Human Factors in Cybersecurity, Vol. 168, 2025, 22–28
https://doi.org/10.54941/ahfe1006138

Exploring How College Students' Mental Models of Cybersecurity Threats Predict Cyber Knowledge and Hygiene

David Schuster

San José State University, San Jose, CA 95192, USA

ABSTRACT

The role of human performance is critical in cybersecurity. Cybersecurity professionals and other employees must respond to unanticipated events successfully to maintain safety. Aviation safety has benefited from decades of human factors research to understand the role of threats and human error. Unfortunately, our present understanding of how to train people to respond effectively to cyber threats remains limited. The goal of this study is to investigate the relationship between threat understanding and engagement in behaviors that increase security, called cyber hygiene. Prior research suggested that an ability to recognize latent threats was associated with performance on a situational judgement test. In the current exploratory, descriptive study, the aim was to replicate that result in the domain of cybersecurity. The results of two studies suggest an association between cybersecurity knowledge and mental models of cyber hygiene but do not offer conclusions about the relationship between mental models and cyber hygiene behavior.

Keywords: Cybersecurity, Cyber threats, Decision making, Cyber hygiene, Mental models, Cybersecurity knowledge

INTRODUCTION

Cybersecurity and transportation systems have in common the potential of compromises to safety and effectiveness because of unforeseen events. The role of human performance is critical in these domains. Cybersecurity professionals, pilots, and other vehicle operators must respond to unanticipated events successfully to maintain safety. Many decades of human factors/ergonomics engineering have resulted in improvements to transportation safety and resulted in frameworks to understand and leverage human performance while mitigating the effects of unforeseen events. Despite this, our understanding of how to effectively train people to perform in cybersecurity roles remains limited. This is an urgent problem in the context of a cybersecurity workforce shortage (CyberSeek, 2025). At the same time, employees across the organization have a critical function in responding to cybersecurity threats. In both technical and nontechnical roles, cybersecurity skills and response to threats can be improved through training. However, there is a gap between the knowledge cybersecurity professionals and other

employees need to maintain security and the present ability of training to provide it. This has been described as a skills gap in the cyber workforce (Furnell, 2021).

The Need for Training

In technical roles, the NICE Cybersecurity Workforce Framework (NIST, 2023) provides an essential step in defining the knowledge and skills necessary to perform cybersecurity work. While the NICE Framework lists targets for training, such training could be developed more strategically with knowledge of how these skills develop. That is, better understanding of the cognitive mechanisms behind cybersecurity professional decision making may reveal opportunities to make training more efficient while also supporting cybersecurity workforce recruitment. Training for cybersecurity skills could be offered earlier and to a broader audience than ones based on current training pipelines, which have been insufficient to meet the demand.

In nontechnical roles, this training need is often met through cybersecurity awareness training, which could benefit from increased knowledge of how understanding of cybersecurity threats predicts behaviors that improve organizational security. Such behaviors are called cyber hygiene (Vishwanath et al., 2020).

Exploring Solutions From Transportation Safety

The theoretical foundation of this work is that frameworks used to improve transportation safety, such as by informing training, can be leveraged to understand the impacts of human behavior in cybersecurity. Aviation safety has benefited from decades of human factors research to understand the role of unforeseen events, which are threats, and human error, exemplified by Helmreich's threat and error management framework (2000). A key element of the threat and error management framework is that threats, such as bad weather, are unavoidable, and human error is one type of threat. Good outcomes require using all available resources to identify, understand, and respond to threats. The training that grew from this approach is called crew resource management (CRM; Helmreich et al., 2000). This human-centric approach has led to now-required training for flight crews and other aviation professionals.

The literature has examples of cybersecurity professionals and other employees being seen not as expert decision makers, but as sources of human error. For example, Verizon (2024) suggested that human error was a factor in 28% of breaches to organizations. Zimmerman et al. (2024) suggested three categories of responses to this finding. Constraining approaches attempt to remove the influence of people from the system or prevent undesirable behavior. Considering approaches acknowledge the role of people and aim to use design to guide behavior. Finally, enabling approaches view people as a resource.

The disconnect between successful enabling approaches that have applied to aviation and those constraining/considering approaches that are common in cybersecurity, is the focus of this study. To be more specific, the goal

of this study is to investigate the relationship between understanding of threats and engagement in behaviors that increase security, called cyber hygiene (Vishwanath et al., 2020). In past work (Schuster et al., 2009), we found evidence that an ability to recognize latent threats was associated with performance on a situational judgement test. In the current study, we aimed to replicate that result in the domain of cybersecurity and extend it to predict cyber hygiene behavior.

Mental Models

One way to describe an individual's understanding of threats is through mental model elicitation. Mental models are "mechanisms whereby humans are able to generate descriptions of system purpose and form, explanations of system functioning and observed system states, and predictions of future system states" (Rouse & Morris, 1986, p. 351). In this context, an individual's mental model is their representation of cybersecurity threats, and it provides a mechanism to understand the current context and make predictions about what will happen next. The ability to measure mental models allows researchers to elicit knowledge and describe the development of expertise. This study aimed to examine the relationship between the quality of an individual's representation of threats, their mental model, with their cybersecurity knowledge and cyber hygiene. The first hypothesis was that the ability to distinguish cybersecurity threats from irrelevant surface features is associated with greater knowledge in the domain. The second hypothesis was that this ability is also associated with engagement in cyber hygiene behaviors.

STUDY 1

Method

Our method was modeled after the approach of Schuster et al. (2009). In our study, 107 undergraduate students at a comprehensive university in the Western United States signed up for an online study for course credit, resulting in $N = 85$ participants. Participants completed two card sorts using a 40-card deck. Each card in the deck described a scenario that implied a cybersecurity threat in one of five categories related to cyber hygiene (from Vishwanath et al., 2020): Storage and device hygiene, transmission hygiene, Facebook and social media hygiene, authentication and credential hygiene, and email and messaging hygiene. Each scenario also included two irrelevant surface features that were directly stated. These were the owner of the device (work, school, or personal) and the operating system (Mac, Windows, or Linux). For example, one storage and device hygiene card for a personal Windows computer was: *You find a USB-stick in the common area of your dorms. You plug it into your personal computer running Windows so that you can identify who the USB-stick belongs to*. Each threat category appeared on six to nine cards in the deck, and the surface features varied in frequency between 12 and 14 of the cards.

During the study, participants completed an eight-item cybersecurity knowledge test adapted from Olmstead and Smith (2017). To measure

cyber hygiene, participants completed the Security Behavior Intentions Scale (Egelman & Peer, 2015). Participants were then asked to do two card sorts using the same 40-card deck, one as a free sort, with labels they supplied, and the second as a guided sort, with the threat categories provided as pile labels. In the free sort, participants were asked to place the cards into piles based on how they viewed the scenarios relating to each other. Participants were also asked to write labels that described their piles. In the guided sort, the researchers first provided the threat categories as pile labels and asked participants to sort the cards into piles using the threat categories as labels.

To score each card sort, agreement was calculated between it and perfect sorts according to each dimension of the scenarios. First, a pairwise list of all possible combinations of two cards was generated. If two cards appeared in the same pile, the pair was scored with a one; if not, the pair was scored with a zero. In this way, each card sort was represented as a binary array, and any two binary arrays of the same sort can be correlated to measure agreement. Answer key card sorts were made for the threat categories and each of the two surface features.

Results

Multiple regression was used to explore whether the guided and free sort's agreement with the threat classification scores predicted the knowledge test performance. This model was significant, $F(2, 82) = 4.56, p = .013, R^2 = 0.1$. However, only the free sort significantly predicted the knowledge score ($b = 3.54$, $t(82) = 2.33$, $p = .023$). In a second model, guided and free sort scores did not significantly predict SeBIS, ($F(2, 82) = 1.04$, $p = .359$, $R^2 = 0.025$). A correlation matrix was run as an exploratory analysis and diagnostic check; three correlations were significant. First, there was a significant correlation between agreement with the threat sort in the free sort and knowledge score, $r(85) = .28$, $p = .01$. Second, the agreement with the operating system sort in the free sort was also negatively related to knowledge score, $r(85) = -.25$, p = .02. Third, there was a negative relationship between agreement with the operating system sort in the free sort and agreement with the threat sort in the free sort, $r(85) = -0.38$, $p < .01$, but this was to be expected due to agreement with one card sort making agreement with a different arrangement less likely. While the results failed to evidence a relationship between sort performance and cyber hygiene, they did suggest that knowledge is associated with a tendency to free sort based on latent threat categories.

STUDY 2

Method

Study 2 was run in a subsequent term with four additional cards to more closely equate the representation of features across the card sort. Due to a misconfiguration, one storage and device hygiene card with a Linux personal computer was omitted, for a total of 44 cards in the Study 2 deck. In Study 2, 305 undergraduate and graduate students at a comprehensive university in

the western United States signed up for an online study, resulting in $N = 218$ (217 undergraduate and 1 graduate) participants with complete data.

Results

Multiple regression was used to explore whether the guided and free sort scores predicted the knowledge test performance. This model was significant, $F(2, 215) = 18.89$, $p < .001$, $R^2 = 0.15$. In this model, both the guided sort ($b = 2.89$, $t(215) = 4.90$, $p < .001$) and the free sort ($b = 1.86$, $t(215) = 2.07$, $p = .04$) significantly predicted the knowledge score. In the second model, guided and free sort scores did not significantly predict SeBIS ($R^2 = 0.003$, $p = .743$).

A correlation matrix was run as an exploratory analysis and diagnostic check; there were six significant relationships, which are reported in Table 1. The most relevant of these are a relationship between agreement with the threat sort in the guided sort with agreement with the threat sort in the free sort. There was also a significant relationship between agreement with the threat sort in the free sort and the knowledge test. While there was a significant relationship between the SeBIS and the knowledge test, the observed strength of the relationship was negligible.

Table 1: Correlations, means, and standard deviations from study 2. * denotes p <.05, ** denotes p <.01.

Variable	N	M	SD	1	2	3	4	5	6
1. Guided sort: Threat	218	0.38	0.19	--					
2. Free sort: Threat	218	0.10	0.12	0.30**	--				
3. Knowledge	218	6.24	1.68	0.36**	0.23**	--			
4. SeBIS	218	56.07	6.86	0.04	0.04	0.16*	--		
5. Free sort: Location	218	0.01	0.08	−0.01	−0.09	−0.09	−.06	--	
6. Free sort: OS	218	0.06	0.24	0.05	−0.36**	0.07	0.03	−0.16*	--

As with the first study, the results suggested that knowledge is associated with a tendency to complete the free sort based on latent threat categories. In the second study, the results additionally suggest the ability to identify the latent threat categories in the guided card sort is associated with cybersecurity knowledge. The results did not provide evidence of a significant relationship between sort performance and cyber hygiene.

CONCLUSION

Taken together, the results suggest an association between cybersecurity knowledge and mental models of cyber threats but do not offer conclusions about the relationship between mental models and cyber hygiene behavior. The results extend those of Schuster et al. (2009) to the cybersecurity domain; in that study, we found a relationship between operation knowledge and threat identification in the domain of driving. In the present study, we again found a relationship between domain knowledge and threat identification.

Although the sample size was adequate, this exploratory study had several limitations that likely affected statistical power. Although the card sort was based on categories found in the literature, the resulting measure had

unknown psychometric properties. It may be the case that the categories and scenarios of the card sorts do not adequately sample from the domain. Because one or more cards were omitted from each study, the features in the scenarios were represented with unequal frequency across the card sort. One way to improve the performance of the card sort in future research would be to conduct pilot card sorts with domain experts, such as cybersecurity professionals, and to use the results of these sorts to select the cards and develop the answer key. In this research, the opposite was done; the scenarios were derived from the answer key. Future research would also be stronger with this approach applied to the knowledge test. It was derived from published sources, but without expert review it is difficult to assert its content validity. These results neither suggest an association between cyber hygiene and mental models nor disconfirm one; that said, it could be possible that other factors predict cyber hygiene more than cybersecurity threat knowledge.

This study also demonstrates how card sorting can be used to elicit participant knowledge in various domains. While the results failed to evidence a relationship between sort performance and cyber hygiene behavior, they did suggest that mental model elicitation can be useful to predict operational knowledge. The pattern of results also suggests avenues to clarify these relationships in a study with validated measures, perhaps with the assistance of domain experts and in research with a population of working professionals.

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.

The author is grateful to the student members of the Virtual Environments, Cognition, and Training Research Lab at San Jose State University for their assistance with data collection.

REFERENCES

CyberSeek (February 14, 2025) 'Heatmap'. https://www.cyberseek.org/heatmap.html

Egelman, S., & Peer, E. (2015) "Scaling the security wall: Developing a security behavior intentions scale (SeBIS)," in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 2873–2882). ACM. http://dl.acm.org/citation.cfm?id=2702249

Furnell, S. (2021) "The cybersecurity workforce and skills," in Computers & Security 100, 102080. https://doi.org/10.1016/j.cose.2020.102080

Helmreich R. L. (2000) "On error management: Lessons from aviation," in BMJ, 320, 781–785.

NIST (2023) "Workforce Framework for Cybersecurity', NICE Framework". Available at: https://www.nist.gov/system/files/documents/2023/06/05/NICE%20Framework%20(NIST%20SP%20800-181)%20_one-pager_508 Compliant.pdf

Olmstead, K. & A. Smith. (March 22, 2017). What Americans know about cybersecurity. (2017). http://www.pewresearch.org//2017/3/22/what-the-public-knows-about-cybersecurity/

Rouse, W. B. and Morris, N. M., 1986. On looking into the black box: Prospects and limits in the search for mental models. *Psychological bulletin*, *100*(3), p. 349.

Schuster, D, M Harper-Sciarini, M Curtis, F Jentsch, and R Swanson. "The Relationship between Conceptual Understanding and Performance." Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting, 2009. https://doi.org/10.1177/154193120905302605

Verizon (2024) 2024 Data Breach Investigations Report. https://www.verizon.com/business/resources/reports/dbir/

Vishwanath, A., Neo, L. S., Goh, P., Lee, S., Khader, M., Ong, G., & Chin, J. (2020). Cyber hygiene: The concept, its measure, and its initial tests. Decision Support Systems, 113160.

Zimmermann, V., Schöni, L., Schaltegger, T., Ambuehl, B., Knieps, M., Ebert, N. (2024) "Human-Centered Cybersecurity Revisited: From Enemies to Partners," in Communications of the ACM 67, 72–81. https://doi.org/10.1145/3665665