Poster: Development of Situation Awareness Measurement for Cybersecurity Professionals

David Schuster
Department of Psychology
San José State University
San Jose, United States
https://orcid.org/0000-0002-2698-7654

Crystal Fausett
School of Information
San José State University
San Jose, United States
crystal.fausett@sjsu.edu

Maiyi Huang
Department of Industrial and Systems Engineering
San José State University
San Jose, United States
maiyi.huang@sjsu.edu

Sabina M. Patel
Human Factors and Behavioral Neurobiology
Embry-Riddle Aeronautical University
Daytona Beach, United States
https://orcid.org/0000-0003-0707-0602

Jenna Korentsides
Human Factors and Behavioral Neurobiology
Embry-Riddle Aeronautical University
Daytona Beach, United States
https://orcid.org/0000-0002-9833-7036

Joseph R. Keebler
Human Factors and Behavioral Neurobiology
Embry-Riddle Aeronautical University
Daytona Beach, United States
https://orcid.org/0000-0003-2246-7472

Elizabeth H. Lazzara
Human Factors and Behavioral Neurobiology
Embry-Riddle Aeronautical University
Daytona Beach, United States
https://orcid.org/0000-0002-3495-0595

2025 Silicon Valley Cybersecurity Conference (SVCC) | 979-8-3315-3429-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/SVCC65277.2025.11133655

***Abstract*— Better understanding of the implications of many aspects of human behavior, especially that of cybersecurity professionals, can help develop the cybersecurity workforce. Despite efforts aimed at documenting and understanding cybersecurity professionals' knowledge, understanding of how cognition supports human performance in specific cybersecurity tasks remains limited. Specifically, understanding of the elements of situation awareness (SA), defined as goal-directed knowledge, is necessary to support human-centered evaluation, selection, training, and recruitment strategies. In this poster, we propose a framework for developing measures of situation awareness for cybersecurity professionals and a method of capturing initial effectiveness data that does not rely on access to proprietary information. While this will not be a complete solution to the problem of limited access, we aim to shorten the path between observation and measurement by demonstrating a process of creating SA measurement from a ransomware simulation.**

Keywords*— *situation awareness, cybersecurity professionals, human performance measurement

I. Introduction

The effective work of cybersecurity professionals is critical to maintain security across the organization. Even in environments of increasing automation, human cognition and decision making remains essential. Therefore, understanding of the implications of cybersecurity professionals' behavior complements technical approaches to cybersecurity solutions. The NICE Workforce Framework for Cybersecurity [1] provides a starting point to understand cybersecurity professionals' knowledge with the goal of supporting recruitment, selection, training, planning by documenting the tasks, knowledge, and skills associated with cybersecurity roles [1]. The NICE Framework continues to develop iteratively through consultation with diverse stakeholders. While enumerating tasks, knowledge, and skills is of great benefit, understanding of how these aspects of human performance develop and support proficiency remains limited. In other words, without defining the knowledge and skills required for specific tasks and environments, one cannot know what to train, what criteria to use to select candidates, or what populations to target for recruitment. Definition is the first step, and the next is to be able to reliably and meaningfully assess human performance. We suggest that the study of knowledge applied to cybersecurity tasks is, therefore, necessary, as it can lead to validated quantitative assessments in aid of the goals of the NICE Framework.

Specifically, understanding of the elements of situation awareness (SA), defined as goal-directed knowledge [2], is necessary to support human-centered evaluation, selection, training, and recruitment strategies. In a widely cited model by Endsley [3], SA involves perception of relevant elements, comprehension of their meaning, and projection of their status into the future.

Because measurement of SA involves knowledge required for a task, the best measures are task specific, increasing the cost of their development. Further, SA measures of operational environments typically require a high level of access to observe and interview cybersecurity professionals, which is challenging given the sensitive nature of security operations data. A solution provided by [4] was to serve as both employee and researcher. Not having that level of access, our approach has been to observe incident response in a purpose-built simulation. In this work, we aim to shorten the path between observation and measurement by demonstrating a process of creating SA measurement from a ransomware simulation.

II. Method

A. Participants

Participants were N = 12 high school students from the western United States. All were participants in their school's cybersecurity club or events, and eight had previously participated in the school's cybersecurity team. None of the participants reported prior experience with Splunk, which was the security information and event management (SIEM) tool the simulation was based on, nor the dataset used to create the simulation.

B. Scenario

The scenario was developed based on past published contest materials from Splunk's BOSS of the SOC competition [5] . We previously described the method of constructing this scenario [6], which consisted of 38 slides describing a novice's investigation of a ransomware scenario. Each slide described one step in the investigation. The purpose of the novice's performance of the scenario was not to illustrate best practices. Instead, it was to provide a baseline for critique. In this study, the purpose of the SA measure was to capture novice participants' understanding of the developing scenario.

C. Measures

To measure situation awareness, we first created a matrix following the approach of [7]. Endsley's three levels were used as rows and five categories of knowledge were used as the columns, resulting in 15 cells. We adapted the five areas of terrain (geographical, physical, logical, cyber persona, and supervisory) from [7] to apply to a corporate environment, resulting in: Where or what affected, nature of the attack, inferred information, patterns, and critique.

We then created objective questions and an answer key, filling in information in each category as it was revealed throughout the scenario.

Following this, we populated situation awareness global assessment technique (SAGAT)-style pauses [8] with questions from each level. We planned three pauses for these questions and did not tell participants when these would occur nor which kinds of questions would be asked. The 15 items that resulted are as follows:

At 2:22 minutes, we asked:

1. What are all the indicators of an attack on the system?
2. What files were affected by this cybersecurity attack?
3. What happened to the files?
4. What information needed to understand this attack is missing?
5. What malware might this be?
6. What should you do next?

At 7:20 minutes, we asked:

7. What destinations did this workspace communicate with?
8. Exactly what type of malware is it?
9. What information needed is still unknown?
10. What type of file caused the downloads?
11. What would you do next, based on the information gathered so far?

At 10:20 minutes, we asked:

12. What were the potential ramifications of this cybersecurity attack?
13. What would happen if this attack on the system was unresolved (as in the company does nothing)?
14. What was the motive of this cybersecurity incident?
15. What can the attackers do or not do with the files?

Each item was scored dichotomously (1 or 0) by one author who did not perform the data collection. Responses that matched the answer key were scored with a 1, and all other responses were scored with a 0.

To explore the performance of our measure compared to other methods, we also included the Situation Awareness Rating Technique (SART) [9] and the NASA Task Load Index (NASA TLX) as a measure of workload [10].

D. Procedure

Participants watched a prerecorded video of the slideshow presentation. For each slide, the narrator of the video read the text of the slide aloud. At each pause, the researcher stopped the video and said, "Please answer some questions about the current state of the scenario. If you do not know, you can say so." If participants said they did not know, the researcher asked where they would look to find this information. Participants were also asked to rate their confidence from 1 (very low confidence) to 5 (very high confidence). During the session, researchers transcribed the participant's response to each item.

III. Results

Participants answered 0 to 6 of 15 items correct, with an average of 3 (M = 2.92, SD = 1.93). Examining the descriptive statistics for each item revealed that Items 2, 6, 10, and 13 were not answered correctly by any participant. The most frequently correct item was Item 5, with 51% of participants answering correctly.

Reliability was calculated for the SA (Cronbach's α = .57, n = 15 items) and confidence (Cronbach's α > .99, n = 15 items) measures. Consequently, reliability for the SA measure was poor, while reliability for the confidence measure was excellent.

We examined a correlation matrix with SA, confidence, the SART [9] SA measure, and NASA TLX workload measure [10] along with order of presentation, and participants' age. When adjusted for the number of correlations in R, none of the relationships were statistically significant, leading to an inconclusive result. Without this adjustment, there would have been significant relationships between confidence and SA (r = .63, p = .03), and between SA and the TLX (r = -.62, p = .04). It is important to note that the adjustment for multiple comparisons is conservative, our sample size was low, and these correlations were run in an exploratory fashion.

IV. Conclusion

We have demonstrated a method for creating SA measurement from a ransomware simulation. Two features of our method increase its utility; first, employing published contests can partially mitigate the need to access proprietary data to create a scenario. Of course, access to such data within an organization could lead to even more meaningful assessment. Second, our application of emerging SA frameworks from other researchers (e.g., [7]) helped us to be more systematic in our definition of KSAs.

The major limitation of this work is that our small sample size limits the evidence of this measure's reliability and validity. The reliability we observed was poor, and the pattern of correlations was inconsistent, which is logical due to low statistical power and unreliability. These problems can be resolved in future research by collecting a larger sample, including more raters, and ensuring interrater reliability. Consideration is also needed for the population against which the measure is validated; even the most favorable result would have left unanswered questions about reliability and validity with participants more diverse in their expertise.

In all, we have provided a framework that scientists and practitioners embedded in organizations may be able to leverage to bridge the gap between cybersecurity professionals' cognition and performance.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 1553018. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

[1] NIST, "Workforce Framework for Cybersecurity," NICE Framework, May 05 2023. Available: https://www.nist.gov/system/files/documents/2023/06/05/NICE%20Framework%20(NIST%20SP%20800-181)%20_one-pager_508Compliant.pdf

[2] R. Rousseau, S. Tremblay, and R. Breton, "Defining and modeling situation awareness: A critical review," in *A Cognitive Approach to Situation Awareness: Theory and Application*, S. Banbury and S. Tremblay, Eds., Burlington, VT: Ashgate Publishing Company, 2004, pp. 3–21.

[3] M. R. Endsley, "Situation awareness global assessment technique (SAGAT)," *IEEE 1988 National Aerospace and Electronics Conference*, vol. 3, pp. 789–795, 1988, doi: 10.1109/NAECON.1988.195097.

[4] H. J. Ofte, "The awareness of operators: a goal-directed task analysis in SOCs for critical infrastructure," *Int. J. Inf. Secur.*, vol. 23, no. 5, pp. 3253–3282, Oct. 2024, doi: 10.1007/s10207-024-00872-6.

[5] Kovar, R., "What You Need to Know About Boss of the SOC," [Blog post], 2023, https://www.splunk.com/en_us/blog/security/what-you-need-to-know-about-boss-of-the-soc.html.

[6] D. Mabie and D. Schuster, "Lessons Learned in Leveraging Existing Simulations for Cybersecurity Training, Evaluation, and Research," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 64, no. 1, pp. 425–429, Dec. 2020, doi: 10.1177/1071181320641095.

[7] J. H. Wong, K. Van Orden, B. R. Abrams, R. M. Iden, and J. Viraldo, "A Framework for Measuring Situation Awareness in Cyberspace Operations," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 65, no. 1, pp. 358–362, Sep. 2021, doi: 10.1177/1071181321651059.

[8] M. R. Endsley, *Situation awareness measurement: how to measure situation awareness in individuals and teams*. in Users' guides to human factors and ergonomics methods. Washington, DC: Human Factors and Ergonomics Society, 202

[9] R. M. Taylor, "Situation awareness rating techinique (SART): The development of a tool for aircrew systems design," *Aerospace Medical Panel Symposium*, vol. 3, pp. 1–17, 1990.

[10] Hart, S. G., and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical approach. In P. A. Hancock and N. Meshakti, Eds., Human mental workload, pp. 139-183. North-Holland: Elsevier Science.