - Many key decision points may exist here:
        - What did Opr2 do first to the machine controls?
        - *Opr2 says later his initial thoughts were to get air into the line to keep the machines from starving for air. He accomplished this by running the bleed valves (located below the HMIs) to move the inlet bleeds open*

Figure B-6 CTA Observation of Key Decision Point and Operator Responses

Task Networks

Post CTA four operator task networks were developed to support the experimental design. Each is reflective of the operator's activity and contributed directly to the experimental design and simulator coding. All task networks are shown below.

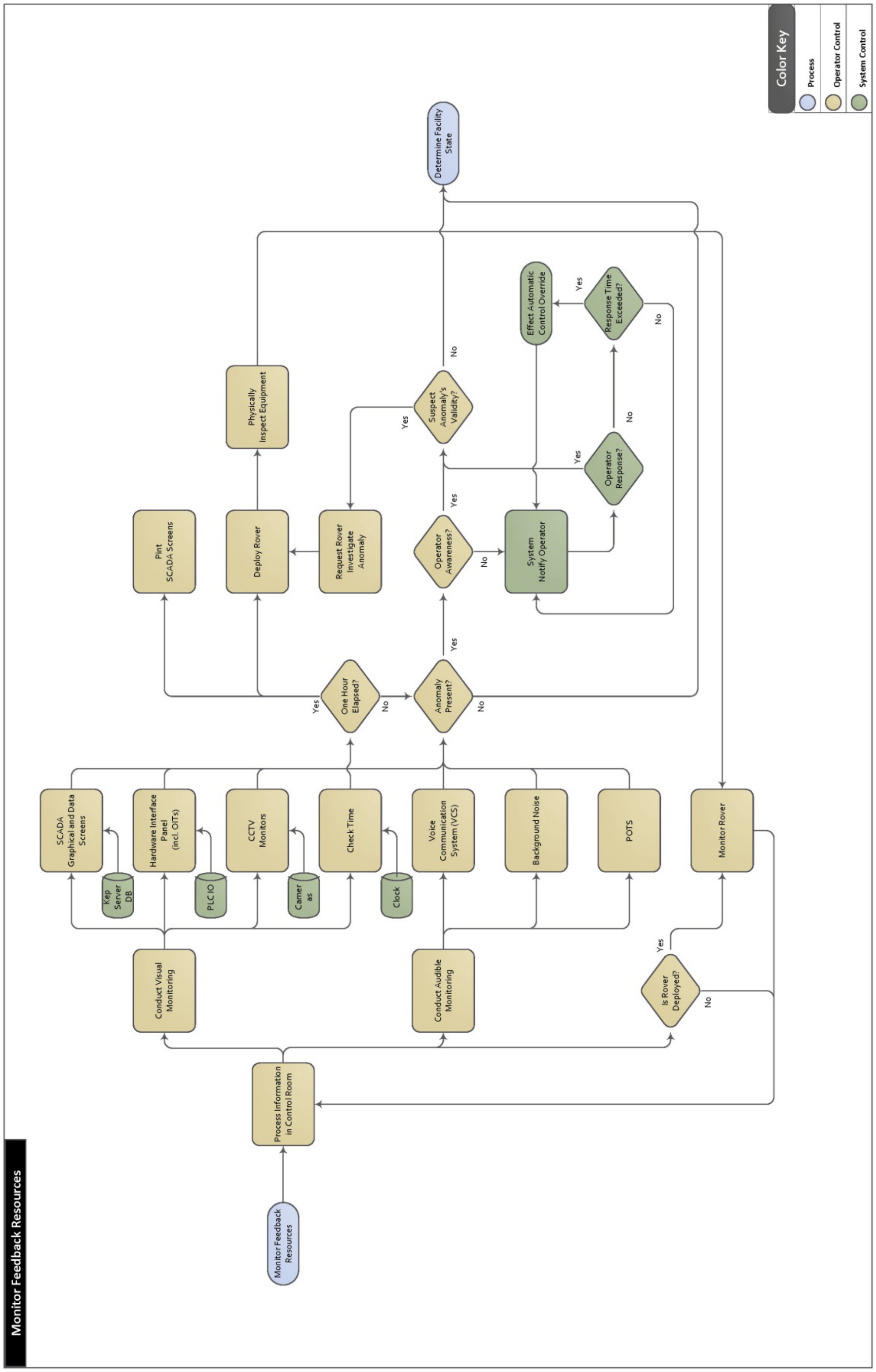


Figure B-7 Task Network: Monitor Feedback Resources

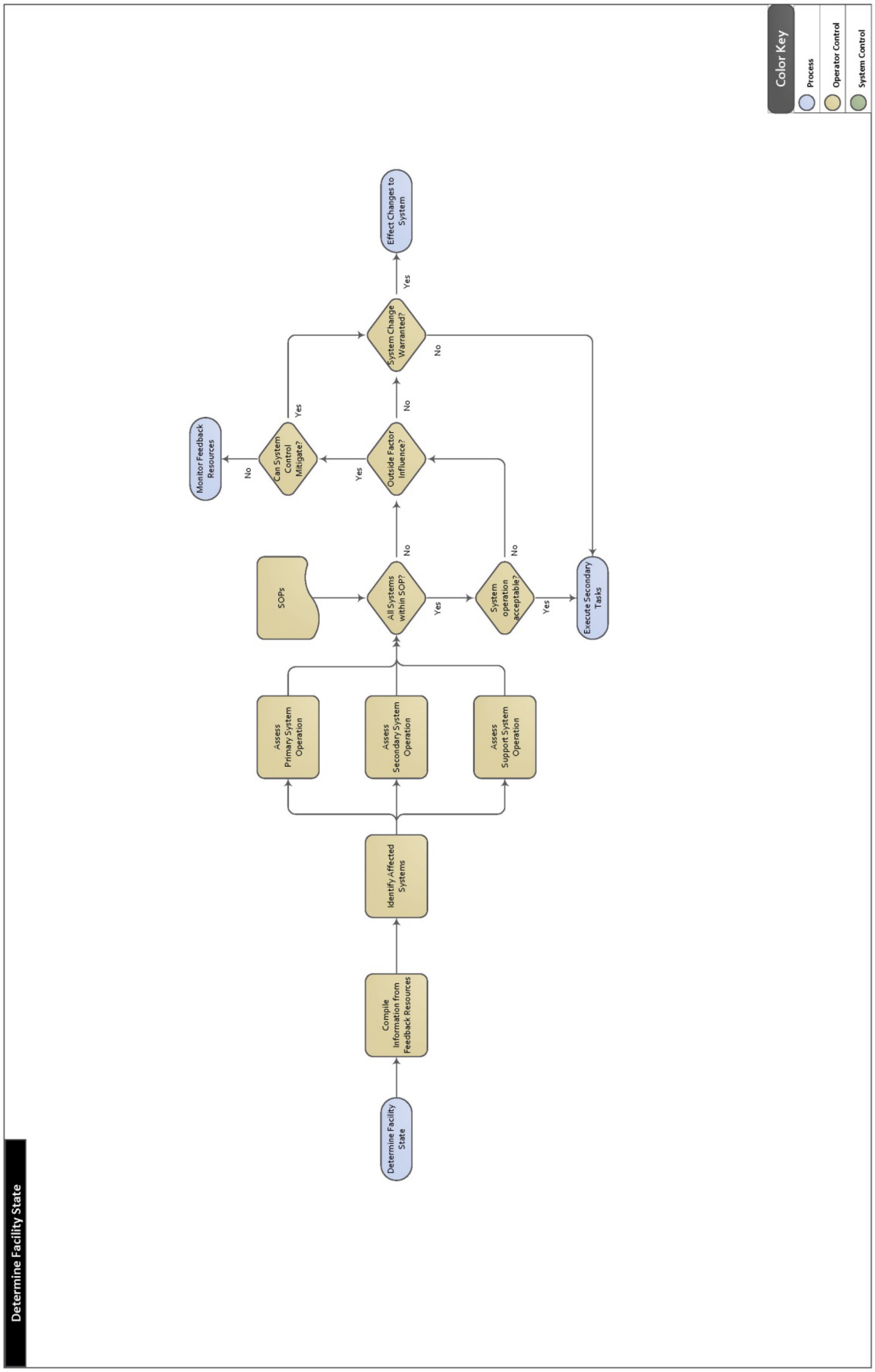


Figure B-8 Task Network: Determine Facility State

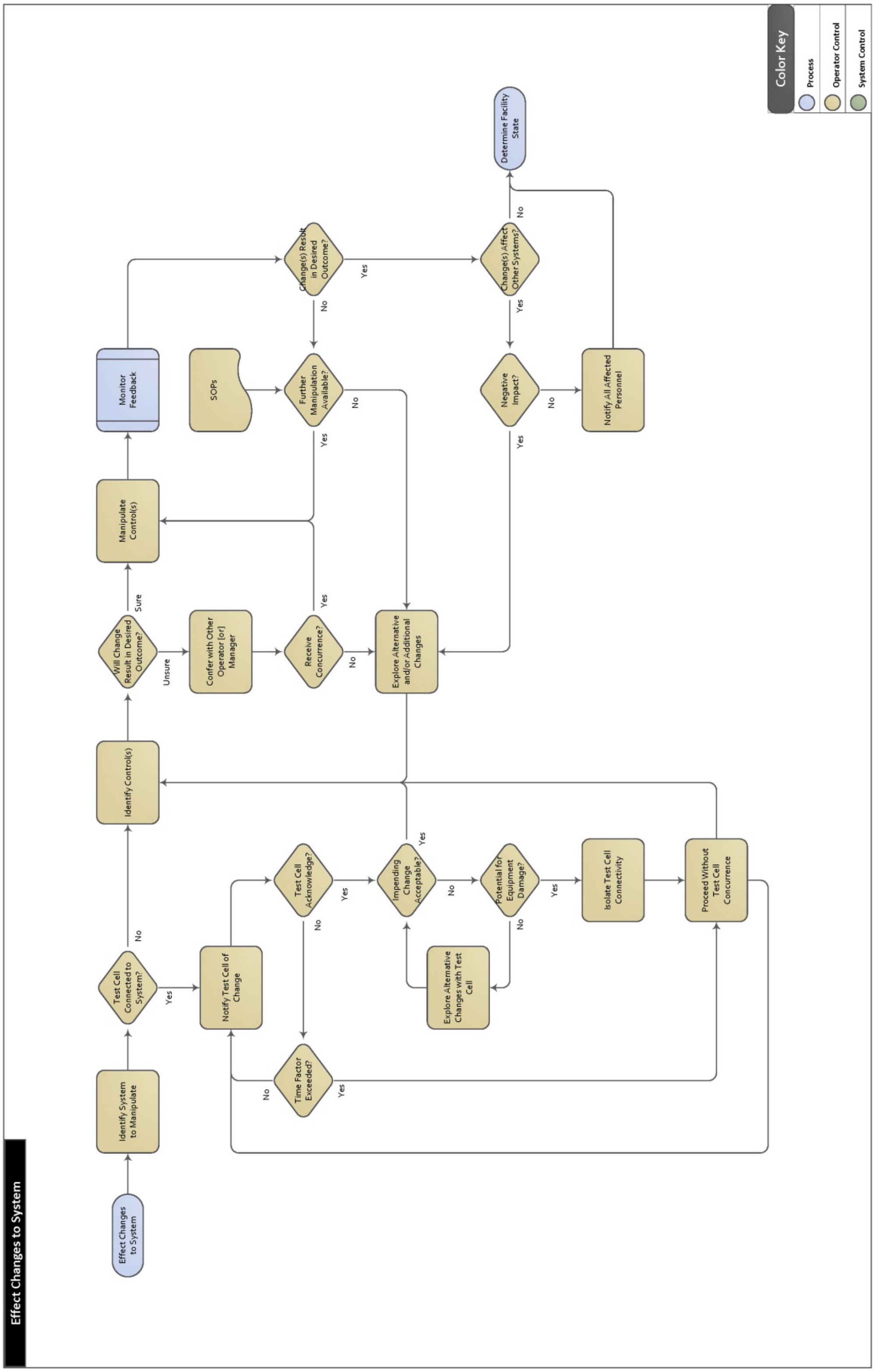


Figure B-9 Task Network: Effect Changes to System

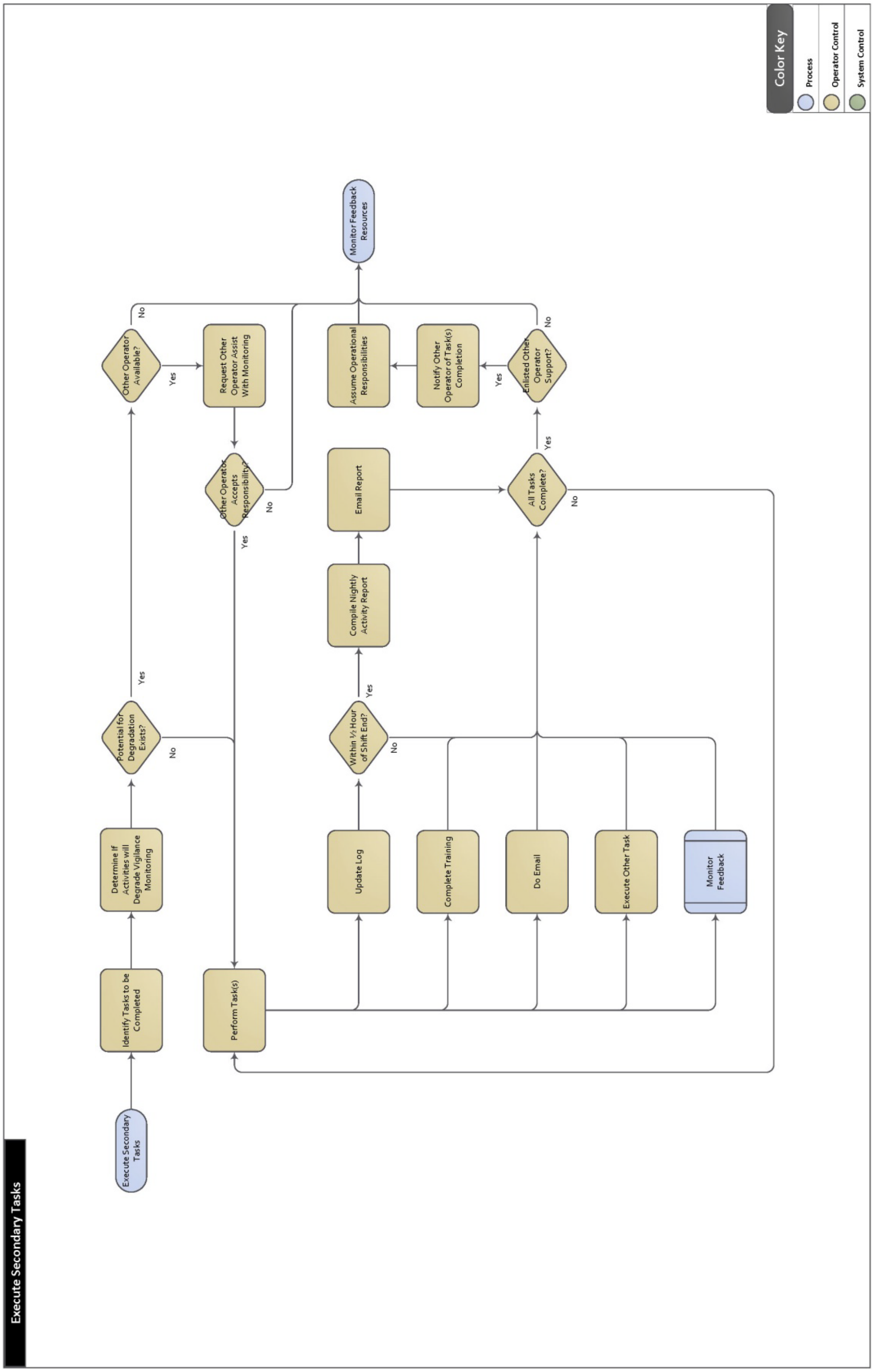


Figure B-10 Task Network: Execute Secondary Tasks

Appendix C – AFIT IRB Exemption Request Approval

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY (AETC)

4 Apr 2014

MEMORANDUM FOR CAPT CHRISTINA RUSNOCK

FROM: John J. Elshaw, Ph.D.
AFIT IRB Research Reviewer
2950 Hobson Way
Wright-Patterson AFB, OH 45433-7765

SUBJECT: Approval for exemption request from human experimentation requirements (32 CFR 219, DoDD 3216.2 and AFI 40-402) for Display Design Construct for Process Control Monitoring.

1. Your request was based on the Code of Federal Regulations, title 32, part 219, section 101, paragraph (b) (2) Research activities that involve the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior unless: (i) Information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) Any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

2. Your study qualifies for this exemption because you are not collecting sensitive data, which could reasonably damage the subjects' financial standing, employability, or reputation. Further, the demographic data you are collecting cannot realistically be expected to map a given response to a specific subject.

3. This determination pertains only to the Federal, Department of Defense, and Air Force regulations that govern the use of human subjects in research. Further, if a subject's future response reasonably places them at risk of criminal or civil liability or is damaging to their financial standing, employability, or reputation, you are required to file an adverse event report with this office immediately.

4/7/2014

X

JOHN J. ELSHAW, PH.D.
AFIT RESEARCH REVIEWER

Appendix D – Experiment Documents

Experiment Checklist

Prior to Participant Arrival

Before Leaving Office

- ☐ Print out all requisite docs and handouts
- ☐ Assign a Participant ID and write this in on all relevant docs
- ☐ Determine Latin Square order for this Participant
  - o Organize all docs in the pre-determined order
- ☐ Take pens, pencils, and any other supplies as needed

In Experiment Room

- ☐ Lay out all docs in a manner conducive to presentation during the training and trials
- ☐ Start Concept on laptop and go online with PLC
  - o Animate the reference data editor
    - ▪ Monitor FBD's: Trial_Training [and] Trial Control (verify ALL timers are reset)
- ☐ On Desktop PC start Experiment Training PPT and GraphWorx Training Menu
  - o Put the PPT in presentation mode
- ☐ Turn off both displays to prevent burn in
- ☐ Unplug Phone in room
- ☐ Turn off your cell phone

Upon Participant Arrival

Initial Encounter in Experiment Room

- ☐ Hang "TEST IN PROGRESS" sign on and lock door to prevent inadvertent entry from outside
- ☐ Discuss facility protocol, emergency egress procedures, and nearest exits
  - o Cover the fact this is an operational facility and equipment can start at any time
  - o Provide hearing protection, if necessary
- ☐ Ask if Participant needs to use the restroom
  - o State the following: "We'll be in this room for approximately 3 hours, bathroom breaks may be taken in between trials but are discouraged."
  - o There will be a bathroom / stretch break at the end of the training session
- ☐ Ask if Participant has any special accommodations or needs that may preclude them from completing the experiment in the estimated 3 hour timeframe
- ☐ Request Participant turn off all electronic devices and stow them in the locker provided
  - o Also remove any watches or other timing devices – being sure to silence all items that may cause an interruption
- ☐ **Administer:**
  - o **Informed Consent Document**
  - o **Demographics Questionnaire**
    - ▪ Be sure to review responses / ask for clarification before going any further

- **Color Blindness Test**
  - Make note of Participant results (for each plate) in the log

Begin Experiment Phase

- ☐ Conduct Training
  - Turn on Displays and go through PPT with Participant
    - Make sure Participant comprehends the training material
  - Be sure to offer multiple opportunities to ask questions
  - Conclude training by offering additional training modules, if they so desire
- ☐ Scrub PLC registers by kicking reset bit ON/OFF; also "Reset_All_Timers" bit, if nec.
- ☐ Begin Experimental Trials
  - Monitor timer values to ready for SAGAT pauses
    - Note: *Be conscious of not making noises (or lack thereof) to allow the Participant to predict when the pause is going to occur*
- ☐ **Administer:**
  - **SAGAT Queries during Pauses**
  - **TLX at conclusion of each trial**
    - BEGIN NEXT TRIAL
      - *Pay mind to Latin Square order!*
- ☐ Perform data entry on SAGAT returns and trial tests (when available)
  - Be sure to add participant ID to all paper copies they turn in
- ☐ Update log and observation document as the trials progress
- ☐ Conclude Experiment

Post Experiment

- ☐ **Administer:**
  - **Post Experiment Questionnaire**
  - **Usability Survey (Pilot Study Only)**
    - Inquire about everything to include questionnaires, surveys, etc.
- ☐ Review administered documents for completion
  - Request clarification where necessary
- ☐ Provide a copy of the ICD to the participant
  - Be sure to put their Participant ID on their copy only
- ☐ Plug the landline telephone back in
- ☐ Remind the Participant to collect their belongings from the locker
- ☐ Ensure / Remind Participant NOT to discuss the experiment with anyone as this could preclude them from participation in the future
- ☐ Thank the Participant again and escort to the exit

Informed Consent Document

Greetings! You are being asked to take part in a research study carried out by Dr. Christina Rusnock and James Bowden, Air Force Institute of Technology / Systems Engineering and Management (AFIT/ENV). This form explains the study and your part in it if you decide to join. Please read the form carefully; take as much time as desired. Ask the researcher to explain anything you do not understand. You can decide not to join the study. If you do join the study, you can change your mind later or quit at any time without any penalty or loss of services or benefits.

Study Title: Display Design Construct for Process Control Monitoring

Primary Researchers:

Name	Title/Department	E-mail	Telephone
Christina Rusnock, PhD	Assistant Professor of Systems Engineering, AFIT/ENV	christina.rusnock@afit.edu	937.255.3636, x4611
James Bowden	Master's Student in Systems Engineering, AFIT/ENV	james.bowden.4@us.af.mil	937.255.3101

What is this study about? The purpose of this study is to evaluate different ways of displaying information on a computer screen. The objective is to help find the best way of displaying process control information that improves a user's situation awareness (SA) and task performance during a process control monitoring activity that requires human monitoring. You are being asked to take part in this study in order to collect data to meet this objective. Depending on the time it takes to complete all of the study's activities, taking part will require about 120 minutes of your time – give or take 15 minutes.

What will I be asked to do if I am in this study? If you take part in the study, you will participate in four 15 to 30-minute computer monitoring sessions using different display types during each session. In addition to monitoring the computer displays you are also being asked to complete a separate reading comprehension task at the same time you are engaged in the monitoring activity. For all sessions there will be pauses at predetermined times so the researcher can ask you questions either automatically through the computer interface or face-to-face about what is happening on any of the displays. Answers will be treated by the researcher as For Official Use Only (FOUO) and not shared with anyone outside of the research effort for not only these but all information you choose to provide as a part of this study.

Are there any benefits to me if I am in this study? You are not expected to benefit directly from participation in this research study, however the participant who performs the best for all tasks combined will be awarded a $25 gift card. Only one card will be awarded at the conclusion of the study and the winner may choose to remain anonymous. The main benefits of this study will be to help in determining how display designers can make process control displays better so operators have an increased awareness of what is going on around them. This could potentially lead to a better work environment for anyone who monitors displays and works in the process control industry.

Are there any risks to me if I am in this study? Because this research requests your subjective opinion, and will gather information about the way that you detect and perceive information, some of your responses might be considered sensitive and may cause discomfort. For this reason you may refuse to answer any question at any time, and likewise may opt out of the study at any time without question. It is also important to note that it might still be possible for a reader of the final written product to attribute results to a given individual and/or organization.

You will be given an opportunity to review information provided and make a reasoned judgment of the risks of divulging such information.

Will my information be kept private? The data for this study will be kept confidential to the extent allowed by federal and state law. No published results will identify you, and your name will not be associated with the findings. Under certain circumstances, information that identifies you may be released for internal and external reviews of this study. The digital file containing the survey and data collection results, as well as the study write-up will be secured on a password-protected computer assigned to the researcher.

Additionally, data collected may be released for future studies. If released it will be sanitized and anonymized as required to ensure that no personally identifiable information (PII) is present. This data could be used for studies that will assess your SA abilities. Again, if this takes place, the data will be anonymized as previously stated.

Your information will only be released, if requested, to authorized members of the AFIT Institutional Review Board (IRB), to ensure research compliance with federal and state law. Your information will not be released to any other entity. The results of this study may be published or presented at professional meetings, but the identities of all research participants will remain anonymous. The data for this study will be kept for three years, as required by AFIT policy, after which time the digital file containing all personal data will be destroyed and all remaining data will remain completely anonymized.

Are there any costs or payments for being in this study? There will be no costs or payments to you for taking part in this study.

Who can I talk to if I have questions? If you have questions about this study or the information in this form, please contact the researcher using the contact information provided above. If you have questions about your rights as a research participant, or would like to report a concern or complaint about this study, please contact the WPAFB Institutional Review Board at (937) 255-3636, x4543 or e-mail HumanSubjects@afit.edu, or regular mail at: Wright Research Site IRB, 711 HPW/IR, 2245 Monahan Way, Wright-Patterson AFB, OH 45433

What are my rights as a research study volunteer? Your participation in this research study is completely voluntary. You may choose not to be a part of this study. There will be no penalty to you if you choose not to take part. You may choose not to answer specific questions or to stop participating at any time.

What does my signature on this consent form mean? Your signature on this form means that: a) you understand the information given to you, b) you have been able to ask the researcher questions and state any concerns, c) the researcher has responded to your questions and concerns, d) you believe you understand the research study and the potential benefits and risks involved.

Statement of Consent: I give my voluntary consent to take part in this study. I will be given a copy of this consent document for my records.

Signature of Participant	Printed Name of Participant	Date

Statement of Person Obtaining Informed Consent: I carefully explained to the person taking part in the study what he or she can expect. I certify that when this person signs this form, to the best of my knowledge, he or she understands the purpose, procedures, potential benefits, and potential risks of participation. I also certify that he or she: a) speaks the language used to explain this research, b) reads well enough to understand this form, c) does not have any problems that could make it hard to understand what it means to take part in this research.

Signature of Researcher	Printed Name of Researcher	Date

Version: March 2014 Page 2 of 2

Demographics Questionnaire

Demographics Questionnaire

Participant # ______________ Date ______________ Age ______________ Gender _____

Education Level ____________________ Occupation ____________________

1. Where do you currently use a computer? (*Circle all that apply*)

   Home Work Other______________ Do Not Use (*Skip to question #6*)

2. What type of computer do you use? (*Circle all that apply*)

   Desktop Laptop Other______________

3. How many hours per day on average do you spend using a computer? ______________

4. In a typical day how many hours do you spend using a computer for:

   ___Web Browsing ___Social Networking ___Email ___Typing Documents

   ___Editing Photos ___Playing Games ___Data Entry ___ Other ______________

5. Which of the following best describes your expertise with computers? (check √ one)

   _____ Novice
   _____ Good with one type of software package (such as word processing or slides)
   _____ Good with several software packages
   _____ Can program in one language and use several software packages
   _____ Can program in several languages and use several software packages

6. What type of electronic device (not a computer) do you use the most? ______________

   a. How many hours per day do you spend specifically using this type device? ______________

7. How many hours per day do you watch a television? ______________

8. What type of programming do you primarily watch on television? ______________

9. For each of the following questions, indicate a time in hours to describe the longest period of time you could comfortably spend:

   Watching television_____ Browsing the Web_____

   Playing Video Games_____ Typing a document_____

   Sitting in one spot_____ Looking at a computer display_____

   Reading a book_____ Talking on a phone_____

10. For each of the following questions, circle the one response that **best** describes you.

    How often do you:

Use a mouse?	Daily, Weekly, Monthly, Once every few months, Rarely, Never
Use a touch screen?	Daily, Weekly, Monthly, Once every few months, Rarely, Never
Use a touchpad?	Daily, Weekly, Monthly, Once every few months, Rarely, Never
Use software that primarily displays data (e.g. spreadsheets)?	Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use software that primarily displays graphics (e.g. photo editing)?
Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use a smart device (e.g. smart phone, tablet)?
Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use E-mail? Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use a computer to create a document?
Daily, Weekly, Monthly, Once every few months, Rarely, Never

Use an electronic device (not a computer) to type?
Daily, Weekly, Monthly, Once every few months, Rarely, Never

11. Do you play computer and/or video games? (Circle One) YES NO (*Skip to question #13*)

12. Which type(s) of computer/video games do you play most often? (*Circle all that apply*)

First person shooter Real time strategy Casual games Other_______________

13. Are you in your usual state of health, physically? YES NO
If NO, please briefly explain:

14. How many hours of sleep did you get last night? _______________

15. What hand are you? LEFT RIGHT BOTH

   a. For the following please indicate with an "L" or "R" what hand you primarily use to:

   ___Write with ___Throw a ball ___Use a computer mouse

   b. (**LEFT HANDED ONLY**) Please briefly explain how you are most comfortable using a computer mouse (i.e. hold mouse in left hand with button swap feature; hold mouse in right hand without button swap, etc.) Please let an experimenter know if you require any special accommodation

16. Do you have any physical impairment(s) affecting your vision, dexterity, and/or general ability to use a computer? If yes to any, please provide details below your response.

Color blindness?	YES	NO
Eyesight / corrected vision?	YES	NO
Difficulty viewing computer displays?	YES	NO
Difficulty using input devices (mouse, keyboard)?	YES	NO
Other(s) – not specified above?	YES	NO

17. Have you had any previous computer programming experience?

YES NO

If yes, please briefly detail your experience level to include number of years:

18. Do you have any experience using a computer for industrial process control operations?

YES NO

If yes, in what capacity and for how long:

Operator Maintainer Both ________ years

Did or does this experience involve long periods of monitoring the process using a computer display? If so, during a normal shift (8-hours) how much time did or do you spend monitoring the process via a computer display?

________ hours / shift

19. Do you have any experience with any type of monitoring activity?

YES NO

If yes, in what capacity and for how long:

Capacity:__________________________ ________ years

Did or does this experience involve long periods of monitoring using a computer display? If so, during a normal monitoring period how much time did or do you spend monitoring the display at one time?

________ hours

SAGAT Questions – (Trial #, Pause #)

T1P1
Did any of the processes experience a deviation? Yes No
For the last deviation (if any), was this the first time for that type (e.g. temp)? Yes No N/A
The next deviation will be: High Low
Name at least two related characters, places, or things in any of the reading task excerpts:

T1P2
Specifically which process number(s) experienced a deviation since the last pause? N 1 2 3 4 5 6 7 8
If you saw deviation(s) since last pause was this the first time for that type (e.g. temp)? Yes No N/A
Will the process rank ordering from best to worst change? Yes No
What types of deviation(s) have you seen since last pause? Temp Press Vibe Flow N/A

T1P3
Name the two process numbers that are running: Best _____ , _____ Worst _____ , _____
Indicate a process number that has **never** experienced a deviation (List no more than 2): _____ , _____
Which process number will experience a deviation next (List no more than 2)? _____ , _____
Will you have the reading comprehension task completely finished by the end of the trial? Yes No

T1P4
How many deviations total have there been since last pause? _____
If there was/were deviation(s) since last pause, were these: High Low Both N/A
Name the color associated with “vibration”: Pink Teal Gray Violet
What is the plot associated with any one of the reading task excerpts?

T2P1
Name the two process numbers that are running: Best _____ , _____ Worst _____ , _____
For the last deviation (if any), was this the first time for that type (e.g. vibe)? Yes No N/A
Will the process rank order from best to worst change? Yes No
What types of deviation(s) have you seen thus far? Temp Press Vibe Flow N/A

T2P2
How many deviations total have there been since last pause? _____
If you saw deviation(s) since last pause was this the first time for that type (e.g. temp)? Yes No N/A
Which process number(s) will experience deviation(s) next (List no more than 2)? _____ , _____
Name at least two related characters, places, or things in any of the reading task excerpts:

T2P3
If there was/were deviation(s) since last pause, were these: Left of Center Right of Center Both N/A
The next deviation will be: Left of Center Right of Center
Indicate a process number that has **never** experienced a deviation (List no more than 2): _____ , _____
Will you have the reading comprehension task completed by the end of this trial? Yes No

T2P4
Name the color associated with “pressure”: Pink Teal Gray Violet
Did any of the processes experience a deviation since last pause? Yes No
Specifically which process number(s) experienced a deviation since the last pause? N 1 2 3 4 5 6 7 8
What is the plot associated with any one of the reading excerpts?

T3P1

For the last deviation (if any), was this the first time for that type (e.g. flow)? Yes No N/A
Which process <u>number</u> will experience a deviation next (List no more than 2)? _____ , _____
Name at least two related characters, places, or things in any of the reading task excerpts:
What types of deviation(s) have you seen since this trial began? Temp Press Vibe Flow N/A

T3P2

Specifically which process number(s) experienced a deviation since the last pause? N 1 2 3 4 5 6 7 8
How many deviations total have there been since last pause? _____
Will you have the reading comprehension task completed by the end of the trial? Yes No
Indicate a process <u>number</u> that has **never** experienced a deviation (List no more than 2): _____ , _____

T3P3

If there was/were deviation(s) since last pause, were these: High Low Both N/A
The next deviation be: High Low
Did any of the processes experience a deviation since last pause? Yes No
Will the process rank order from best to worst change? Yes No

T3P4

If you saw deviation(s) since last pause was this the first time for that type (e.g. flow)? Yes No N/A
Name the two process numbers that are running: Best _____ , _____ Worst _____ , _____
Name the color associated with “flow”: Pink Teal Gray Violet
What is the plot associated with any one of the reading excerpts?

T4P1

How many deviations total have there been for this trial? _____
Specifically which process number(s) experienced a deviation since the trial began? N 1 2 3 4 5 6 7 8
If you saw deviation(s) was this the first time for that type (e.g. press)? Yes No N/A
The next deviation will be: Left of Center Right of Center

T4P2

Did any of the processes experience a deviation since last pause? Yes No
If you saw deviation(s) since last pause was this the first time for that type (e.g. vibe)? Yes No N/A
Will the process rank order from best to worst change? Yes No
Name at least two related characters, places, or things in any of the reading task excerpts:

T4P3

If there was/were deviation(s) since last pause, were these: Left of Center Right of Center Both N/A
Will you have the reading comprehension task completed by the end of this trial? Yes No
Indicate a process <u>number</u> that has **never** experienced a deviation (List no more than 2): _____ , _____
Which process <u>number</u> will experience a deviation next (List no more than 2)? _____ , _____

T4P4

Name the two process numbers that are running: Best _____ , _____ Worst _____ , _____
What is the plot associated with any one of the reading excerpts?
What types of deviation(s) have you seen since last pause? Temp Press Vibe Flow N/A
Name the color associated with “temperature”: Pink Teal Gray Violet

Post Experiment Questionnaire

Post Experiment Questionnaire

Participant Number (Assigned By Researcher):

1. Overall, rate your performance on:

    The **process monitoring** tasks

    Excellent Very Good Good Fair Poor

    The **reading comprehension** tasks

    Excellent Very Good Good Fair Poor

    Both tasks combined

    Excellent Very Good Good Fair Poor

2. Overall, how often were you were you able to:

    Pay close attention to the **process monitoring** tasks?

    Always Very Often Sometimes Rarely Never

    Pay close attention to the **reading comprehension** tasks?

    Always Very Often Sometimes Rarely Never

    Detect process deviations?

    Always Very Often Sometimes Rarely Never

    Understand information about the processes and system?

    Always Very Often Sometimes Rarely Never

    Predict what any given process or the system was going to do next?

    Always Very Often Sometimes Rarely Never

3. Was one task (monitoring or reading) easier to accomplish than the other? If so, explain.

4. Were there times when you were bored? If so, when did this occur and what did you do (if anything) to overcome the boredom?

1

5. What is the longest period of time you would be able to perform the activity that comprised this experiment?

    a. Is there something that could be changed to increase the time you've indicated above? If so, please explain.

6. Which activity (the monitoring or reading) consumed the majority of your attentional resources and why?

7. Which of the following BEST describes the workload strategy you used to accomplish both the primary (process monitoring) and secondary (reading) tasks? (*Select only one*):

A. Attempted to switch back and forth between both tasks, applying full attention to only one task at a time.

B. Attempted to divide and balance attention between both tasks equally.

C. Attempted to complete the secondary (reading) task as quickly as possible so as to devote full attention to the primary (process monitoring) task when done.

D. Attempted to focus attention mostly on either the primary (process monitoring) or secondary (reading) task and only worked on the other task intermittently or not at all.
NOTE: If you select D please circle which task you focused on most: monitoring [or] reading.

E. Attempted to focus most attention on the secondary task and rely heavily upon visual cues and/or pauses to prompt actions in the primary task.

F. (Other) - Please explain:

8. Did the way you executed the monitoring and reading tasks change over time?

    a. What were the main reasons for this/these change(s)?

9. For the following questions, please use the display reference guide (**circle only one** response for each question):

    Which display design yields the **best** task performance outcomes?

    A B C D

    Which display design yields the **worst** task performance outcomes?

    A B C D

    Which display design is **most** effective for monitoring multiple processes?

    A B C D

    Which display design is **least** effective for monitoring multiple processes?

    A B C D

Which display design provides for the **best overall awareness** about the individual processes and system as a whole? (**Select only one**)

A B C D

What was it about this particular display design that made you select it?

Which of the display types would prove **better for long duration monitoring**? (**Circle no more than 2**)

A B C D

Provide an overall rank order for the display designs indicating 1 for the **best** and 4 for the **worst** design type:

____A ____B ____C ____D

Post Experiment Questionnaire

10. If you could have designed a display for this study, what characteristics would it have and what would it have looked like? (Please provide a sketch if you so desire)

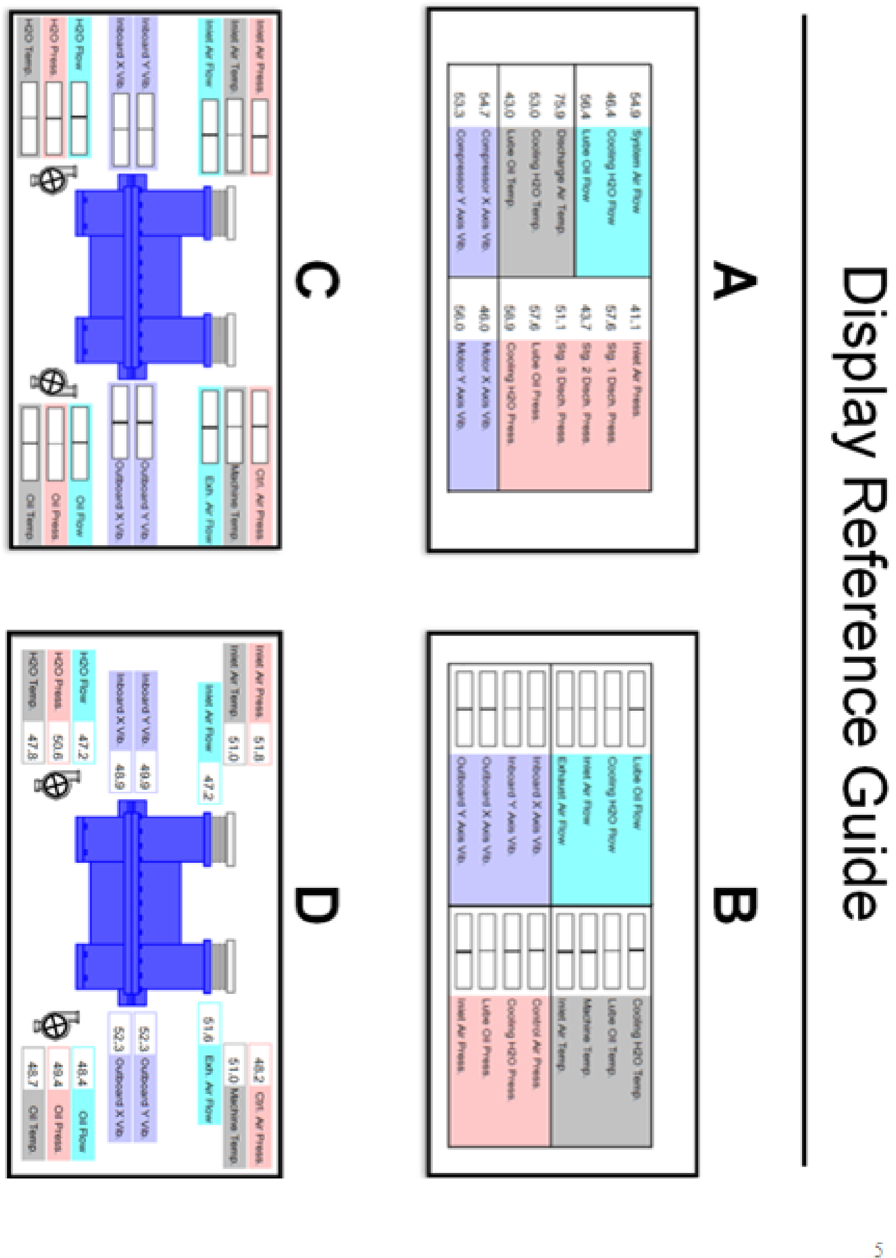
11. What was the most challenging aspect of participating in this experiment?

12. Would it have been possible to perform additional tasks above and beyond the monitoring and reading tasks done for this experiment? If so, how many more tasks and of what kind?

13. What outside factors (if any) may have influenced your participation in this study?

14. What changes would you make to the display designs and/or this study to make it better for future participants?

15. Please provide any additional information or comments you may have about the display designs, the experimental setup, or the experiment as a whole:


Display Reference Guide
A
B
C
D
5

Usability Survey

Display Design Experiment Usability Survey **AFIT SCN: 2014-04**
Expiration Date: 16MAY15

Participant Number: __________

Displays

Select the number that most appropriately reflects your impression of:

Information on the displays

hard to read **easy to read**

1 2 3 4 5 6 7 8 9

The use of colors for differing process variable types

unhelpful **helpful**

1 2 3 4 5 6 7 8 9

Highlighting of process variables to aid in mouse cursor placement

unhelpful **helpful**

1 2 3 4 5 6 7 8 9

Highlighting of process variables to aid in mouse cursor placement

unhelpful **helpful**

1 2 3 4 5 6 7 8 9

Process layouts made sense

never **always**

1 2 3 4 5 6 7 8 9

Amount of information displayed on each individual process

inadequate **adequate**

1 2 3 4 5 6 7 8 9

Amount of information displayed during each trial (all processes together)

inadequate **adequate**

1 2 3 4 5 6 7 8 9

Arrangement of processes on the screens (1-4 across top, 5-8 across bottom)

illogical **logical**

1 2 3 4 5 6 7 8 9

Please write any additional comments about the Displays here:

System Information

Select the number which most appropriately reflects your impression of:

Process variable values / data movement (the simulation) throughout the trials

unbelievable **believable**

1 2 3 4 5 6 7 8 9

Process variable values / data movement related well to the tasks at hand?

never **always**

1 2 3 4 5 6 7 8 9

Number of pauses for each trial was

excessive **adequate**

1 2 3 4 5 6 7 8 9

Terminology on the screen

ambiguous **precise**

1 2 3 4 5 6 7 8 9

Messages appearing on screen

confusing **clear**

1 2 3 4 5 6 7 8 9

Positioning of areas to click on the screen (e.g. predict button location, etc.)

inconsistent **consistent**

1 2 3 4 5 6 7 8 9

Performing an operation led to a predictable result

never **always**

1 2 3 4 5 6 7 8 9

Quantity of deviations experienced in each trial

unacceptable **acceptable**

1 2 3 4 5 6 7 8 9

Frequency of deviation appearances

unacceptable **acceptable**

1 2 3 4 5 6 7 8 9

Please write any additional comments about this topic here:

Learning

Select the number which most appropriately reflects your impression of:

Learning to operate the mouse

difficult **easy**

1 2 3 4 5 6 7 8 9

Getting started

difficult **easy**

1 2 3 4 5 6 7 8 9

Time to learn the experimental setup and interface

slow **fast**

1 2 3 4 5 6 7 8 9

Remembering the different process variable types (temp/press/vib/flow)

difficult **easy**

1 2 3 4 5 6 7 8 9

Remembering how to “predict”

difficult **easy**

1 2 3 4 5 6 7 8 9

Remembering what penalties exist / caused a reduction in score

difficult **easy**

1 2 3 4 5 6 7 8 9

Tasks can be performed in a straight-forward manner

never **always**

1 2 3 4 5 6 7 8 9

Steps to complete a task follow a logical sequence

never **always**

1 2 3 4 5 6 7 8 9

Please write your comments about this topic here:

System Capabilities

Select the number which most appropriately reflects your impression of:

Response time for operations

inadequate **adequate**

1 2 3 4 5 6 7 8 9

The rate at which information is displayed

inappropriate **appropriate**

1 2 3 4 5 6 7 8 9

The system is reliable

never **always**

1 2 3 4 5 6 7 8 9

Ease of operation is consistent throughout

never **always**

1 2 3 4 5 6 7 8 9

You can accomplish tasks with minimal instruction

never **always**

1 2 3 4 5 6 7 8 9

The reading tasks were

inappropriate **appropriate**

1 2 3 4 5 6 7 8 9

The reading tasks provided a good distraction away from the primary task

never **always**

1 2 3 4 5 6 7 8 9

Please write your comments about this topic here:

Training Session

Select the number which most appropriately reflects your impression of the training sessions you received.

Training sessions were

confusing **clear**

1 2 3 4 5 6 7 8 9

The terminology used in the training

confusing **clear**

1 2 3 4 5 6 7 8 9

Information in the training was easily understood

never **always**

1 2 3 4 5 6 7 8 9

Amount of time given for training

inadequate **adequate**

1 2 3 4 5 6 7 8 9

Please write your comments about this topic here:

Environment

Select the number which most appropriately reflects your impression of:

Lighting in the experiment room

uncomfortable **comfortable**

1 2 3 4 5 6 7 8 9

Temperature of the experiment room

uncomfortable **comfortable**

1 2 3 4 5 6 7 8 9

Location of the operator during the experiment

Inappropriate **appropriate**

1 2 3 4 5 6 7 8 9

Experimenter presence and activity during trials was

distracting **acceptable**

1 2 3 4 5 6 7 8 9

Please write any additional comments about this topic here:

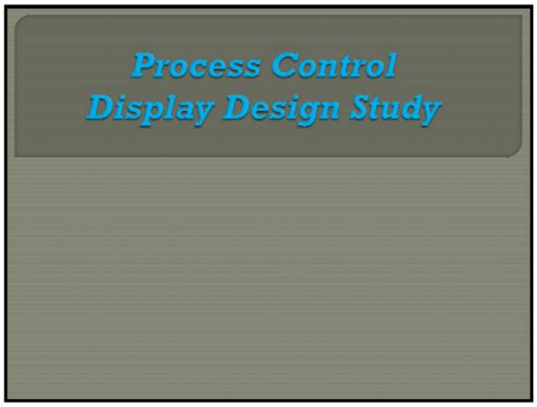
Overall Impressions

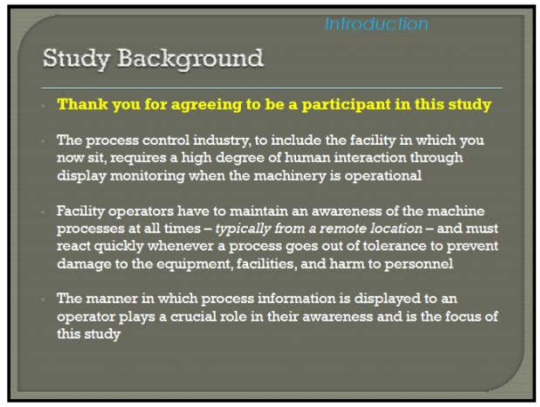
Select the number that most appropriately reflects your overall impression of using the experimental setup.

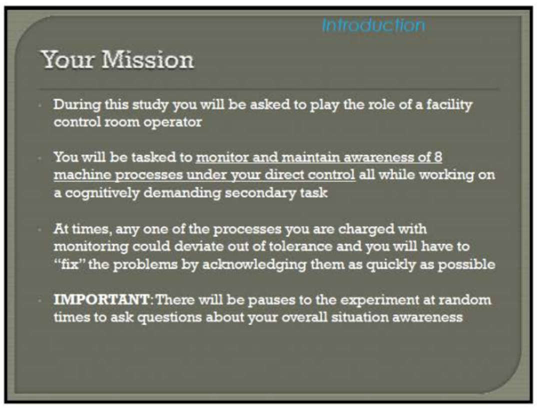
bad										**good**
	1	2	3	4	5	6	7	8	9	
frustrating										**satisfying**
	1	2	3	4	5	6	7	8	9	
dull										**stimulating**
	1	2	3	4	5	6	7	8	9	
difficult										**easy**
	1	2	3	4	5	6	7	8	9	

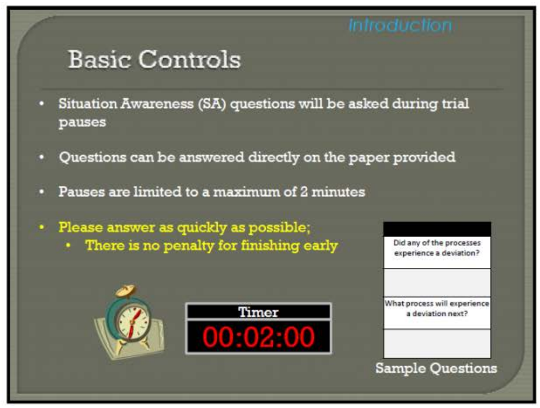
Please list below or discuss with the experimenter any additional comments/concerns you may have about your experience during this study. Your feedback is needed to refine the experimental setup (to include this and any other questionnaires) before proceeding to the next phase. Your input is valued and very much appreciated.

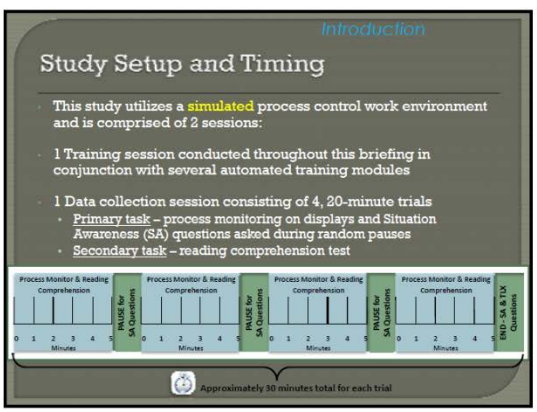
Experiment Training Slides

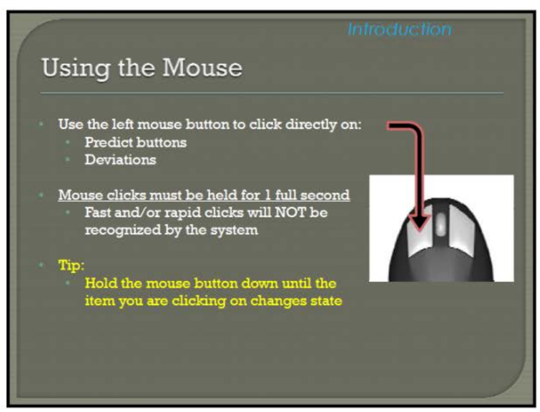


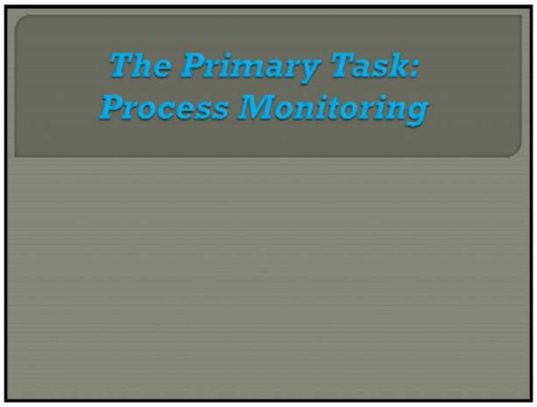


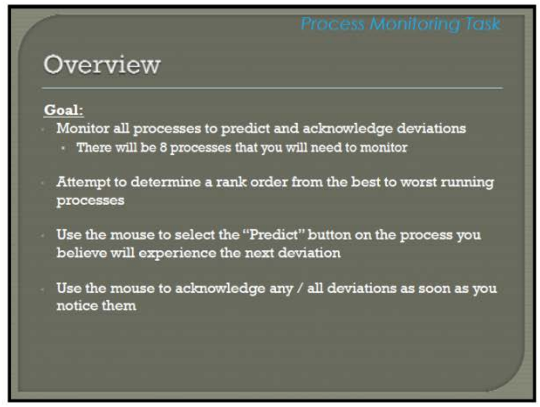


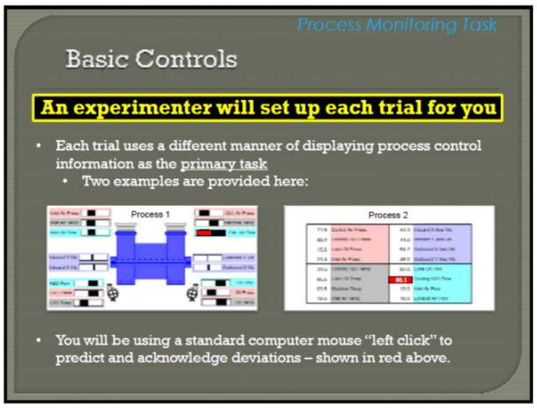


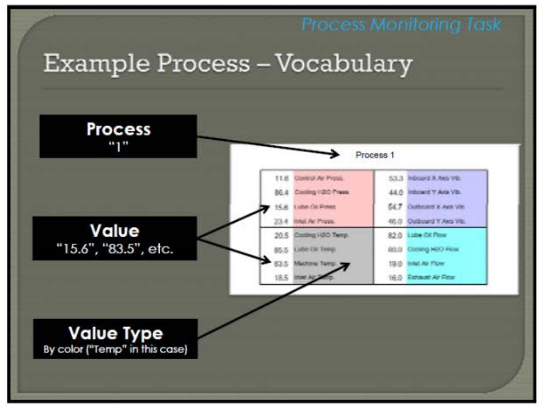


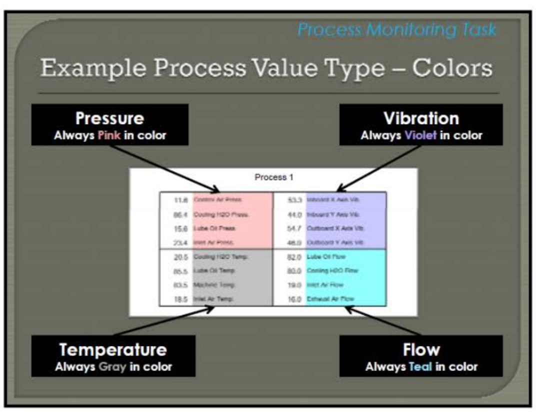


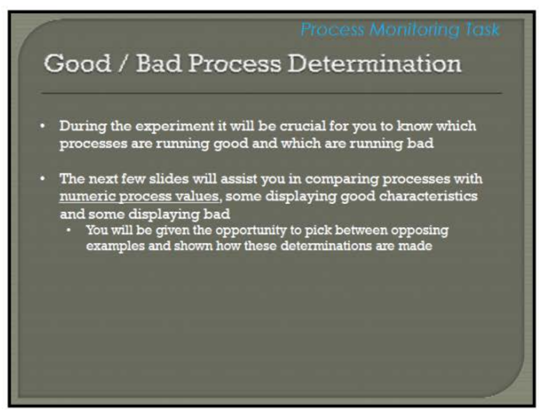

The Primary Task:
Process Monitoring

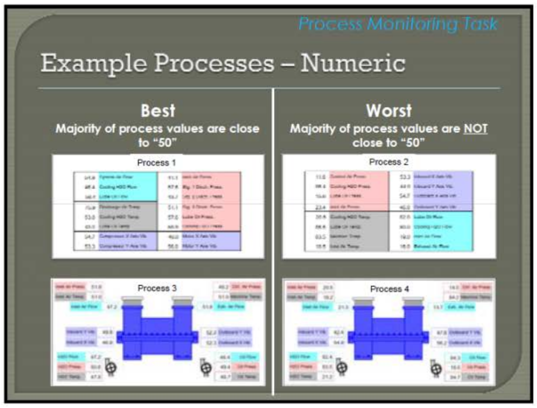

Process Monitoring Task
Overview
Goal:
Monitor all processes to predict and acknowledge deviations
There will be 8 processes that you will need to monitor
Attempt to determine a rank order from the best to worst running processes
Use the mouse to select the "Predict" button on the process you believe will experience the next deviation
Use the mouse to acknowledge any / all deviations as soon as you notice them

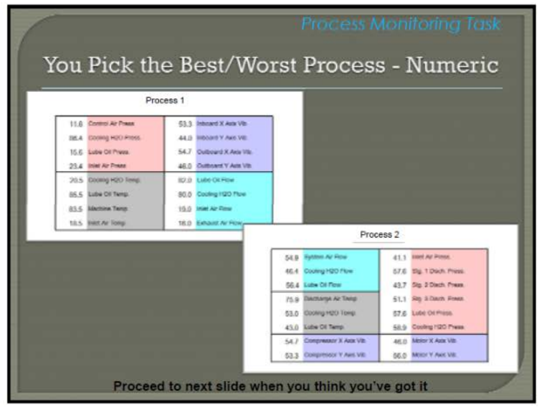

Process Monitoring Task
Basic Controls
An experimenter will set up each trial for you
Each trial uses a different manner of displaying process control information as the primary task
Two examples are provided here:
Process 1
Process 2
You will be using a standard computer mouse "left click" to predict and acknowledge deviations – shown in red above.

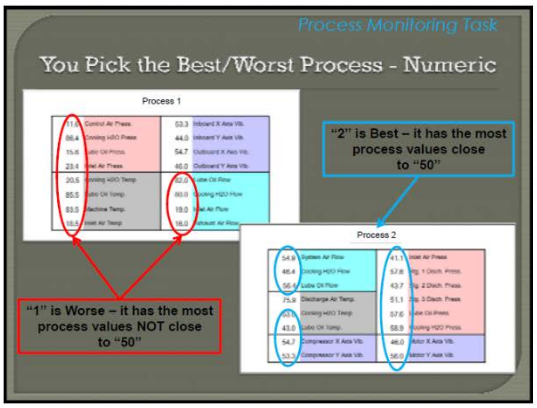

Process Monitoring Task
Example Process – Vocabulary
Process
"1"
Value
"15.6", "83.5", etc.
Value Type
By color ("Temp" in this case)
Process 1

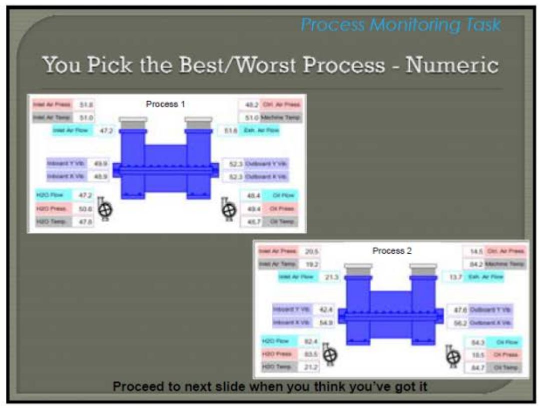

Process Monitoring Task
Example Process Value Type – Colors
Pressure
Always Pink in color
Vibration
Always Violet in color
Process 1
Temperature
Always Gray in color
Flow
Always Teal in color

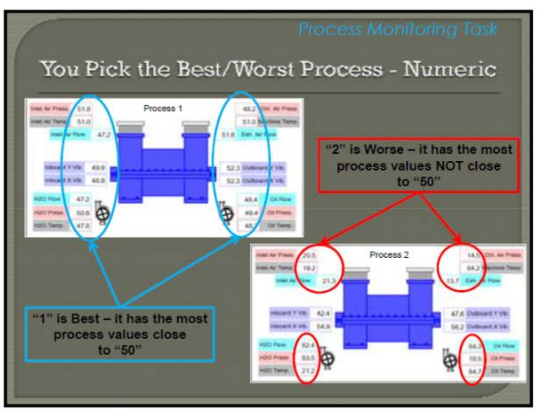

Process Monitoring Task
Good / Bad Process Determination
During the experiment it will be crucial for you to know which processes are running good and which are running bad
The next few slides will assist you in comparing processes with numeric process values, some displaying good characteristics and some displaying bad
You will be given the opportunity to pick between opposing examples and shown how these determinations are made

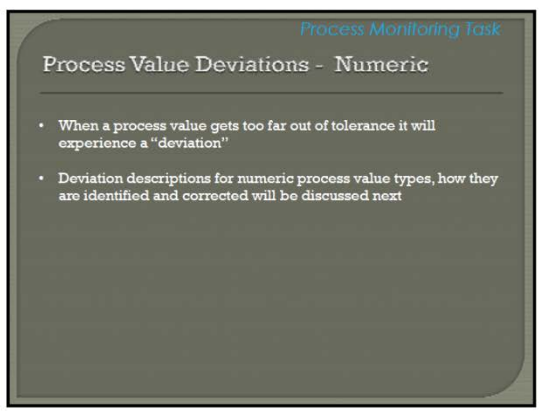

Process Monitoring Task
Example Processes – Numeric
Best
Majority of process values are close to "50"
Worst
Majority of process values are NOT close to "50"
Process 1
Process 2
Process 3
Process 4


Process Monitoring Task
You Pick the Best/Worst Process - Numeric
Process 1
Process 2
Proceed to next slide when you think you've got it


Process Monitoring Task
You Pick the Best/Worst Process - Numeric
Process 1
"2" is Best – it has the most process values close to "50"
Process 2
"1" is Worse – it has the most process values NOT close to "50"

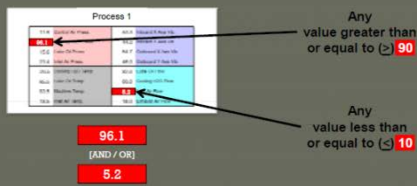

Process Monitoring Task
You Pick the Best/Worst Process - Numeric
Process 1
Process 2
Proceed to next slide when you think you've got it


Process Monitoring Task
You Pick the Best/Worst Process - Numeric
Process 1
"2" is Worse – it has the most process values NOT close to "50"
Process 2
"1" is Best – it has the most process values close to "50"


Process Monitoring Task
Process Value Deviations - Numeric
When a process value gets too far out of tolerance it will experience a "deviation"
Deviation descriptions for numeric process value types, how they are identified and corrected will be discussed next

Process Monitoring Task

Example Process Deviations - Numeric



- Deviation characteristics:
  - Can occur in any process at any time
  - Will be more frequent in the worst running processes
  - Turn and stay red until they are acknowledged by left mouse click

Process Monitoring Task

Predicting Deviations

- Now that you are familiar with good/bad process identification and deviations you are ready to predict in which process a deviation may occur
- How to predict a deviation and prediction characteristics will be discussed next

Process Monitoring Task

Example Prediction

- Prediction characteristics:
  - Can predict any process , but once selected cannot unselect
  - Ability to predict goes away if any deviation is present
  - Turn and stay blue until a deviation occurs
  - No penalty for predicting an incorrect process

STOP

Please let the researcher know if you have any questions and you are ready to practice:

Numeric Processes

Training Modules 1 and 2

Process Monitoring Task

Good / Bad Process Determination

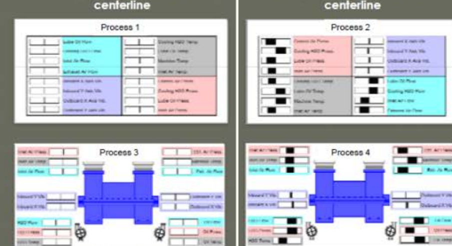
- Now that you are familiar with numeric type process values let's explore another type, graphic
- The next few slides will assist you in comparing processes with graphic process values, some displaying good characteristics and some displaying bad
  - You will be given the opportunity to pick between opposing examples and shown how determinations are made

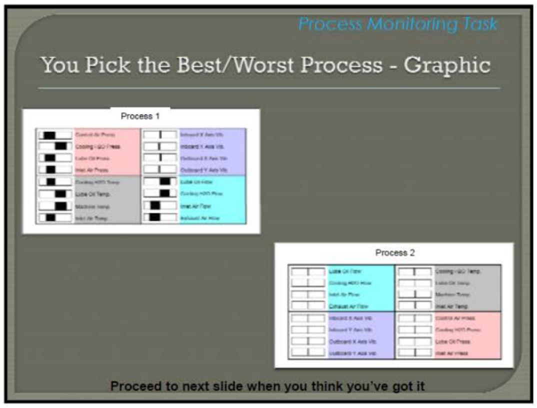
Process Monitoring Task

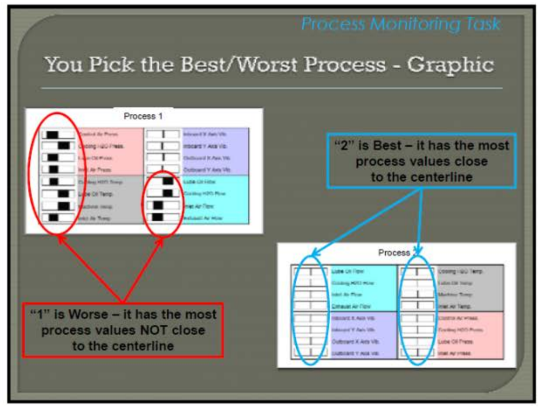
Example Processes – Graphic

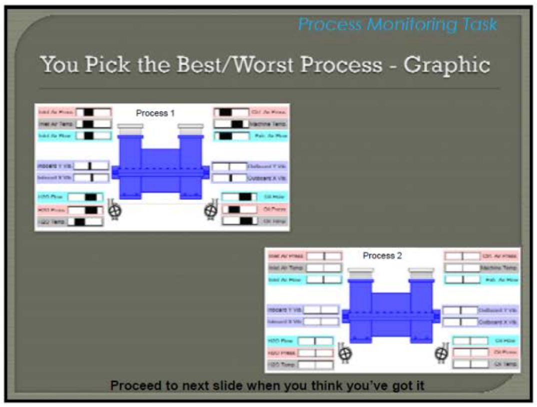
Best
Most process values near centerline

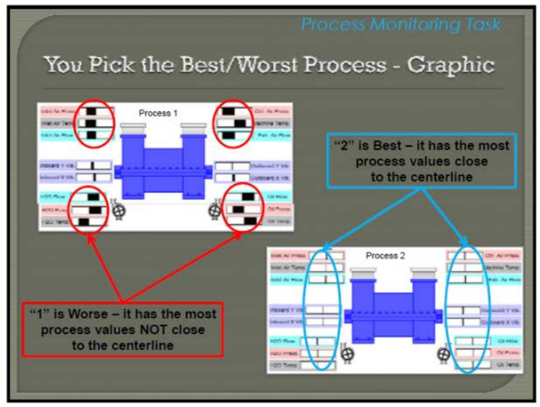
Worst
Most process values far from centerline

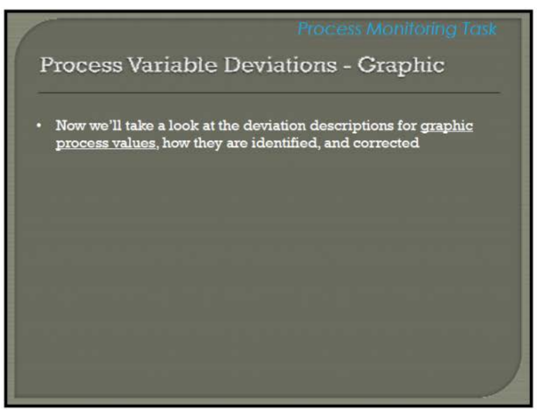


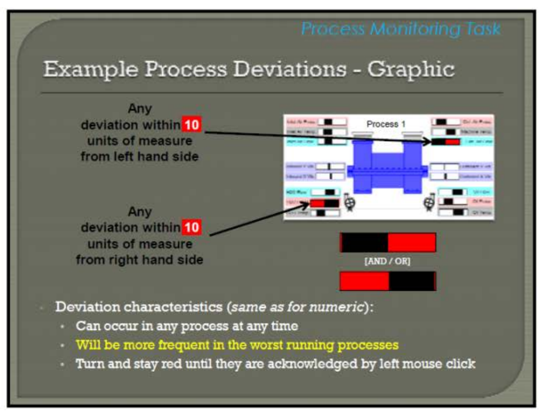

Process Monitoring Task
You Pick the Best/Worst Process - Graphic
Process 1
Process 2
Proceed to next slide when you think you've got it


Process Monitoring Task
You Pick the Best/Worst Process - Graphic
Process 1
"2" is Best – it has the most process values close to the centerline
"1" is Worse – it has the most process values NOT close to the centerline
Process


Process Monitoring Task
You Pick the Best/Worst Process - Graphic
Process 1
Process 2
Proceed to next slide when you think you've got it


Process Monitoring Task
You Pick the Best/Worst Process - Graphic
Process 1
"2" is Best – it has the most process values close to the centerline
"1" is Worse – it has the most process values NOT close to the centerline
Process 2


Process Monitoring Task
Process Variable Deviations - Graphic
• Now we'll take a look at the deviation descriptions for graphic process values, how they are identified, and corrected


Process Monitoring Task
Example Process Deviations - Graphic
Any deviation within 10 units of measure from left hand side
Process 1
Any deviation within 10 units of measure from right hand side
[AND / OR]
Deviation characteristics (same as for numeric):
• Can occur in any process at any time
• Will be more frequent in the worst running processes
• Turn and stay red until they are acknowledged by left mouse click

Process Monitoring Task

Remember Prediction?

- Prediction characteristics:
  - Can predict any process , but once selected cannot unselect
  - Ability to predict goes away if any deviation is present
  - Turn and stay blue until a deviation occurs
  - No penalty for predicting an incorrect process

STOP

Please let the researcher know if you have any questions and you are ready to practice:

Graphic Processes

Training Modules 3 and 4

The Secondary Task: Reading Comprehension

Secondary Task

Overview

Purpose:
The secondary task (reading comprehension) is intended to mimic the workload experienced by a control room operator who has many responsibilities in addition to just process monitoring

- Secondary tasks typically include: responding to requests for information, reporting on equipment, doing email, etc.

Goal:

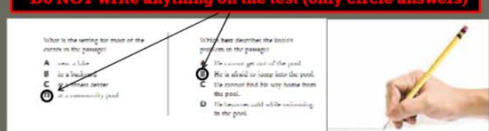
- Answer all questions correctly in the allotted timeframe
- Remain fully aware of the primary task (monitoring the processes)
  - Remember the comprehension task is a secondary, lower priority task than the primary task
- Resist the urge to guess or pick answers at random
  - There will be a penalty assessed for wrong answers on this task only

Secondary Task

Basic Controls

- The secondary task is a multiple choice comprehension test based on a given reading passage.
- This task is accomplished at the same time as the primary task
- 2 pencils are provided and answers may be changed at any time until the 20-minute trial ends.
- Make sure you double check your test!

Do NOT write anything on the test (only circle answers)

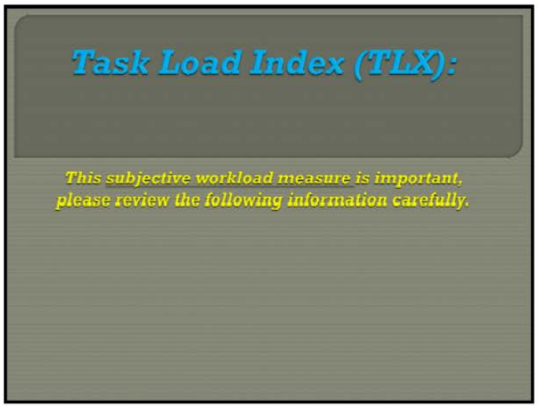


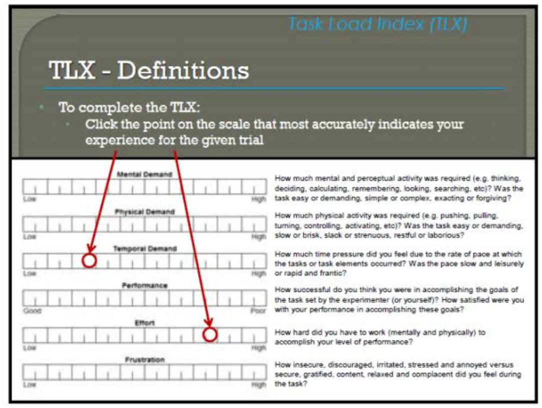
STOP

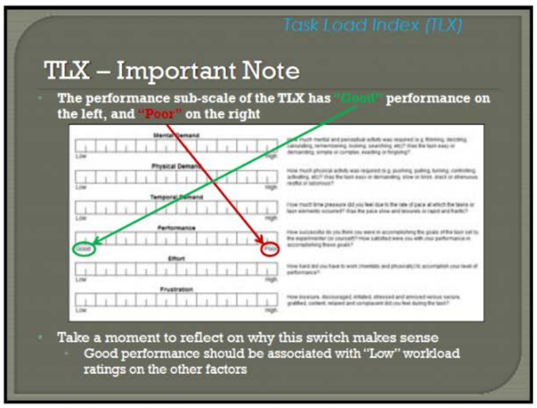
Please let the researcher know if you have any questions and you are ready to practice:

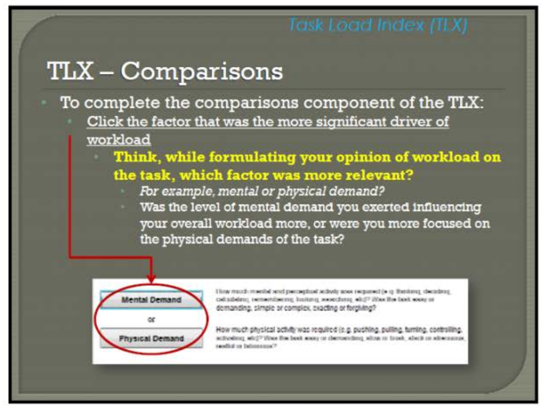
Secondary Task

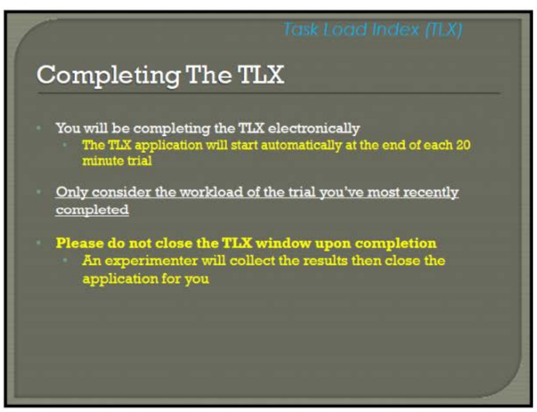
Handouts

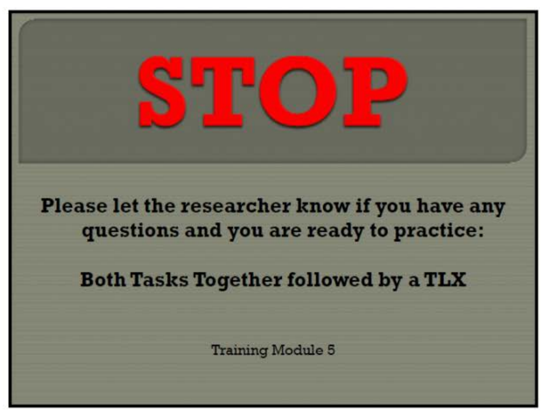

Task Load Index (TLX):
This subjective workload measure is important, please review the following information carefully.


Task Load Index (TLX)
TLX - Definitions
To complete the TLX:
Click the point on the scale that most accurately indicates your experience for the given trial
Mental Demand
How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical Demand
How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal Demand
How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance
How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort
How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration
How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?
Low
High
Good
Poor


Task Load Index (TLX)
TLX – Important Note
The performance sub-scale of the TLX has "Good" performance on the left, and "Poor" on the right
Take a moment to reflect on why this switch makes sense
Good performance should be associated with "Low" workload ratings on the other factors


Task Load Index (TLX)
TLX – Comparisons
To complete the comparisons component of the TLX:
Click the factor that was the more significant driver of workload
Think, while formulating your opinion of workload on the task, which factor was more relevant?
For example, mental or physical demand?
Was the level of mental demand you exerted influencing your overall workload more, or were you more focused on the physical demands of the task?
Mental Demand
or
Physical Demand


Task Load Index (TLX)
Completing The TLX
You will be completing the TLX electronically
The TLX application will start automatically at the end of each 20 minute trial
Only consider the workload of the trial you've most recently completed
Please do not close the TLX window upon completion
An experimenter will collect the results then close the application for you


STOP
Please let the researcher know if you have any questions and you are ready to practice:
Both Tasks Together followed by a TLX
Training Module 5

Scoring:
How points are earned

Scoring

The Competition

- Throughout this experiment, you will be engaged in a competition based on your individual trial scores
- The top scorer at the end of the experiment will be awarded:
  - **A $25 gift card to Texas Roadhouse**
- Participant ID numbers will be used to maintain anonymity
  - Four data collection trials means you will have four chances to achieve top scores!

Scoring

Overview

- Trial Scoring
  - Maximum points possible per trial: 1000

  Primary task – process monitoring total value: 800:

  400 for Predicting and Acknowledging deviations
  400 for Situation Awareness responses during pauses

  Secondary task – reading comprehension value: 200

Scoring

Primary Task Point Structure – Deviations

- Predictions Value: 200
  - Must predict accurately to get full score
  - There is **NO PENALTY** for a false prediction
    - So be sure to predict (even if you must guess) whenever this option is available!
- Deviations Acknowledgement: 200
  - Time based scoring so acknowledge as quickly as possible
    - Every 5-seconds without an acknowledgement reduces awarded point value
    - Clicking on process values that are not exhibiting a deviation results in a penalty / reduction in score!

Scoring

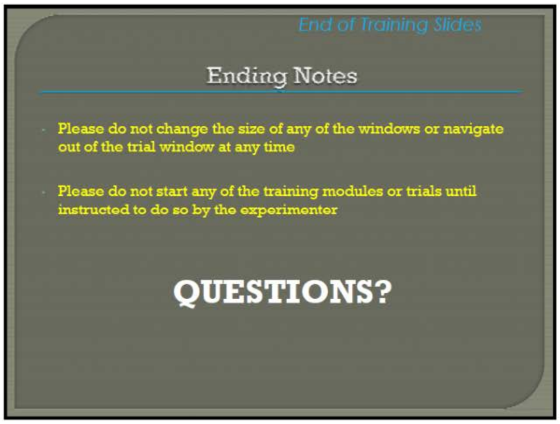
Primary Task Point Structure – SA

- Situation Awareness Responses: 400
  - Administered during pauses
  - You will be asked to recall things you've seen or done during the trial
    - Answer to the best of your ability
  - Unanswered questions are scored as 0 points so be sure to answer all of the questions in the 2-minute maximum allotted time

Scoring

Secondary Task Point Structure

- Reading Comprehension Test Value: 200
  - 8 questions worth 25 points each
- Incorrect responses nominally reduce score, whereas unanswered questions are simply 0 points
- Tips (for the reading task **ONLY**):
  - If you have no idea what the answer is: DON'T GUESS, Leave it blank
  - If you are able to eliminate two or more of the multiple choice options then circle the best one


End of Training Slides
Ending Notes
Please do not change the size of any of the windows or navigate out of the trial window at any time
Please do not start any of the training modules or trials until instructed to do so by the experimenter
QUESTIONS?

Appendix E – ANOVA Interval Plots

Appendix E contains ANOVA interval plot figures produced by the Minitab 17 statistical analysis software. Several of the figures in this appendix appear in other areas of the document, but quite a few do not. In most cases where these results do not appear, this was done for the purposes of brevity and with consideration toward space constraints dictated by a majority of paper call submittal guidance. These limitations resulted in the omission of ANOVA plot data that did not present findings of significance. This appendix presents both the included and omitted information here to aid the reader in forming a more comprehensive picture of the study's results.

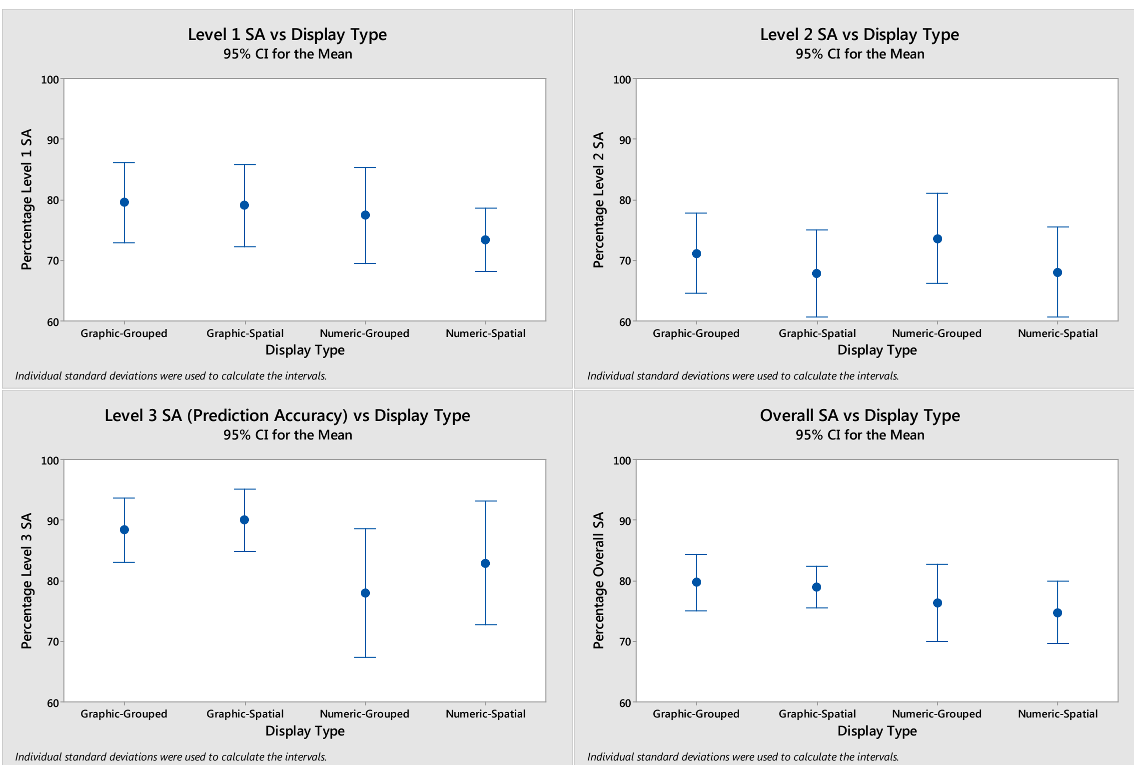


Figure E-1 Display Construct Influence on SA

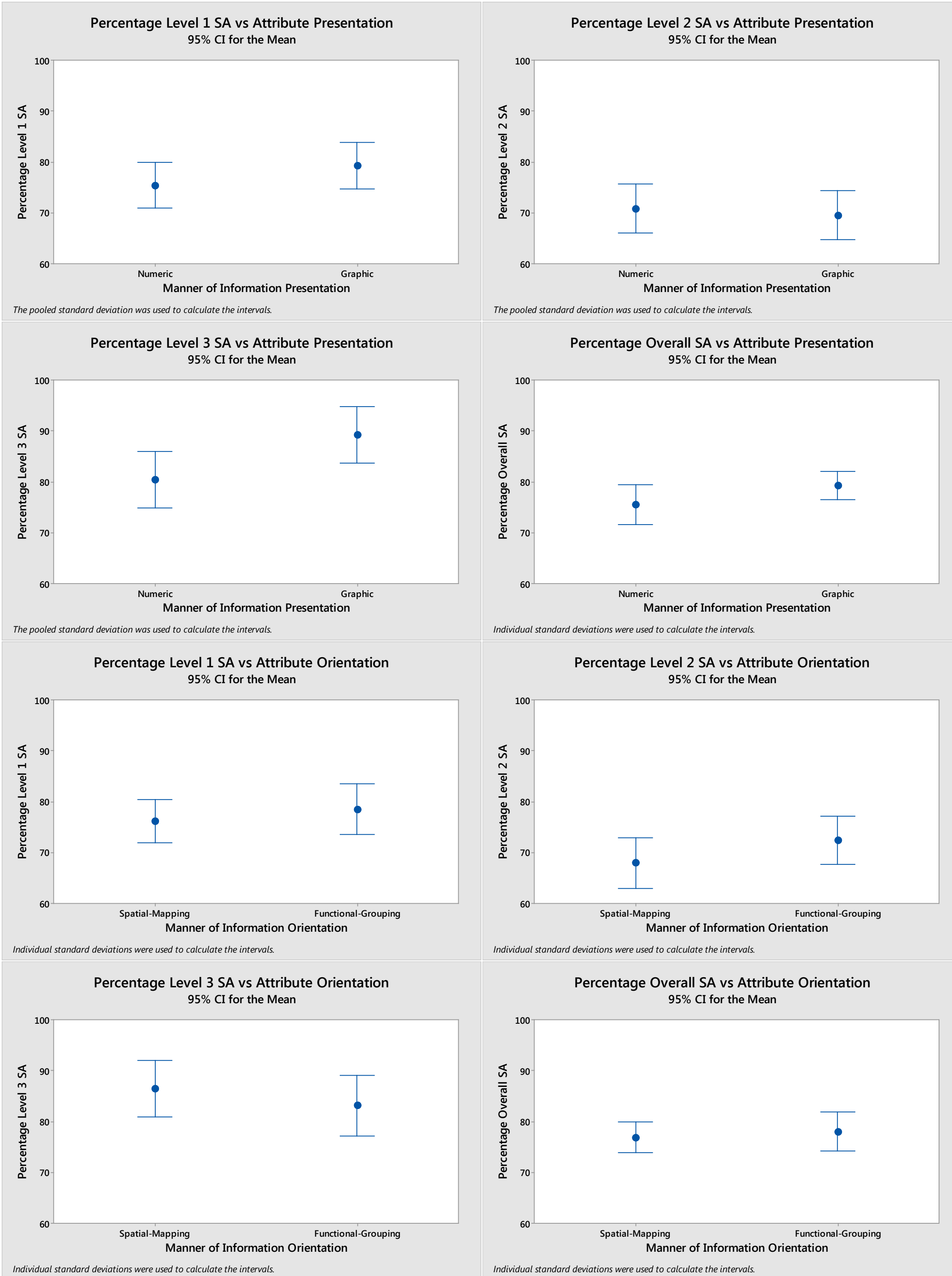


Figure E-2 Individual Display Attribute Influence on SA

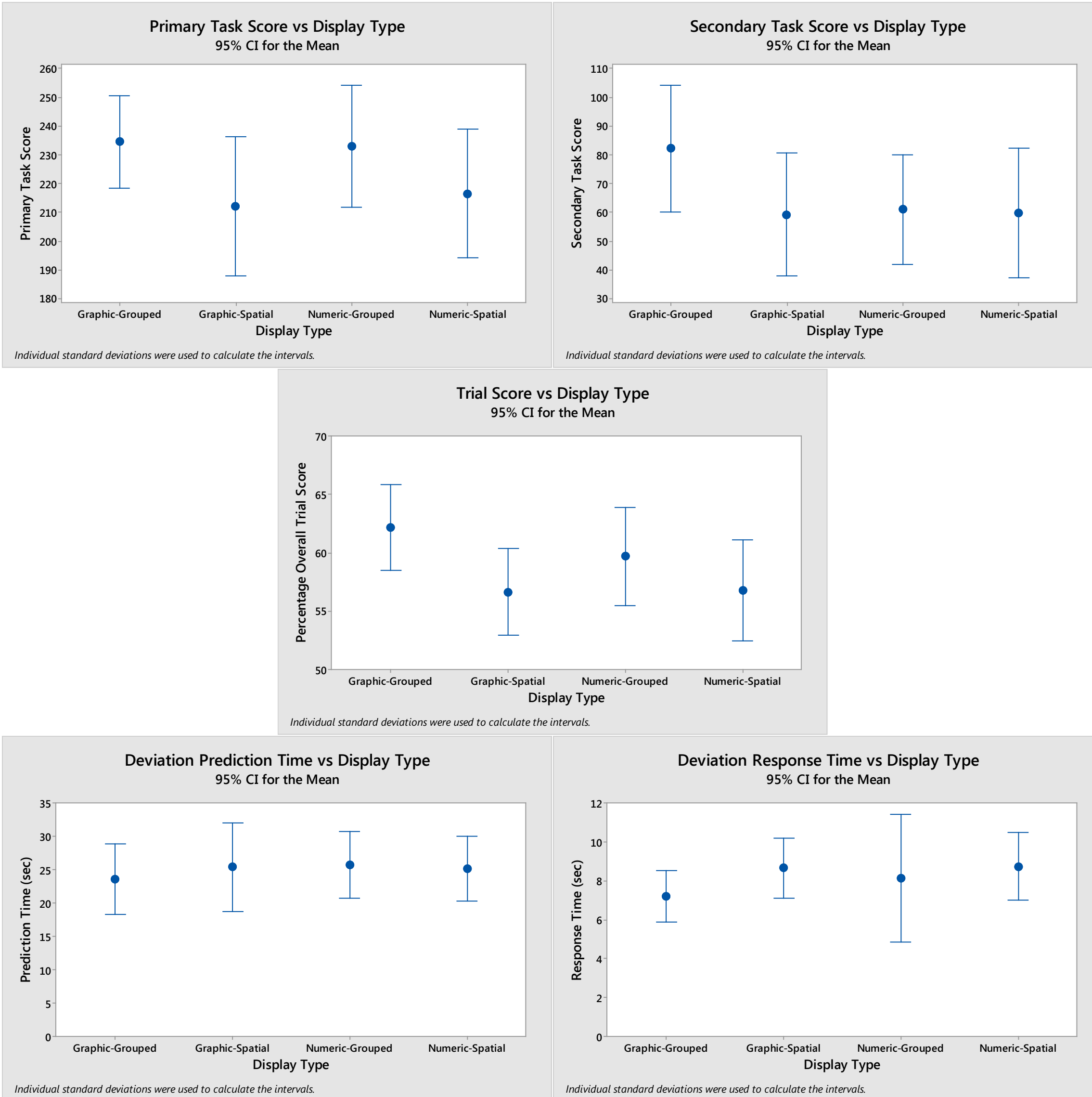


Figure E-3 Display Construct Influence on Performance

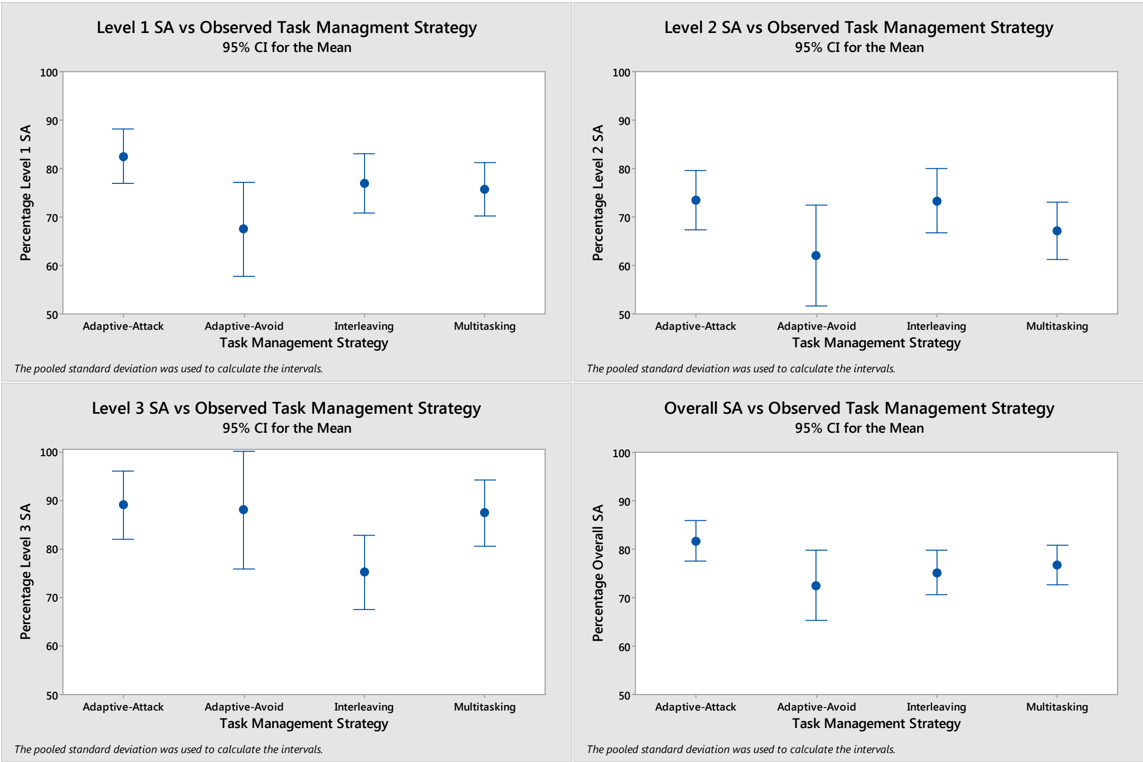


Figure E-4 Task Management Strategy Influence on SA

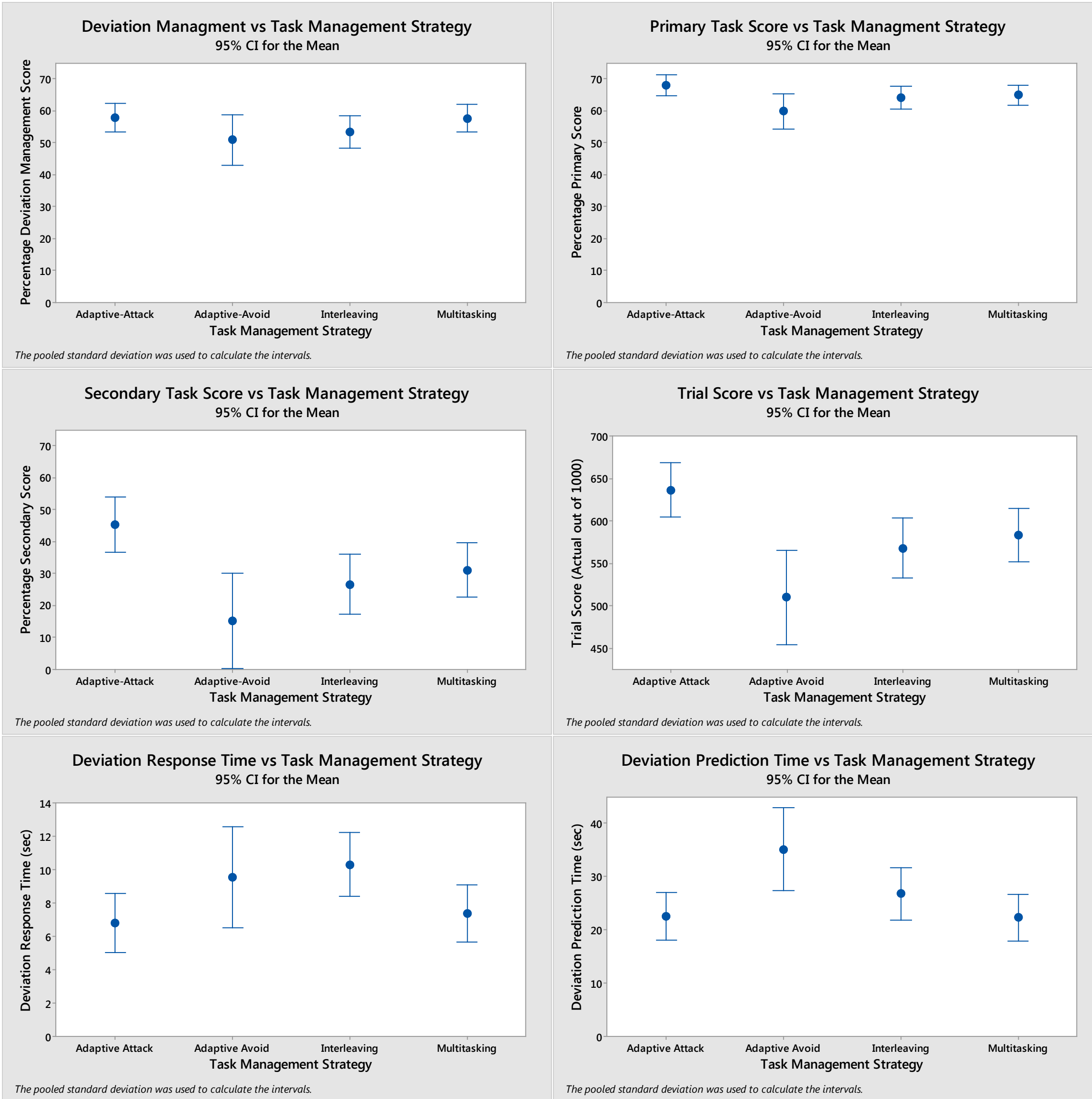


Figure E-5 Task Management Strategy Influence on Performance

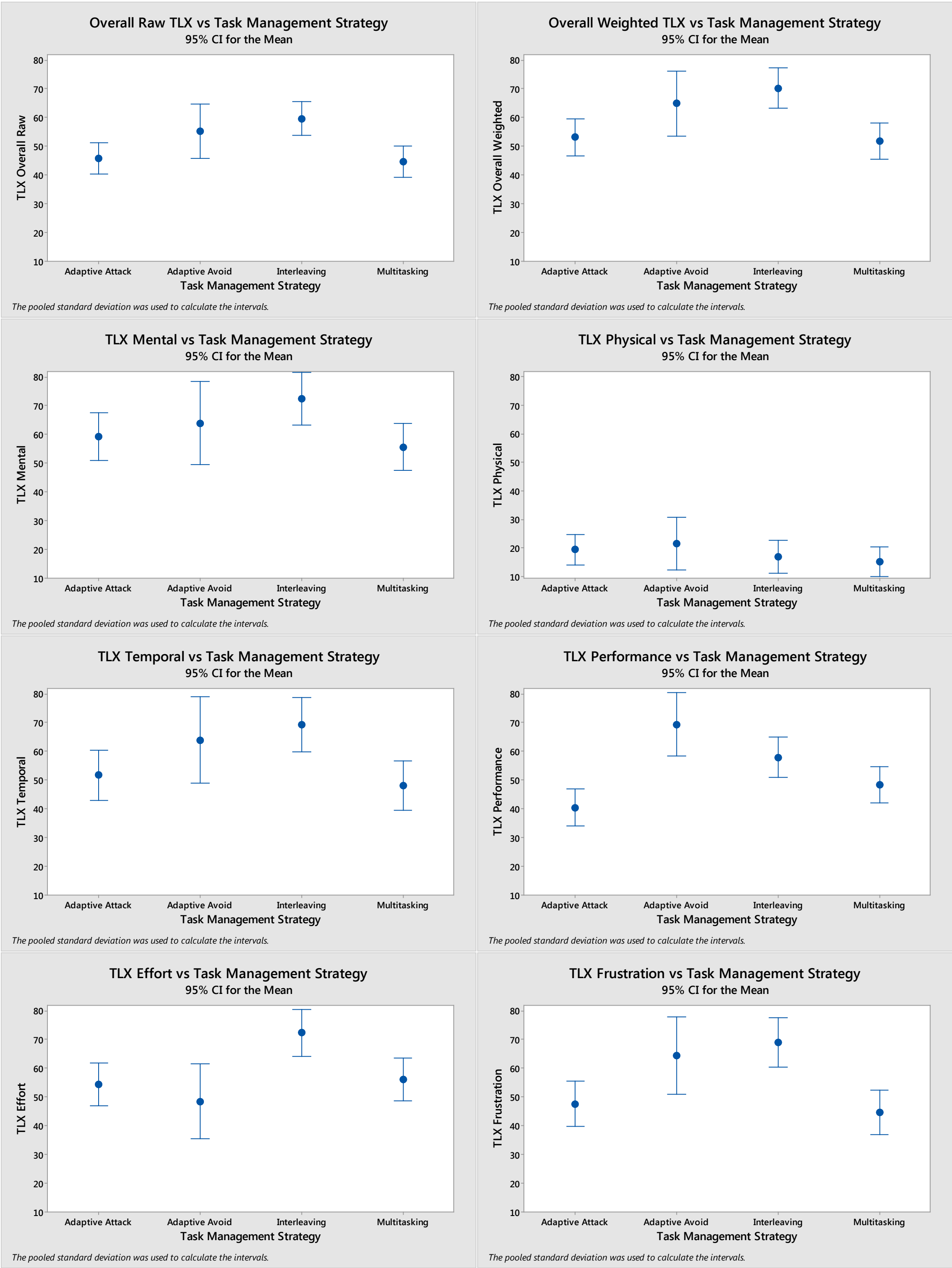


Figure E-6 Task Management Strategy Influence on Workload

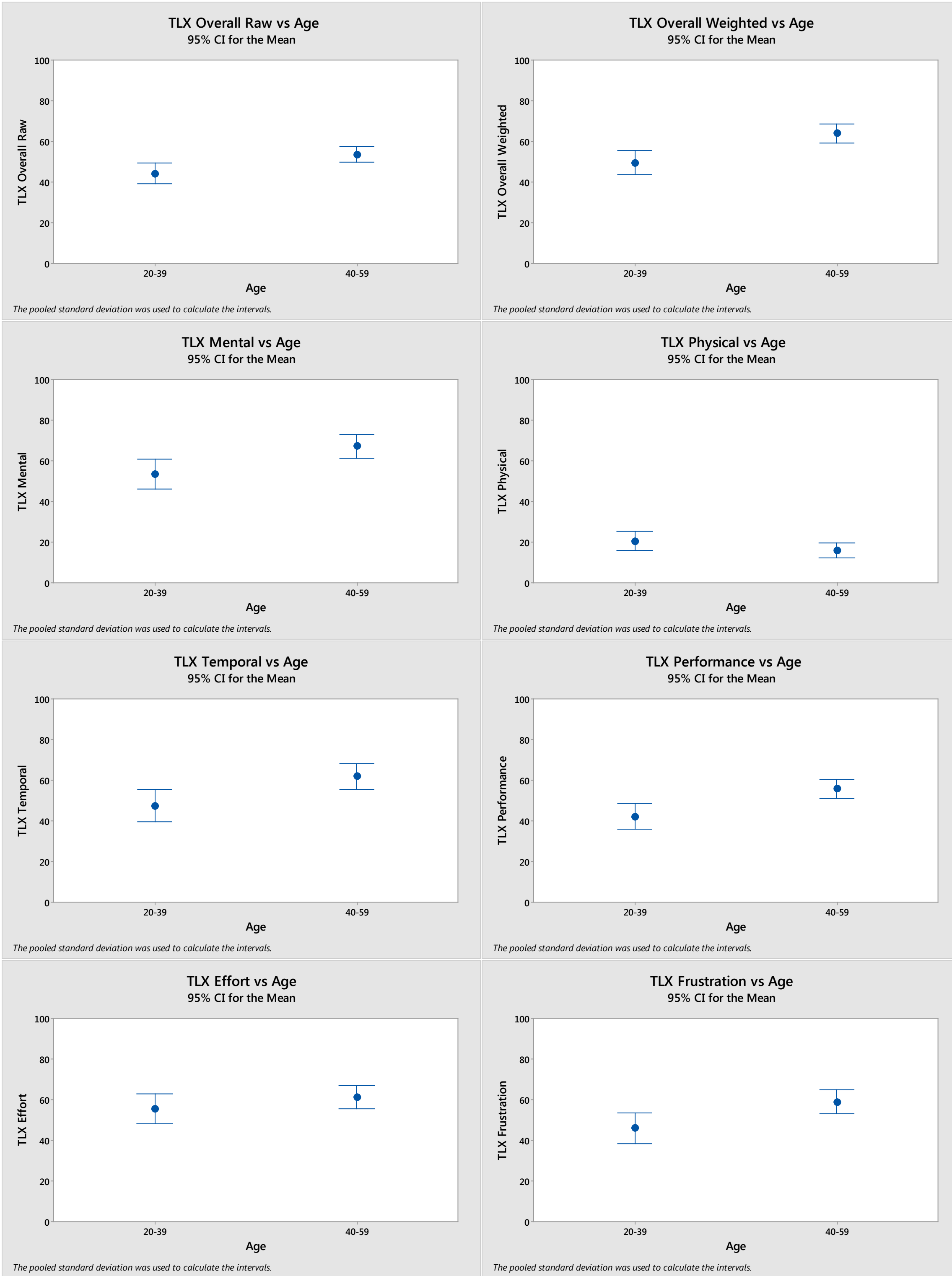


Figure E-7 Age Influence on Workload (split at 40 years)

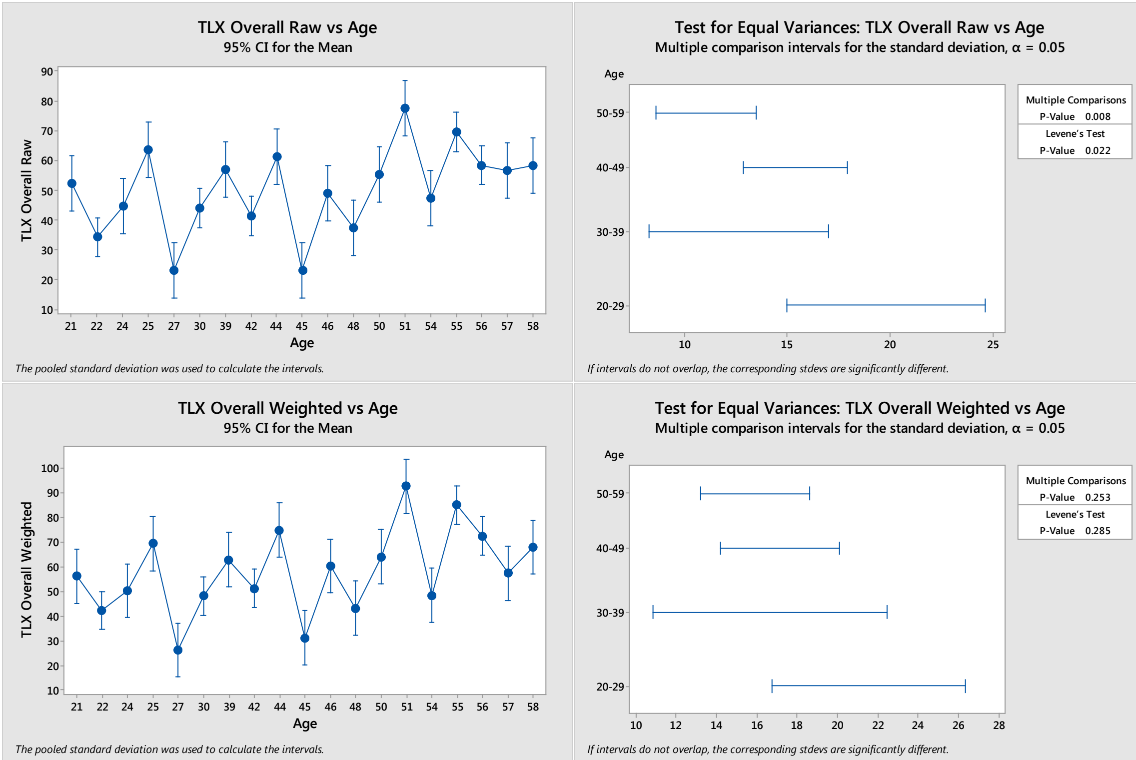


Figure E-8 Age Influence on Raw and Weighted TLX Overall

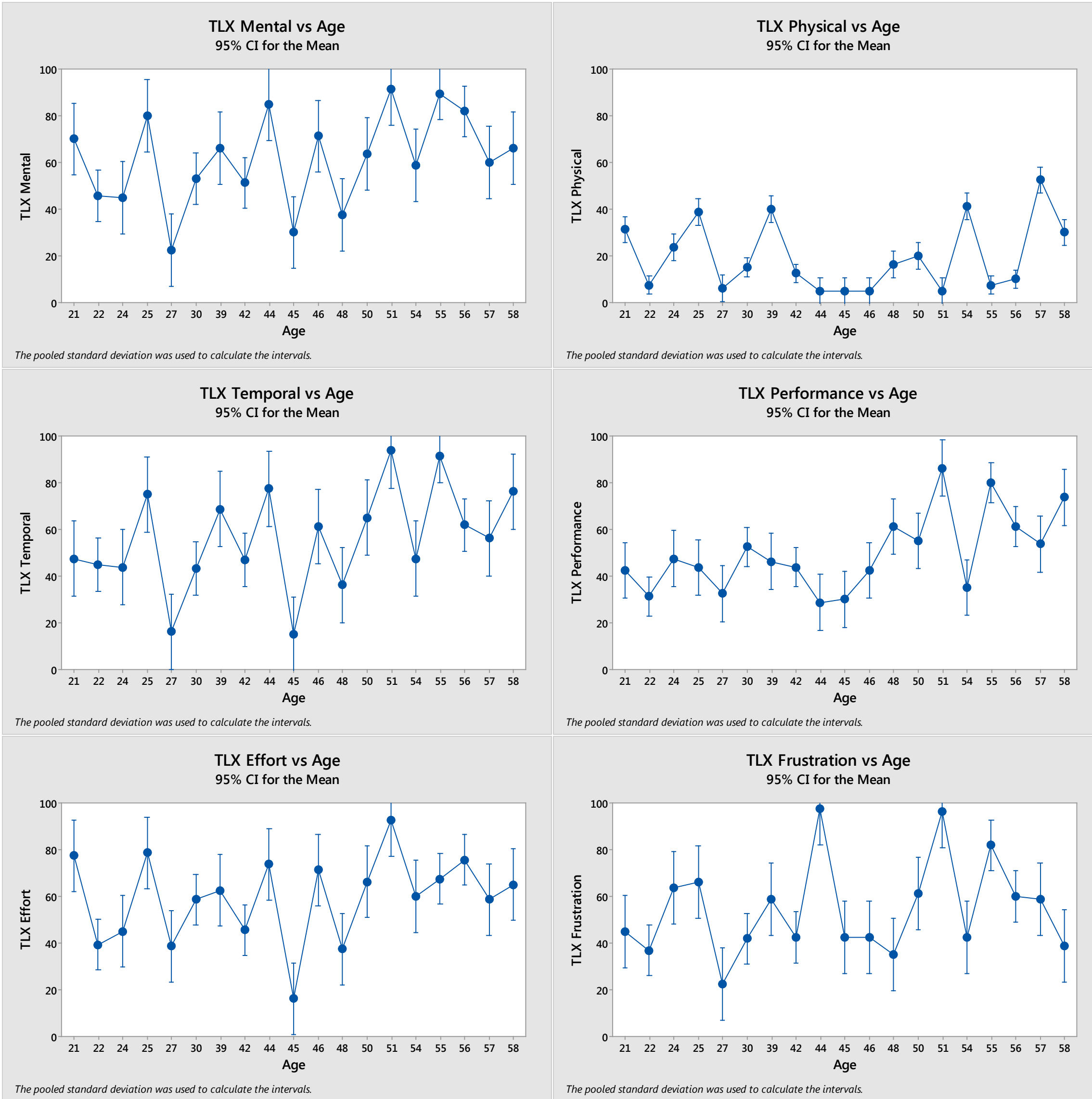


Figure E-9 Age Influence on Workload – Individual TLX Factors
Confidence Interval bars restricted to TLX full scale, 0-100

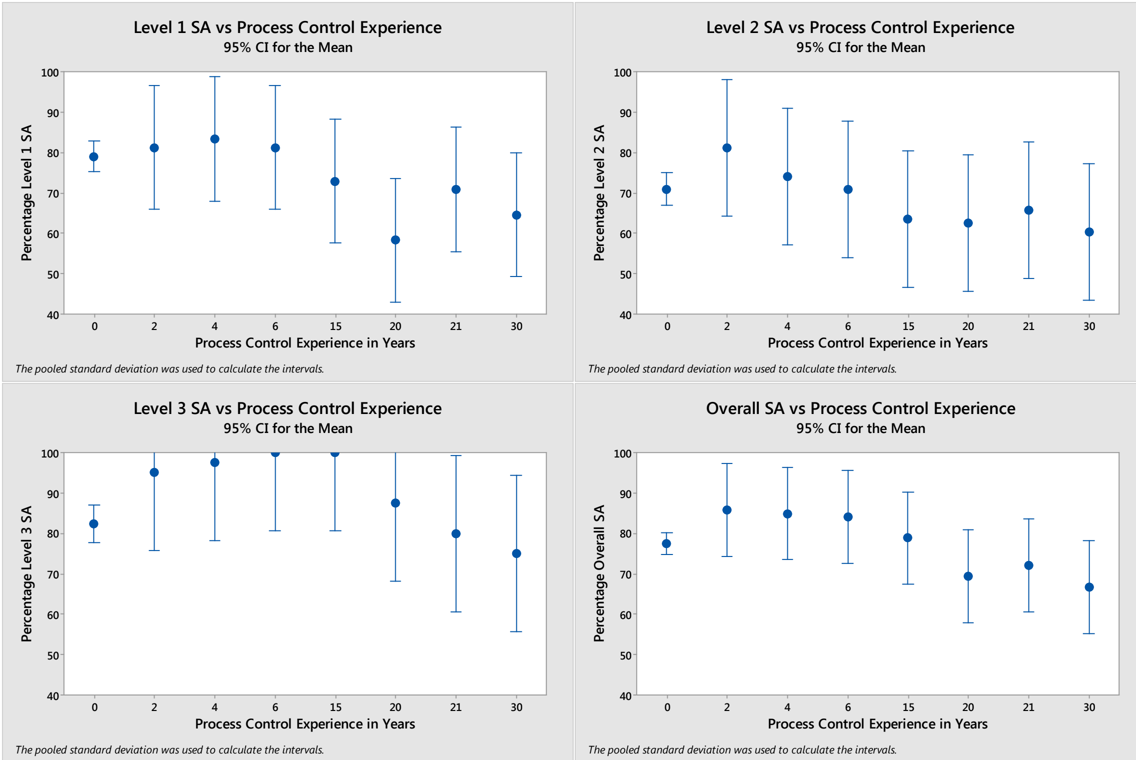


Figure E-10 Process Control Experience Influence on SA
Confidence Interval bars restricted to SA values less than 100%

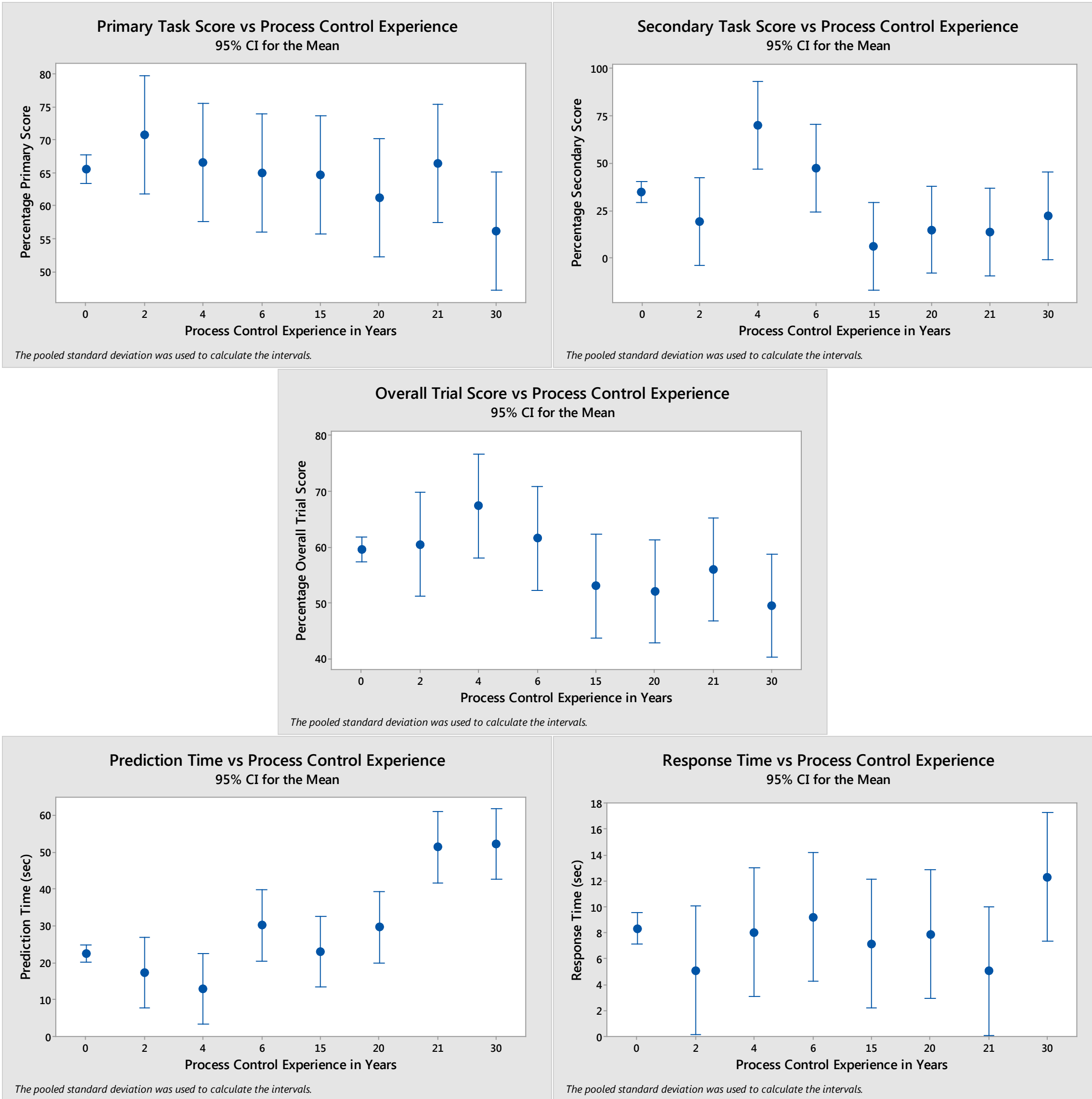


Figure E-11 Process Control Experience Influence on Performance

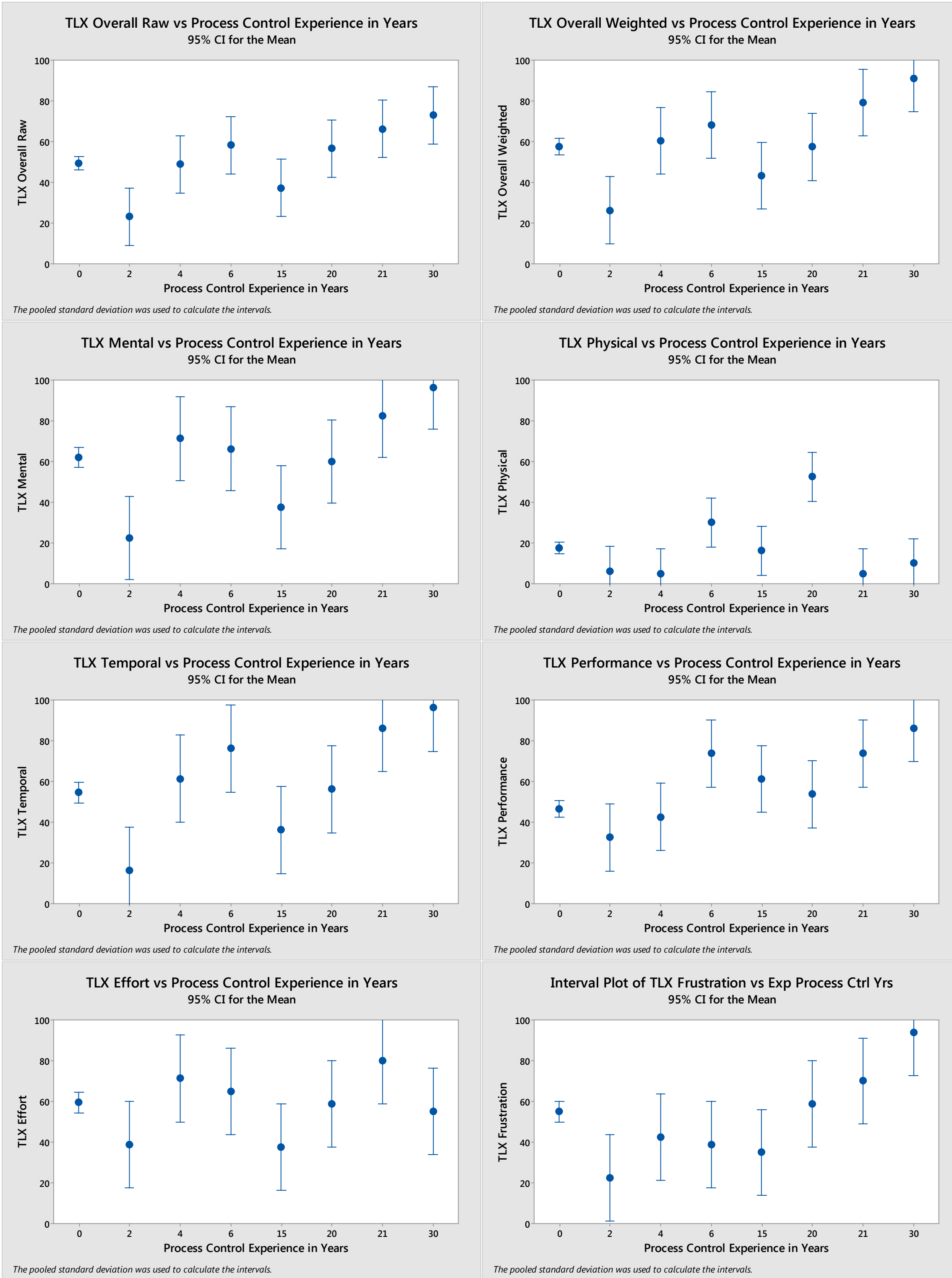


Figure E-12 Process Control Experience Influence on TLX Workload
Confidence Interval bars restricted to TLX full scale, 0-100

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 074-0188*
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**			
1. REPORT DATE *(DD-MM-YYYY)* 26-12-2014	**2. REPORT TYPE** Master's Thesis		**3. DATES COVERED** *(From – To)* October 2011 – December 2014
TITLE AND SUBTITLE Influences of Display Design and Task Management Strategy on Situation Awareness, Performance, and Workload in Process Control Environments			**5a. CONTRACT NUMBER**
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Bowden, James R, Mr.			**5d. PROJECT NUMBER**
			5e. TASK NUMBER
			5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way Wright-Patterson AFB OH 45433-7765			**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT-ENV-MS-14-D-29
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory / Aerospace Systems Directorate 2130 Eighth Street, Bldg. 45, Area B Wright-Patterson Air Force Base, OH 45433 937-255-4210 and curtis.kessler@us.af.mil ATTN: Mr. Curtis D. Kessler			**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/RQOE
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **DISTRUBTION STATEMENT A.** APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.			
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.			

14. ABSTRACT

Process control environments demand well informed high performing human monitors to maintain effectual control of multiple processes. Most research aims to satisfy this requirement through the evaluation of competing heuristic-based display design constructs. Contrary to that method, this study takes a novel approach by examining both factors internal and external to the human observer to identify where beneficial outcomes actually reside. External factors explore the underlying design construct attributes, while internal factors focus on the effect of operator task management strategy, age, and experience. Results from this study present several key findings relative to operator situation awareness, performance, and workload. Findings suggest the specific manner in which external information is presented and oriented on a process control room display is inconsequential toward situation awareness and performance. Further, operator preferred task management strategy has a profound effect on their performance and experienced workload, while exhibiting only a mild effect on situation awareness. In most cases, an Adaptive Attack strategy produces desirable results, while an Adaptive Avoidance does not. Interleaving and Multitasking fall between these two extremes. Lastly, findings indicate subject variables, age and experience have negative effects on overall situation awareness and system deviation prediction times.

15. SUBJECT TERMS

human factors, process control, display design, situation awareness, task performance, workload

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Maj Christina Rusnock, PhD, AFIT/ENV
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	188	(937) 255-3636, ext 4611 Christina.Rusnock@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18