Operator Performance in Long Duration Control Operations: Switching from Low to High Task Load

Kristopher M. Thornburg[1], Henricus P.M. Peterse[1,2], Andrew M. Liu[1]
[1]MIT Humans and Automation Lab, Cambridge, MA
[2]Delft University of Technology, Delft, The Netherlands

Long duration, low task load environments are typical for nuclear power plant control rooms, where operators, after hours of operating under a low task load situation, may have to shift to a high task load situation. The effects of time-on-task and boredom due to low task load will be an important consideration for the design of new nuclear power plant control rooms, which will rely more heavily on automation. This paper describes a research study of performance in a simulated nuclear control room environment, where 36 participants responded to an alarm during a 4 hour long experiment where the alarm onset time and the availability of distractions were varied. The results indicate that operators perform better in a sterile environment and that the duration of non-active time before the alarm influences operator performance.

INTRODUCTION

The nuclear industry in the United States, and specifically nuclear power plant control rooms, are undergoing extensive modernization. Recent initiatives promise construction of advanced power plants over the coming years (Schmidt, 2010) which will likely contain new controls and rely more on automation and dynamic information display methods. The operators' role in the new control room will become increasingly passive, in which the operators' primary task will be monitoring the system through electronic displays. The operators will have to perform well in long duration, low task load environments in which rapid but infrequent shifts to high task load may be necessary. Understanding the effect of increased automation in these safety-critical environments such as nuclear power plants is of vital importance but has not been extensively studied within the fields of Human-Computer Interaction or Human Factors.

In long duration, low task load, safety-critical operating environments, the switch from a passive monitoring state to an active alarm resolution status is a critical event. The operator must maintain high awareness of the current operating conditions to quickly identify the problem then respond with the appropriate actions. But operators encounter these high taskload situations relatively infrequently and they may occur at any time during a long working shift, including near the end of a shift when time-on-task may play a role in performance. Previous studies have shown that time-on-task degrades monitoring performance due presumably to lapses in attention (Schroeder, Touchstone, Stern, Stoliarov, & Thackray, 1994; Thackray & Touchstone, 1988). Cognitive abilities such as planning may also be affected by low workload conditions (Rogers & Nye, 1993), so it is likely that time-on-task before the critical switch to high task load will also affect the operator's ability to respond to the events properly. Furthermore, these previous studies examined relatively short periods of monitoring (< 2hrs). Long duration monitoring studies, reflecting current operations in many control rooms, and the effects of boredom have received very little focus (Fisher, 1993; Straussberger & Schaefer, 2006, 2007). Therefore, investigating the influence of the time spent monitoring the system before a high workload event occurred, on the order of hours, is of great importance when considering future control rooms.

The nature of the low task workload period may also affect the operator's ability to respond to the high workload event. In current control environments, (airplane cockpits or nuclear power plants, for example), it has been assumed that it is best that operators have as few distractions available as possible, usually by implementing a "sterile" control environment. For example, the FAA implemented a rule in 1981 to ensure sterile cockpits after accidents occurred due to distractions being available. Even within the sterile control environment, however, operators may still find and devote attention to non-operationally relevant tasks or objects (Barnes & Monan, 1990) With longer time-on-task, it is possible that the allocation of their attention may shift to any available distractions during the period preceding the high workload event.

Therefore, the main goals of this research are to determine if the critical event onset time, the operating condition, or the attention state distribution during the monitoring period have a significant correlation with operator performance, specifically in a long duration, low task load environment in which operators switch from a passive monitoring state to an active control state.

METHOD

To answer these operational questions, a low task load experiment was conducted in the context of a nuclear power plant control room. The experiment was designed as four-hour control task, to replicate longer duration control room shifts. To facilitate this experimental design, a PC-based nuclear power plant control room simulator was developed, called Human Operator Monitoring of Emergent Reactors (HOMER). HOMER was designed and implemented to represent an individual nuclear power plant control task. The primary task of HOMER users was to monitor the main interface and ensure the reactor was functioning properly and to respond to any alarm events as quickly as safely as possible.

Experiment Simulation (HOMER)

HOMER's main interface displays four control loops that form the nuclear power plant (Figure 1). The interface has both graphical and alphanumeric symbols that are based on current industry uses and standards. The interface is designed to allow the operator to effectively monitor the state of the reactor. Alarms are directly visible at the center top of the display and the state of the valves and pumps can be controlled directly by clicking on them. One of four different alarm conditions could occur during the experiment and the subjects followed procedures contained in a notebook binder to resolve the alarm condition. The detailed step-by-step procedures were designed to ensure the proper response in the event of an alarm condition and be completed in approximately 20 minutes.

A separate monitor, positioned to the right of the primary display, displayed other simulation features of HOMER, including a secondary task, pop-up windows for entering periodic subjective ratings, and alternate distractions. The secondary task is a single reactor loop similar to one of the loops in the primary display, with which the participants could practice interacting with the interface. Every 30 minutes, a pop up window appeared on the secondary display with subjective ratings to be filled in by the participant. Potential distractions, like an Internet browser, could also be opened then used on the secondary display during the non-sterile condition.

Experimental Treatments

The operational questions were manifested in two specific experimental treatments: operating condition and critical event onset time. The operating condition was varied to investigate the influence of the environment on operator performance. Half of the subjects operated in a sterile condition, where they were not allowed access to any activity that was not directly related to the task. The other half could access the Internet and were allowed to read or use their own electronic devices (e.g., laptops or cell phones) or books. Three critical event onset times were implemented (1:30, 2:30 or 3:30 hours) to investigate the effect of the monitoring duration on the performance of resolving the alarm. Subjects were equally divided among the six possible treatment conditions. The primary performance measurements were the success rate in clearing the alarm and the time to clear the alarm. Attention states were captured by video cameras mounted at the workstations and measured by the percentage of time spent directed at the primary interface, distracted from the primary interface, and divided between the primary interface and another task. The attention states were determined by coding the participant's gaze direction: (1) directed at the primary interface, (2) divided between the primary interface and something else, or (3) completely

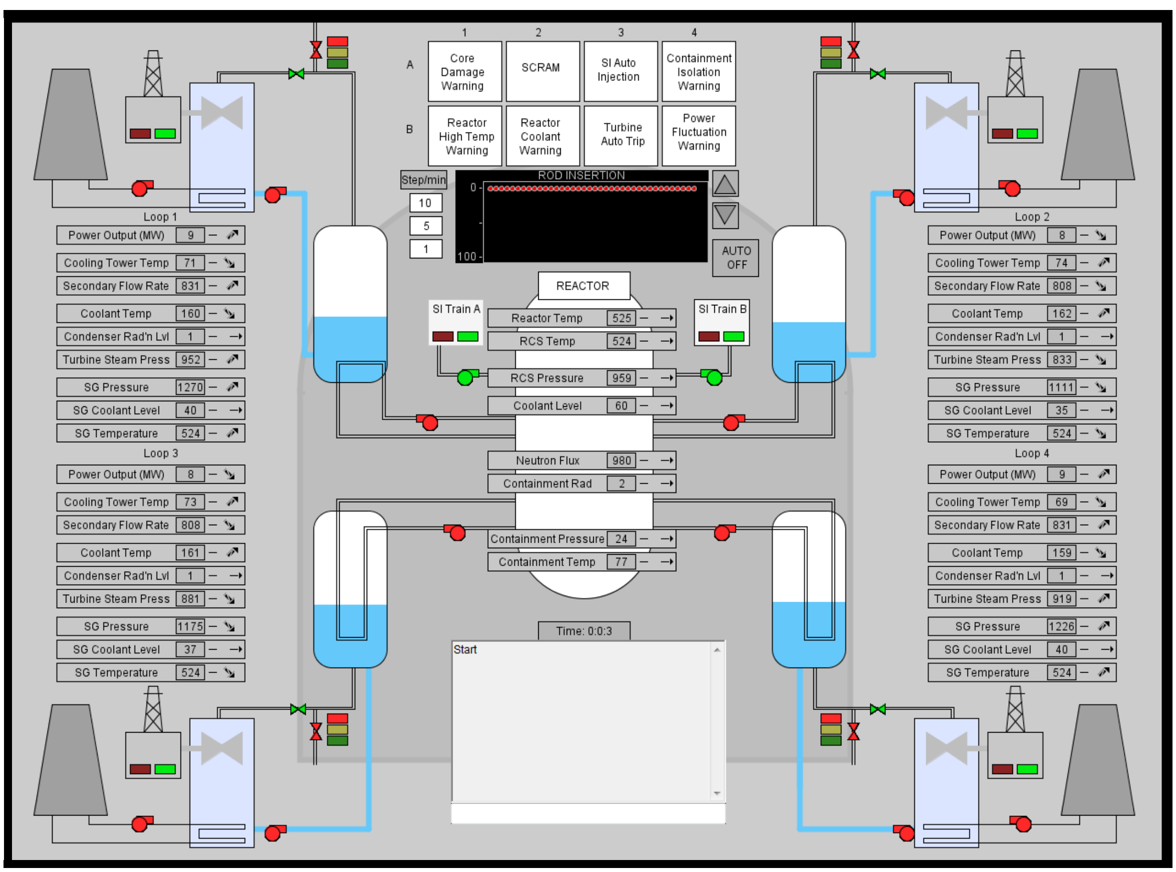

Figure 1: Main HOMER interface depicting four reactor loops. Alarm panel is at the center top of the display. The chat box is at the center bottom of the display.

distracted and unable to see the primary interface.

Experimental Setup

Thirty-six individuals participated in the experiment (15 female, 21 male), with a mean age of 22.6 years (18 years to 29 years old). None of them reported any nuclear power plant experience, but all of them reported some comfort in using computers. Participants were randomly placed in groups of three within the experiment room to simulate the possible social interactions within a nuclear power plant control room but each participant was tasked to monitor and control only his or her own individual reactor through his or her own HOMER simulation workstation. Each participant in the group was given a different alarm condition, so they would not be able to help each other. The participants completed several different activities over the course of the four-hour experiment. The primary task was to monitor the primary control interface for an alarm condition. Participants were instructed to follow the appropriate paper-based procedure to resolve any alarm situations that may arise as fast as possible.

Participants also performed some secondary tasks, including recording specific parameters from the primary control interface to a secondary interface and rating their boredom, workload and fatigue on 5-point Likert scales every 30 minutes. These scales were author-generated and not taken from a standard instrument such as NASA-TLX. This interval for recording the state parameters is a standard procedure for several commercial reactors. On the primary interface, a chat box presented a question every 20 minutes that participants were required to answer. These questions represented typical interactions with the control room supervisor.

Procedure

The experiment began with a short briefing to inform the participants on the general rules and the purpose of the experiment. After the briefing, the participants completed a survey to record demographic and personality characteristics of each participant. Then participants were provided a self-paced training period (approximately 10-15 minutes) facilitated through training slides highlighting the components and possible interactions with the primary and secondary interfaces. They were also given instructions how to identify and use the correct emergency procedures in the event of an alarm. After training, the participants were allowed to ask any questions about the procedures. Following a five minute break, each participant was seated at a workstation with two monitors and provided with a set of paper-based emergency procedures and the experiment began. After the 4 hour experiment session, the participants recorded their feelings about their performance in another survey. Each participant was given $200 in compensation and was eligible to receive a $250 bonus for the best performance.

RESULTS

Overall, 22 out of the 36 participants (61%) were able to clear the alarm in an average time of 51:12 (mm:ss). In the sterile operating condition, 14 participants (77%) cleared the alarm in an average time of 41:27 (mm:ss) whereas only 8 participants (44%) in the non-sterile condition were able to clear the alarm, taking an average of 68:17 (mm:ss). The success rate for participants in the sterile environment was significantly higher than the non-sterile environment (Wald = 3.986, $p = 0.046$), shown by a logistic regression while the average time to clear the alarm event approached significance ($t(20) = -2.006$, $p = 0.059$).

The time required to clear the alarm is significantly different across onset times ($F(2,19) = 6.355$, $p = 0.008$) (Figure 2). A Tukey post-hoc analysis revealed that the alarm clearance time is significantly different between both the first and second onset time ($p = 0.046$) and the first and third ($p = 0.014$). Participants in the first critical onset time group succeeded more often than the groups with onset time 2 and 3 (Wald $X^2 = 6.518$, $p = 0.011$) with completion rates of 92, 58, and 33.3% respectively. Completion rates were higher in the sterile condition for onset time 2 (83% vs 33%) and onset time 3 (50% vs 17%) although the differences were not statistically significant. These completion rates for the 3rd onset group are likely biased by the fixed experiment duration of 4 hours that limited the group to a maximum of thirty minutes to clear the alarm. The fact that most participants took much longer than the expected 20 minutes to resolve the alarm would also further reduce the group's success rate. We suspect that the short training time and confusing format and language of the procedures increased the variability in performance between subjects. Nevertheless, even the additional hour on task for the 2nd onset time groups was sufficient to reduce the success rate especially for the non-sterile condition. This suggests an interesting interaction that distractions have a greater detrimental impact on performance as the time on task increases.

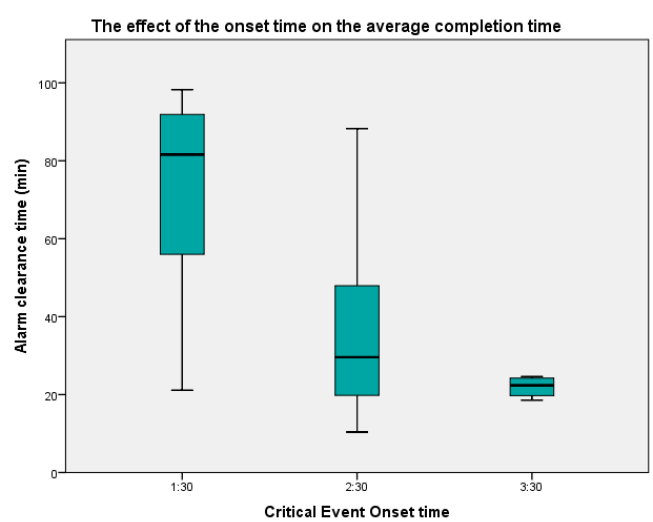

Figure 2: Completion times for each critical event onset time group

Analyses of the distribution of attention indicate that subjects operating in the non-sterile environment spent more time in a divided attention state ($F(1,33) = 7.934$, $p = 0.008$). Interestingly, no other independent variable had an influence on any attention state. The amount of time that participants spent in a certain attention state before the alarm was not a reliable indicator of their performance during the alarm ($F(1,19) = 1.874$, $p = 0.187$).

Another interesting result from the attention state analysis was that participants spent an average of 49% of their time directed solely toward the primary interface. Even considering the time spent attending to the primary interface under the divided attention state (7%), operators spent nearly half of their time attending to something completely unrelated to the task (44%). For subjects in the non-sterile condition, they obviously spent time with their devices or books, but subjects in the sterile condition often simply diverted their attention to other parts of the laboratory environment. Other research has demonstrated similar divisions among attention states in long duration, low task load supervisory control environments (Hart, 2010).

CONCLUSIONS

Operator performance, in the context of our simulated control room environment, is indeed influenced by the operating condition, where a sterile environment promotes better performance as compared to a non-sterile environment. These results support the widely held assumption that the minimization of distractions available to the operators of nuclear power plants has a positive effect on human performance.

However, the results also suggest that the duration of the non-active time before an alarm condition influences operator performance. A longer non-active time of just one hour has a negative impact on the operators' performance success rate, especially in the condition where distractions are present. This suggests that it may be favorable to have shorter working shifts, or frequent breaks for operators, though the effects of those mitigations in a long duration, low task load environment require further examination. The analysis of the attention states showed that the amount of time spent in a specific attention state before an alarm condition does not influence the performance during the alarm condition. Further detailed analysis is needed to see if an interaction between distractions and time on task is also reflected in the distribution of attention.

The overall conclusion regarding the design of more modern nuclear power plant control rooms or any long duration, low task load control environments is that it may be favorable for operator performance to minimize the amount of distractions potentially available for the operators, while keeping working shifts relatively short. The attention states of operators do shift slightly depending on the operating conditions, however, these relatively small changes do not reliably predict performance.

ACKNOWLEDGMENTS

This research was sponsored by the Nuclear Regulatory Commission. This research does not represent official policy, direct or implied, of the Nuclear Regulatory Commission and is solely the opinion of the authors. We would like to thank Dr. Amy D'Agostino, Prof. M. L. Cummings, Dr. Charles Oman and the members of the MIT Humans and Automation Laboratory that helped make this research possible.

REFERENCES

Barnes, V. E., & Monan, W. P. (1990, October). *Cockpit Distractions: Precursors to Emergencies.* Paper presented at the Human Factors Society 34th Annual Meeting.

Fisher, C. D. (1993). Boredom at work: A neglected concept. *Human Relations, 46*(3), 395-417.

Hart, C. S. (2010). *Assessing the Impact of Low Workload in Supervisory Control of Networked Unmanned Vehicles.* (S.M.), Massachusetts Institute of Technology, Cambridge, MA.

Rogers, M. D., & Nye, L. G. (1993). Factors Associated with Severity of Operations Errors at Air Route Traffic Control Centers. In M. D. Rogers (Ed.), *An Examination of the Operational Error Database for Air Traffic Control Centers* (pp. 243-256). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Schmidt, J. (2010). $8.3B in loan guarantees for nuclear reactors, from http://www.usatoday.com/tech/news/2010-02-16-obama-nuclear-power-plant_N.htm

Schroeder, D. J., Touchstone, R. M., Stern, J. A., Stoliarov, N., & Thackray, R. (1994). Maintaining Vigilance on a Simulated ATC Monitoring Task Across Repeated Sessions (pp. 14): Civil Aeromedical Institute, Federal Aviation Administration.

Straussberger, S., & Schaefer, D. (2006). Monotony in Air Traffic Control - Contributing Factors and Mitigation Strategies (pp. 272): EUROCONTROL Experimental Centre.

Straussberger, S., & Schaefer, D. (2007). Monotony in Air Traffic Control. *International Journal of Engineering and Operations, 15*(3), 183-207.

Thackray, R. I., & Touchstone, R. M. (1988). An Evaluation of the Effects of High Visual Taskload on the Separate Behaviors Involved in Complex Monitoring Performance (pp. 13): Civil Aeromedical Institute, Federal Aviation Administration.