Naval Research Laboratory
Washington, DC 20375-5320

NRL/MR/5580--19-9886

Evaluating and Optimizing Operator Performance within Supervisory Control Environments

Cyrus K. Foroughi

Information Management and Decision Architectures Branch
Information Technology Division

May 28, 2019

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
28-05-2019

2. REPORT TYPE
NRL Memorandum Report

3. DATES COVERED *(From - To)*
Oct 2017 - Oct 2018

4. TITLE AND SUBTITLE
Evaluating and Optimizing Operator Performance within Supervisory Control Environments

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
NISE

5d. PROJECT NUMBER
55-N2N2-08

5e. TASK NUMBER

5f. WORK UNIT NUMBER
N2N2

6. AUTHOR(S)
Cyrus K. Foroughi

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/MR/5580--19-9886

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375-5320

10. SPONSOR / MONITOR'S ACRONYM(S)
NRL/NISE

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This is the final report for Cyrus Foroughi's Karles Fellowship. This report highlights the major findings and accomplishments of the work.

15. SUBJECT TERMS

Supervisory control	Automation	Automation failures
Workload	Attention	Eye-tracking

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Cyrus K. Foroughi
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited	Unclassified Unlimited	6	(202) 767-3879

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

This page intentionally left blank.

Evaluating and Optimizing Operator Performance within Supervisory Control Environments

Objective:

The task demands placed on current naval operators has changed dramatically over the past decade. The recent shift towards unmanned systems (e.g., UAVs) where operators are monitoring multiple tasks across multiple screens raises important questions about human limitations in attention and workload. Such limitations are a concern to the Navy because mission performance will suffer when operators are not able to attend to the proper tasks and/or when operators are cognitively overloaded. Thus, the goals for this research were 1) to evaluate operator performance in supervisory control environments with measures of workload and attention, and 2) to make recommendations about how operators should be allocated/assigned to optimize the operator-system interaction.

Technical Approach:

I developed a series of experiments to evaluate operator performance with measures of attention and workload across a wide range of different supervisory control scenarios. I used the U.S. Naval Research Laboratory's Supervisory Control Operations User Testbed (SCOUT; Sibley, Coyne, & Thomas; 2016) and Simulated Multiple Asset Routing Planning Testbed (SMART) to manipulate many performance-relevant features including workload, stress, fatigue, and boredom. Further, I used eye-tracking to get real-time measures of attention (i.e., gaze) and workload (i.e., pupillometry).

Benefit to Science and Technology:

This work was in line with the 2011 Naval S&T Plan, which specifically requested research to better understand "task allocation/assignment, planning, coordination, and control for heterogeneous systems." By evaluating operator performance with measures of workload and attention in a supervisory control environment, an environment similar to those currently employed by the Navy, we now have a better understanding of task performance reductions as a result of limitations in human attention and workload.

Summarized Findings/Accomplishments:

This work resulted in multiple refereed papers and conference publications that increased our understanding of 1) how humans interact with emerging automated technologies, 2) how humans make decisions in supervisory control environments, and 3) how eye-tracking can be used to assess where humans direct their attention in supervisory control environments.

Manuscript approved January 10, 2019.

Findings:

During my one-year Karles' Fellowship, I created and worked on existing lines of research that were aimed at bettering our understanding of how humans interact with emerging technologies, specifically, supervisory control environments. Below I will highlight the findings from these projects.

Given the rapid growth of automation throughout the DoD, there is an immediate need to better understand how humans monitor multiple automated tasks concurrently. In one study, participants completed an experiment within SCOUT where they were to monitor three automated sensor search feeds. The automated sensor search feeds were set to different levels of reliability (viz., 70%, 85%, and 95%). Eye-tracking data were collected to provide a real-time index of attention allocation and workload. The performance data revealed that participants' failed to detect Automation Misses approximately 2.5 times more than Automation False Alarms across both conditions. The eye data revealed that participants spread their attention equally across all three of the automated sensor feeds for the duration of the experiment. The eye data provide objective, attention-based data in support of a system-wide trust approach as the default position of an individual monitoring multiple automated tasks. To my knowledge, this is the first study to provide objective, attention-based support for system-wide trust theory. Surveys and subjective measures have been previously used.

Given that automated systems of the future are likely to be highlight reliable (e.g., 99.9%), there is an immediate need to understand how humans can detect very rare automation failures. In a second study, participants completed a different experiment within SCOUT where they were monitoring three 100% reliable automated sensor search feeds. At two specific time points, the automation failed resulting in one Automation Miss and one Automation False Alarm. Eye-tracking data were collected to provide a real-time index of attention allocation and workload. The performance data revealed that participants' were not good at detecting the rare-event automation failures: detecting approximately 1/3 of the Automation Misses and 2/3 of the Automation False Alarms. The eye data revealed that individuals with a more complex visual search strategy were able to detect the automation failures compared to those who had a more routine pattern. Similar to the last study, research in this area is lacking and to the best of my knowledge this is the first study to use eye-tracking to assess how humans detect rare-event automation failures, thus advancing the science in this area considerably.

In another line of work, I investigated how individuals make decisions involving risk and uncertainty within a spatially focused route planning task involving multiple simulated unmanned vehicles and objectives. Participants were instructed to create twenty-four route plans within SMART, half of which included a combination of risky route suggestions and enhanced icons. No performance differences were detected in providing enhanced icons/visualizations. Interestingly, participants overall performance improved when provided risky, sub-optimal route suggestions. These results could have occurred due to multiple factors, including the multi-objective route planning task being too complex or the scenarios lacking sufficient variability. An additional interpretation of these findings is that humans are really poor at performing these complex multi-objective tasks, and failed to comprehend the uncertainty and risk inherent in their decisions.

Another line of work was aimed at assessing the viability of using low-cost eye-tracking to help understand human behavior and attention. The recent availability of low-cost eye-tracking hardware provides researchers a fruitful opportunity to collect additional human subject data for under $1000. This study was aimed at investigating whether low-cost eye-tracking is capable of replicating a large effect showing a relationship between resting pupil size and working memory capacity. Navy and Marine Corps student pilots participated in this study and we were granted access to their aviation selection test scores. The study demonstrated the capability of the Gazepoint GP3 system to detect the pupillary light reflex within every participant. However, in contrast to findings from other researchers, analyses revealed a negative correlation between resting pupil size and partial Operation Span scores, and no correlation between resting pupil size and two cognitive components of the aviation selection test. These findings, in addition to other reasons discussed herein, suggest that the Gazepoint GP3 system's millimeter pupil size measurements should not be used in isolation to compare values between individuals. They also suggest the need for further investigation of the relationship between baseline pupil size and working memory capacity.

Overall, this work highlighted the need to understand limits in human attention. Humans have real limitations and systems need to be designed to accommodate for these limitations. Assuming that humans will automatically fill any new role that spawns out of the emergence of technology is a naïve thought that is not recommended. All of these projects have ongoing follow-up studies that further aim to tease out many of the nuanced findings.

Publications/Presentations:

Coyne, J., Foroughi, C., Brown, N., & Sibley, C. (2018, September). Evaluating decision making in a multi-objective route planning-task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 217-221). Sage CA: Los Angeles, CA: SAGE Publications.

Sibley, C., Foroughi, C., Brown, N., & Coyne, J. T. (2018, September). Low Cost Eye Tracking: Ready for Individual Differences Research? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 741-745). Sage CA: Los Angeles, CA: SAGE Publications.

Foroughi, C. K., Ambinder, M., Armijo, A., Kozachuk, J., & Woo, H. (2018, September). Human Factors and Video Game Research: A Discussion Panel. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 2047-2050). Sage CA: Los Angeles, CA: SAGE Publications.

Foroughi, C.K., Devlin, S., Sibley, C., & Coyne, J. T. (submitted). Rare Event Automation Failures in a Simulated Supervisory Control Environment. *International Symposium on Aviation Psychology*.

Foroughi, C.K., Sibley, C., Rovira, E., Pak, R., & Coyne, J. T. (submitted). Detecting Automation Failures in a Simulated Supervisory Control Environment: Evaluating System-Wide and Component-Specific Trust Theories. *Ergonomics*.

Foroughi, C.K., Devlin, S., Sibley, C., & Coyne, J. T. (manuscript in preparation). Attention Allocation When Detecting Rare Event Automation Failures. *Journal of Experimental Psychology: Human Perception and Performance*.