Assessing Cognitive Load and Usability for CEMA Training Using COBWebS

CDT Mason B. Rockman, CDT Keenan W. O'Shea, Robert H. Thomson, Ph.D.
United States Military Academy
West Point, NY 10996
mason.rockman@westpoint.edu, keenan.o'shea@westpoint.edu, robert.thomson@westpoint.edu

Michael W. Boyce, Ph.D., Nathan L. Vey, J. Allen Geddes
Combat Capabilities and Development Command – Soldier Center (CCDC-SC)
Simulation and Training Technology Center (STTC)
12423 Research Parkway, Orlando, FL 32826
michael.w.boyce11.civ@mail.mil, nathan.l.vey.civ@mail.mil, james.a.geddes2.civ@mail.mil

Mark K. Rusboldt
Program Executive Office for Simulation, Training, and Instrumentation (PEO STRI)
12211 Science Drive
Orlando, FL 32826-3224
mark.k.rusboldt.civ@mail.mil

LTC Chad T. Bates, Ph.D., CW3 Eric T. Colon
United States Army Cyber Command
8825 Beulah Street, Fort Belvoir, VA 22060
chad.t.bates.mil@mail.mil, eric.t.colon2.mil@mail.mil

Keywords:
CEMA, Training, West Point, Cognitive Load, Usability

ABSTRACT: *The Army's Modernization Priorities describe a need for network hardware, software, and infrastructure to allow soldiers to fight when the electromagnetic spectrum is denied or degraded, as well as the need to improve human performance and decision making through training starting at the soldier level through the Synthetic Training Environment. Modeling and Simulation technologies can provide an avenue to support this type of training, especially in the growing area of Cyberspace and Electromagnetic Activities (CEMA). Cyber Operations Battlefield Web Services (COBWebS) is a training technology that can provide Cyber effects to the end user. This research aims to provide information about the effectiveness of the COBWebS user interface design to support CEMA training. The study is a repeated-measures design where participants will experience two different levels of information (high information = very detailed, low information = streamlined version of information) in a counterbalanced design. Participants will be given an overview of the purpose of COBWebS and the type of tasks they will be doing during the study. They will then be given a set of tasks to complete using COBWebS. The participants will use a talk-aloud protocol where they are talking through their decisions and what they are looking at, take subjective usability and workload questionnaires, as well as undergo a semi-structured interview at the end of each session. The hypothesis is that as the level of complexity increases, the participants will exhibit attentional narrowing, focusing on specific areas most relevant to the task at hand.*

1. Introduction

A soldier is tasked with protecting critical government information from threats and being able to respond with countermeasures. In the cyberspace domain, this is known as offensive and defensive cyberspace operations. Some would call this cyber warfare, others would call it Cyberspace Electromagnetic Activities (CEMA), while others call it Information Operations or Information Warfare. Regardless of the term, the important aspect for Army Modeling and Simulation (M&S) is the information. Information may come in a variety of formats and displays, but independent of the platform, the question becomes: how can we make sure our soldiers are properly trained to detect and mitigate potential threats?

An existing Army program looking to train for CEMA is One Semi-Automated Forces (OneSAF). OneSAF is a computer forces simulation provided by the Program Executive Office for Simulation, Training and Instrumentation (PEO STRI) that provides entity-level models and behaviors that are both semi-automated and fully automated applications designed to achieve Army readiness. OneSAF supports the training, test and evaluation, analysis, intelligence, acquisition and experimentation communities by providing the latest physics-based modeling and data collection, and reporting capabilities. OneSAF models real-world representations of platforms, soldiers, equipment, logistical supplies, communications systems and networks, emerging threats, and aviation assets to achieve the level of fidelity required for an application or scenario. Their vision is to satisfy all US Army modeling and simulation communities and other government agencies' needs for a common, reusable, computer-generated (ground) forces entity-level simulation. [1]

Due to the growing prevalence of the cyberspace domain within the Army operational training environment, current research looks to better understand the debate on military cyber personnel competencies by investigating how cyber defense operator's level of self-regulation predicts cognitive load within different frameworks. Current cyber exercises have the capability to adjust complexity, however, it is unclear how adjusting specific parameters can change from cyber situation awareness and cognitive load. The Cyber Operations Battlefield Web Services (COBWebS) project is a step towards a larger understanding and integration of cyber training within the Army in addition to exploring the adjustment of those specific parameters.

COBWebS enables the incorporation of CEMA effects into training exercises that utilize constructive simulations, such as OneSAF, to stimulate live mission command devices. COBWebS accomplishes this by affecting tactical messages as they flow between the simulation and the tactical network, providing the white cell training staff with the ability to intercept, deny, delay, and inject forged data, resulting in the effects of CEMA attacks on the trainees' mission command devices that they are using for situational awareness and decision making purposes. The trainee must recognize that their command and control devices are under a cyber-attack, and then make decisions accordingly to mitigate the threat and still accomplish their mission. [2]

2. Literature Review

Gutzwiller et al.'s 2016 experiment can be used as a framework for the tasks (Denial of Service, Information Delay and Forged Route) we are looking to address. Gutzwiller et al. (2016) conducted a preliminary experiment that sought to create a cyber-defender task and evaluate the results. Gutzwiller et al. (2016) evaluated the impact the situation and interface had on situation awareness, but the methods and results can also be applied to cognitive load. In this experiment, the researchers came up with four considerations for their methodologies. First, to further understand their participants, they gave them a demographic survey to complete. Second, they knew that this task required experts in the field. Third, they used a knowledge audit which allowed them the flexibility to capture an uncertain environment along multiple dimensions and increase their understanding of how the participants developed awareness in the system. Fourth, they used the interactive method of concept mapping which was selected to prompt participants to visualize their knowledge [3]. These considerations can be used to help form the methodology of this experiment. This study found that the different tiers of cognitive load was hard to control due to the differences in available human resources.

Additionally, Gutzwiller (2019) wrote a paper about the importance of situation awareness (SA) [4]. In his paper, he defines SA as paying attention to a current or future situation with goals or accomplishments in mind. This paper is extremely relevant since Gutzwiller (2019) finds research that supports the idea that situation awareness can be improved with a unique interface design. Although he looks at situation awareness, this can be easily applied to

elements of cognitive load and how changes to the current interface can help improve the current status of the system and its usability for future users.

Josok et al. (2019) investigated how cyber defense operator's level of self-regulation contributed to their performance in actual operations. Josok et al. (2019) hypothesized that higher levels of self-regulation predicted higher levels of cognitive agility. They also identified that in prior research, cognitive agility had been linked to performance in defensive cyberspace operations. Therefore, the researchers expected a positive association between levels of self-regulation and displays of cognitive agility. At the conclusion of their study, data showed that higher levels of self-regulation were associated with displays of cognitive agility. Their study helps explain how understanding factors contributing to cyber operator performance are needed to improve education and training programs for military cyber personnel. Their study also identifies self-regulation as an important contributing factor to cognitive agility and how it can be a key determinant in evaluating individuals' cyber operating performance. [5]

Knox et al. (2018) presents how methods of Slow Education and Accelerated Learning could lead to greater Cognitive Agility in cyber defense operations [6]. Their study builds on earlier empirical and theoretical human factors research in cyber defense. Particularly the Hybrid Space Framework, Cognitive Agility, and the Orientate-Locate-Bridge model (OLB) for socio-technical communication. The Hybrid Space Framework visualizes the intersection between cyber-physical and strategic-tactical dimensions, which allows for the application of psychological concepts in training. Cognitive agility, as they outline in their paper, demonstrates an individual's metacognitive ability to understand, monitor, and regulate the use of flexible cognitive strategies that help performance on a given task. The OLB model dissects the steps of improved grounded communication based on shared mental models in complex hybrid environments. These earlier models and studies help provide a common framework for cognitive processes that can contribute to an improved understanding of governance in the cyberspace domain. Their article focuses on the impact of how Slow Education approaches to cyberspace domains can support improved sensemaking and understanding. Their study ultimately furthers the discussion relating to cognitive agility and adaptations to cyber education capable for improving cyberpower understanding and control.

Because our experiment will be focusing on the impact of different levels of information on cognitive load, we looked at an article provided by Paas et al. [7]. Their article looks at Cognitive Load Theory: the interaction between information structures and cognitive architecture. Cognitive load theory is mainly concerned with the learning of complex cognitive tasks, where learners are often overwhelmed by the number of information elements and their interactions that need to be processed simultaneously before meaningful learning can occur. The main focus of CLT is the instructional control of this high load and understanding how to attain meaningful learning in complex cognitive domains. Thus, the theory suggests that learning happens best under conditions that are aligned with human cognitive architecture. Understanding this concept will help us determine the correct amount of information to display for each of the three levels we are going to present to the user.

3. Research Question and Hypotheses

This research will look specifically at the COBWebS system, which can simulate various cyber-attacks on command and control communications. Phase I of this project will look to focus specifically on cognitive load and its impact on user's ability to complete tasks in an accurate and timely manner. The independent variable of interest is the level of information, which will be comprised of two levels. Both levels will differ in the amount of information provided to a user (low and high information), with the intent of impacting a user's cognitive load as it pertains to time on task and accuracy with each of the scenarios. This project will focus on three different scenarios which we will put the participants through. Those scenarios are: Denial of Service, Forged Route, and Information Delay. Phase II of this project will look to apply a trust scale to help build on the research done during phase I. In addition, phase II will look to help provide qualitative feedback on the current framework and capabilities of COBWebS. The end state of this project is to conduct an empirical investigation to provide recommendations for future COBWebS development.

For this project, we have developed two key hypotheses, one pertaining to each of the information loads. We believe that participants will perform worse in the low information load scenario compared to their performance in the high information load scenario. This assumes that the low information scenario will decrease cognitive load and decrease accuracy; whereas, the high information scenario will increase cognitive load and accuracy. Overall, with the increase

of information provided to the user, cognitive load will increase, ultimately reducing time on task and an increase of accuracy when completing the required task for each scenario.

In this experiment we will be utilizing the System Usability Scale to help gather feedback in the system. To validate the proficiency of this scale, we looked at a study by Borsci, Federici, & Lauriola [8]. In their study, they verified the two-component structure of usability and learnability. They viewed whether practitioners would take advantage of these components and extract additional information from the data the scale provided. They concluded that there could not be more data extracted than what the scale was meant to measure. Thus, validating the use of this scale. Additionally, we will also run the participants through a NASA-TLX scale which will measure the perceived cognitive load of the user. In order to validate the legitimacy of this scale, we reviewed an article by Hart [9]. In this article she tests the credibility of the scale and determines it is a quick and effective way to measure the perceived cognitive load of the user after completing a series of questions aimed at determining the perceived mental and physical loads of the tasks in which the participant conducted.

In order to give additional feedback on the system interface, we completed a heuristic evaluation using Neilsen's ten usability heuristics. Nielsen's ten usability heuristics are:

- Visibility of system status
- User control and freedom
- Consistency and standards
- Error prevention
- Flexibility and efficiency of use
- Help and documentation
- Recognize, diagnose, and recover from errors
- Aesthetic and minimalist design
- Match between system and the real world
- Recognition rather than recall

Our heuristic evaluation focused on the three tasks in which we plan to test our participants on (Denial of Service, Forged Route, and Information Delay). In our evaluation we found that the heuristic that was most violated was the

visibility of the system status. In other words, the system did not always keep the users informed about what is going on within the interface in a reasonable amount of time. We believe that this is due to the fact the COBWebS is tailored more towards experts. This is extremely important to keep in mind since our targeted group is West Point Cadets, who have minimal-to-no experience or familiarity with this system.

Another major heuristic that was violated was error prevention. Error prevention calls for there to be a good enough design that operator errors don't happen, and the system eliminates error-prone conditions. A major contributing factor to this heuristic violation is the startup of COBWebS. There are many different programs that need to be initiated for the system to run and work properly. The system should be updated to have an autonomous startup process to get it to work properly. As discussed, this information is setting the framework for future studies on COBWebS and allowing the developers to improve upon the heuristic violations. The table below contains the list of Nielsen's usability heuristics with the three tasks which will be tested during this experiment. The letter 'Y' indicates a violation of that heuristic for each task that we determined during our evaluation.

	General Evaluation	Denial of Service	Forged Route	Information Delay
Visibility of System Status	Y	Y	Y	Y
User Control and Freedom	Y		Y	
Consistency and Standards	Y			
Error Prevention	Y		Y	
Flexibility and Efficiency of Use		Y		
Help and Documentation	Y			
Recognize, Diagnose, and Recover from Errors		Y		
Aesthetic and Minimalist Design	Y			
Match Between System and the Real World				
Recognition Rather than Recall	Y	Y		Y

Table 1: Heuristic Evaluation Results

4. Method

4.1 Participants

For this study we will be using a within-subjects design with two different levels of information load (low and high information load). The testing population will be Cadets from the United States Military Academy. The goal is to have at least 20 participants for each group, making at least 40 participants in total. We will be utilizing the Department of Behavioral Sciences and Leadership's participant recruitment system to recruit Cadets. The participant will be from PL100 - Intro to Psychology for Leaders course at West Point.

4.2 Materials

For this experiment, everything will be conducted in one of our Engineering Psychology laboratories using a Dell Laptop loaded with the COBWebS system. In addition, the NASA-TLX and System Usability Scale will be conducted on a separate computer for information collection. Information load will be presented in hard copy format using a pre-written script. No other materials will be required to conduct this experiment.

4.3 Procedure

Participants will walk into the testing area and receive an initial familiarization with the COBWebS system prior to conducting the experiment. After the participant feels comfortable and familiarized with the system, he or she will conduct the experiment on one of the two levels of information. Once the participant has successfully completed the three tasks from the script, he or she will conduct a NASA-TLX to help determine the cognitive load of the user during the experiment. The participant will then provide us some qualitative feedback by completing the system usability scale which will help determine the overall ease of the system. Lastly, we will finish the experiment by asking the participant a series of five questions to help provide feedback pertaining to the overall liking of the system. The questions asked are as follows:

- What did you like about the system?
- What did you find difficult in using the system?

- If at any point during the experiment, where did you get stuck?
- What would you change in the system, if anything?
- Based on the information provided, did COBWebS do what you expected it to do?

5. Future Work / Conclusion

There are many ways in which future work can be applied to the COBWebS interface. This project is specifically looking at the operators behind the interface. If this system is going to be implemented into the daily training schedule for soldiers in the United States Army, then it should be looked at and further developed from the user's perspective as well. As mentioned earlier, future works for this project will look to introduce a trust aspect to help gather more empirical data on the COBWebS project, which in turn will help us provide additional recommendations for future COBWebS development. This project will hopefully provide COBWebS project managers and developers with valuable information in ways that help determine upgrades or enhancements to the current system. Our heuristic evaluation was the first step in providing general feedback on the current usability of the system. After completion of the first phase data collection, COBWebS project managers and developers will have a better understanding of using the system from a very basic/unfamiliar perspective. This data should provide developers with knowledge and ideas as to how they can make the system more applicable to a wider audience. Phase II data collection will provide developers with more empirical data pertaining to trust aspects and a user's willingness to rely on the system for training purposes.

6. References

[1] PEO STRI: "One Semi-Automated Forces (OneSAF)" Retrieved from https://www.peostri.army.mil/onesaf, December, 2019.

[2] Mize, J., H. Marshall, M. Hooper, R. Wells, & J. Truong: "Cyber Operations Battlefield Web Services (COBWebS) – Concept for a Tactical Cyber Warfare Effect Training Prototype" Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, USA, November 2015.

[3] Gutzwiller, R. S., Hunt, S. M., & Lange, D. S.: "A task analysis toward characterizing cyber-cognitive situation awareness (CCSA) in cyber defense analysts" 2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA) (pp. 14-20), IEEE, March 2016.

[4] Gutzwiller, R. S.: "Situation Awareness in Defensive Cyberspace Operations: An Annotated Bibliographic Assessment Through 2015 (No. TR-3184)" NIWC Pacific, San Diego, USA, 2019.

[5] Jøsok, Ø., Lugo, R., Knox, B. J., Sütterlin, S., & Helkala, K.: "Self-regulation and cognitive agility in cyber operations" Frontiers in psychology, 10, 2019.

[6] Knox, B., Lugo, R., Helkala, K., Sütterlin, S., & Jøsok, Ø: "Education for cognitive agility: improved understanding and governance of cyberpower" European Conference on Cyber Warfare and Security pp. 541-XII), Academic Conferences International Limited, June 2018.

[7] Paas, F., Renkl, A., & Sweller, J.: "Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture" Instructional science, 32(1), 1-8, 2004.

[8] Borsci, S., Federici, S., & Lauriola, M.: "On the dimensionality of the System Usability Scale: a test of alternative measurement models", Cognitive processing, 10(3), 193-197, 2009.

[9] Hart, S. G.: "NASA-task load index (NASA-TLX); 20 years later" Proceedings of the human factors and ergonomics society annual meeting, Vol. 50, No. 9, pp. 904-908, Sage CA: Los Angeles, CA: Sage publications, October 2006.