Human factors related to the performance of intrusion detection operators

P. Lif and T. Sommestad

Department of Information and Aeronautics Systems, Swedish Defence Research Agency (FOI)
e-mail: patrik.lif@foi.se

Abstract

Intrusion detection systems are common in contemporary enterprises. These systems are sometimes operated by a single individual as a part time activity; they are sometimes operated by cyber security operation centres in which a group of technology experts with the sole task of monitoring, detecting, analysing and responding to threatening events in the computer network. In either case, human factors and ergonomics should be expected to influence the intrusion detection capability. In this paper, Wickens' model of information processing and human factors concepts and tests are related to the tasks of intrusion detection operators. This model is used to identify both environmental conditions and human capabilities that are relevant for operators' performance as well as experimental setups that can test hypotheses related to these factors. Based on this analysis, it is proposed that the most important factors are attention, vigilance, automation, multitasking and mental workload and tests and measures such as NASA-TLX and eye-movements, should be useful.

Keywords

Human factors, intrusion detection system, system operator, cyber security.

1. Introduction

Intrusion detection systems (IDS) continue to be a promising technology which attracts researchers. It is safe to say that vast majority of this research focuses on improvements of the technical solutions, without considering the human factors related to them. The extant research on human factors related to IDSs is mainly qualitative and descriptive, describing current practice and issues associated with it. In fact, the only quantitative studies found in the literature are the test by Sommestad and Hunstad (2013) , the test by Sawyer et al. (Sawyer et al., 2014) and the test by Ben-Asher and Gonzalez (2015). Sommestad and Hunstad found that intrusion detection operators screening the output of an IDS significantly reduces the portion of false alarms without significantly decreasing the probability that an attack is detected; Sawyer et al. who found that it is more difficult to detect correlations when relevant information is available for less time and when frequently and attacks are rare; Ben-Asher and Gonzalez found that situated knowledge helped operators detecting attacks and that knowledge in cyber security helps the operator to identify the attack type (i.e. the root cause of the attack).

This paper suggests that more experimental research should be focused on intrusion detection operators to further understand their role and factors that determine their efficacy. In addition, it is recommended that this research utilize established methods and tests from the *Human Factors and Ergonomics* domain.

Section two of the paper provide an overview of the work intrusion detection operators and relates this work to Wickens' model of information processing, a model commonly used in the human factors and ergonomics research. Section three relates the work of intrusion detection operators to theories within human factors and ergonomics research and with associated measurement procedures. Section four summarizes the result and present conclusions from the analysis.

2. Operation of intrusion detection systems

This paper is concerned with the activities carried out by intrusion detection operators who focus on identifying analyzing threats to cyber security. Overall, work on intrusion detection can be said to aim at identifying threats and ongoing attacks against the monitored systems and devising a suitable response. It is a subset of the activities usually carried out by operators in cyber security operation centers (Zimmerman, 2014) or a task placed on system security administrators (Werlinger et al., 2008). This section provides a general description of the activities intrusion detection operators carry out in intrusion detection work, based on the three phases for intrusion detection identified by Goodall et al. (2004): monitoring (section 2.1), analysis (section 2.2) and response (section 2.3). In addition, the section describes previous research of human performance in relation to these phases. We relate these tasks to Wickens model of information processing (Wickens, 2013), depicted in Figure 1

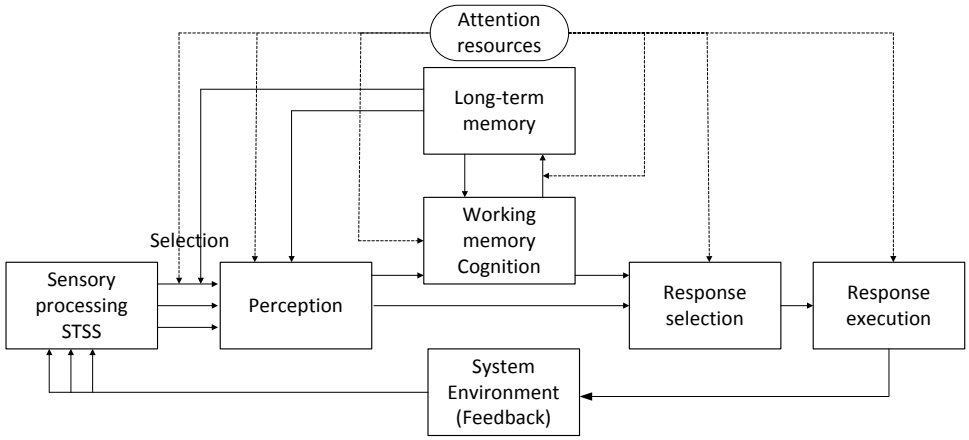

Figure 1: Wickens' model of information processing (Wickens, 2013).

The model describes typical stages or mental operations characterized by the flow of information when humans perform tasks. Events are first processed by our senses, which for IDS-operators is mostly visual. The information may be held in a short term sensory store for up to one second. From the sensation humans only perceive parts of the information available. Perceiving involves determining the meaning of the event. This is affected by past experiences that are stored in our long-term

memories. From perception, the models show two possible paths: one meaning a direct response and one where the information goes to working memory. From working memory the information may be transferred to long term memory for later use and used to select a response. Wickens' four-stage plus memory model also contains a feedback loop: an executed response may change the environment and thereby affect future sensory input. Also of great importance is attention, which act as a filter that selects some elements for further processing and block other elements. Also, attention acts as the "fuel that provides mental resources or energy to the various stages of information processing" (Wickens, 2013).

2.1. The monitoring task

Monitoring is the mundane task of identifying anomalies, irregularities and other signs of active threats in a stream of alerts. This includes, in addition to visual inspections of the output from log systems, keeping track of ongoing threats such as new malware and new software vulnerabilities. In a cyber-security operation center, where tasks are divided among personnel, this is a phase that the tier one typically takes care of (Zimmerman, 2014).

Much of the work performed in the monitoring phase keeping up with the information that flows in and dealing with false positives and false negatives from sensors (e.g. network based IDSs). In Wicken's model, this means that the operator will have to be able to perceive the alerts and decide if this alert (along with related alerts) is worthy to investigate further. Because of this, it has been suggested that the monitoring task bares the signature of a vigilance problem (Mancuso et al., 2014). Thompson et al. (2006) found that intrusion detection alerts is the key resource in the task. This is in line with the findings by Goodall et al. (2004), who just as Thompson et al. (2006), identified that efforts are made to configure the sensors' rulesets so that the right alerts and the right amount of alerts are raised. Other information used to effectively monitor the security posture includes information about ongoing threats (e.g. from email lists and feeds) and situational knowledge related to the own environment (e.g. vulnerability scans, normal business and network configurations) (Goodall et al., 2004).

An operator's performance in this phase can be seen as the ability to identify actual attacks (i.e. alerts worthy of further investigation) without spending much time on events that are benign or not threatening.). The only quantitative study on this phase found in the literature is the one by Sawyer et al. (2014). Sawyer et al. used *NASA Task Load Index* to assess the mental workload during a task pertaining to matching IP-addresses in two columns of a table. They found that, as one would expect, it is more difficult to detect correlations when the table with IP-addresses were rearranged frequently and attacks are rare. The implications from this test on intrusion detection work are not clear, especially since the task does not resemble any monitoring activity performed by those who perform intrusion detection.

2.2. The analysis task

In the analysis phase, the analyst starts with the events that were found to be worthy of further investigation during the triage performed in the monitoring phase. These are analyzed further in order to successfully diagnose it and determine if a response is needed. However, the activity is more unpredictable in terms of duration and frequency than the continuously ongoing monitoring phase.

The information passed from monitoring to analysis, which usually is an alert from a sensor, is during analysis fused together with other information that can help in this work in a deeper analysis aiming at determining if the root cause of the alerts is a threat to the organization. In terms of Wicken's model, the operator will iterate through a series of loops to sense, perceive, decide and act on information. Actions will mainly involve collection of new information sources (e.g. network scans, machine states and traffic logs) and tools to feed this information to (e.g. malware scanners). This will produce new things (e.g. tool output) to perceive and make decision based on. Alternatives will be explored and hypotheses will be tested before a final decision is reached concerning the root cause. According to Goodall et al. (2004), system administrators mainly use knowledge related to intrusion detection (e.g. the sensors), general security expertise and local knowledge about the own environment. Furthermore, other types of system logs are often used in this process (Thompson et al., 2006).

An operator's performance in this phase can be seen as the ability to identify actual attacks along with root causes within a short period of time. This includes the ability to dismiss false alarms and, when possible, identify the root cause for the false alarm. Sommestad and Hunstad (2013) tested an intrusion detection operator's ability to filter out relevant alerts in an experiment based on synthetic data. These results suggest that an operator can reduce the portion of false alarms significantly, while not missing that many actual attacks. Ben-Asher and Gonzalez (2015) performed an experiment where subjects had the task of determining if a system state (e.g. certain network loads and active services) posed a security threat or not. In this simplified IDS task, they found that more knowledge in cyber security facilitated the correct detection of malicious events, decreased the false classification of benign events as malicious and helped operators to identify the type of threat causing the state.

2.3. The response task

In the response phase the operator device a suitable response to the events that were detected during the monitoring phase and analyzed during the analysis phase. The most common forms of responses are: interventions, feedback and reporting (Goodall et al., 2004). Examples of interventions include pulling a network plug, reconfiguring a network, reinstalling a machine or patching a software; feedback usually means tuning the signature ruleset of the IDS; examples of reporting includes letting other know about the threat or escalating the matter to even more in depth analysis (e.g. to collect forensic evidence and pursue legal action). Which of these

types of interventions that is suitable depends not only on the event as such, but also on the operator's role in the organization and the operator's constituency.

Because the proper response is contingent on the cause of the incident and on the operator's authority/responsibility, it is difficult to define an overall performance marker for this activity. It could be that also this phase involves a number of iterations of perceive-decide-act-respond, e.g. where the operator contact asset owners to collect more information and to discuss alternatives. In the end, the action will be the response used. Cichonski et al. (2012) state that the criteria for determining the right containment strategy include: potential damage, need for evidence preservation, availability requirements, the cost of the strategy, and the effectiveness of the strategy. While it seems quite possible to study humans in the response phase there are, to the authors' knowledge, no empirical studies in the extant literature.

3. Human factors and intrusion detection

It is clear that the intrusion detection operator and his/her interaction with the IDS are important for the detection capabilities in most organizations. Human factors methods could be used to understand the operator's role and variables that determine the overall intrusion detection capability. In intrusion detection research, overall intrusion detection capability is typically using signal detection theory, which allows four possible outcomes: hit, miss, false alarm and correct rejection. The purpose with this chapter is to exemplify human factors issues that are relevant in information processing for IDS-operators, and are likely to influence the overall intrusion detection capability. Our analysis suggests that the established concepts and test from human factors research depicted in in Figure should be considered highly relevant. These will be further presented in sections 3.1 to 3.6.

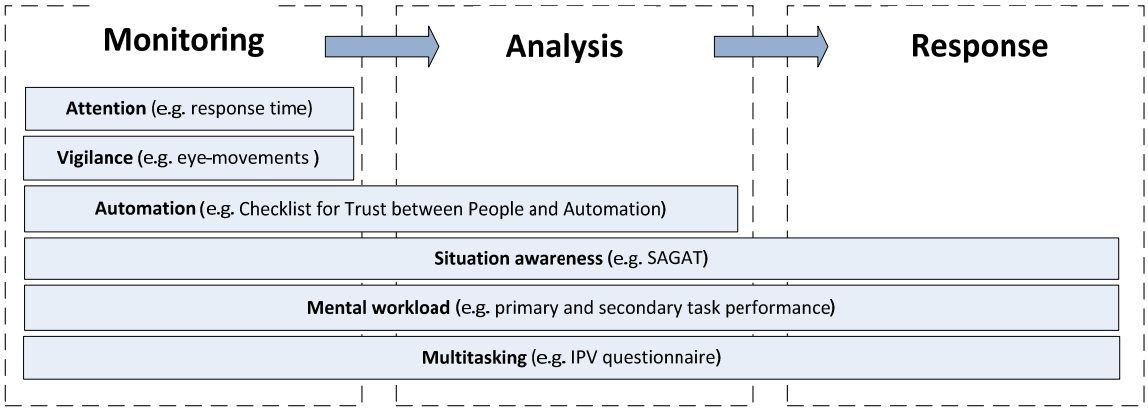

Figure 2: Model presenting human factors concepts and measurement methods that can be used in IDS-operators during monitoring, analysis, and response.

3.1. Attention

As seen in Figure 2 *attention* (Wickens, 2013) affects IDS-operators in three (of four) information processing stages: perception, response selection and response execution. Limitations in human attention are well known and can be described in three categories: 1) selective attention, 2) focused attention, and 3) divided attention (Wickens, Hollands, Banburry, & Parasuraman, 2013). Selective attention is when we select and process the wrong information; focused attention is when we fail to focus on one task, even when we know we should (i.e. we get distracted); divided attention is when we want or need to focus on multiple tasks but it is not possible, i.e. the IDS-operator has to handle information from two sources simultaneously. Visual and auditory attention is different, and should be used with the strength and weaknesses. Visual attention can be improved in the design process of displays, e.g. by close proximity in space or color (Wickens et al., 2013). Auditory attention is different from visual attention since it is omnidirectional and can be used in darkness or even during sleep. Attention could be measured in terms of response time and eye movements. For instance, response time could be measured as the time it takes for the operator to see some alert that he/she is set to look for and eye movement measurements could be used to measure if the operator continuously looks for new alerts on the computer screen.

3.2. Vigilance

Within the vigilance paradigm, an operator is required to work during a long time and detect signals that are intermittent, rare, unexpected and often of low salience (Wickens et al., 2013). As noted by Mancuso et al. (Mancuso et al., 2014), *vigilance* is likely to be important to IDS-operators. IDS-operators' vigilance could be manipulated and tested to better understand its relevance in the monitoring task, and thereby enable optimization of the IDS-system from a human factors perspective. To measure vigilance; psychophysiological methods (e.g. heart rate and skin conductance), response time (e.g. time to answer an alarm), eye-movements (to analyze search patterns, what subjects look at, and what they missed) and NASA-TLX (a subjective workload assessment tool for ratings of operators mental workload in man-machine systems).

3.3. Automation

Experience shows that IDS-operators cannot possibly handle all alarms produced by today's sensors, and some level of *automation* will be necessary. Implemented correct, automation can generate substantial benefits in the monitoring process. However, it may also have negative effects and thereby lower the overall detection capability by suppressing true alerts from the alert feed. In many situations a task analysis should be conducted to weight costs and benefits of automation to determine how and if automation should be implemented. In complex systems it can be hard, or impossible for the operator to understand how automation is done in the system at hand, e.g. how rulesets of the IDS works. Experience from other domains suggests that one important factor is operators' possibility to override automated functions

and that another important part of automation is *trust*. An operator's trust on a system and automation is affected when he/she will conduct actions suggested by the alarm. Depending of the system performance trust is built over time. If there are too many false alarms, it is possible that the operator will completely ignore the alarm and seek other information and act according to experience or other information sources (Wickens et al., 2013). Automation is not all or none, but a continuum from fully manual to full automation. To get an understanding of a system at the level of automation Sheridans & Verplank (1978) developed a scale with ten levels. Also trust between people and systems can be investigated using the Checklist for Trust between People and Automation (Jian, Bisantz, Drury, & Llinas, 2000).

3.4. Situation Awareness

The analysis task is typically solved using a complex process where alternatives are explored and hypotheses are tested. As noted above, this includes keeping track of current and previous alerts, other logs, and the monitored system as well as its environment. Thus, in this process, *situation awareness* (SA) (Endsley, 1995a, 1995b) appears to be necessary for IDS-operators to perform well. SA can be seen as an internal mental model where incoming data from systems, the environment, and co-workers should be integrated. This integrated picture can be used for decision making and actions (in the response phase), but in a complex and dynamic situations it is necessary to update situations awareness continuous.

Good SA for IDS-operators may be problematic for operators supervising huge data sets in very complex network systems that often are geographically distributed. SA is divided in three levels: perception, understanding and prediction. These three levels appear to match well with the level of awareness required during the monitoring (perception), analysis (understanding) and response (prediction) phases. It is applicable to IDS-operators that must perceive attacks, value if it is a real attack, and also be prepared for future attacks.

While SA is often referred to in cyber security research, it is typically used as conceptual tool without explicit measurements of the level of SA (see Franke and Brynielsson (2014). The experiment by Stevens-Adams et al. (Stevens-Adams et al., 2013) is an exception, where participants ability to answer questions concerning attacks and the overall system state was measured using a method developed by the experimenters. More established SA measures can be divided in: requirement analysis, freeze probe recall, real-time probe, post-trial subjective ratings, observer rating, process indices and team SA. Two popular methods are SAGAT (freeze probe technique) and SART (self-rating technique). In SAGAT, participants are interrupted and excluded from sensor information at certain (randomly selected) points in time and asked to state the current and/or future state of system. For example, with SAGAT, the IDS-operator's screen could be shut and the operator could be asked how much traffic that is coming in through the external firewall, which users that are logged on and/or if there is any machine that has been involved in multiple alerts the last 10 minutes. SART is a questionnaire developed for air traffic control for retrospectively measuring ten dimensions of SA: familiarity of the situation, focusing

of attention, information quantity, information quality, instability of the situation, concentration of attention, complexity of the situation, variability of the situation, arousal, and spare mental capacity. Questionnaires like SART uses a 7-grade rating scale that operators answer and this could be adapted to IDS-work, or used as-is.

3.5. Multitasking

In the response task, *multitasking* is likely to be an issue. Multitasking is closely connected to attention, but it has more emphasis on the actual execution of the task and the interaction between perceptual, cognitive and motor processes is in focus in multitasking. This interaction may cause interference, and system effectiveness may be reduced since the operator cannot handle all tasks or process the information available. Perhaps the most obvious interruption is perceptual or motor (e.g moving the mouse or hit buttons), when an operator only can focus on one screen and normally only handle one motor task at a time. Cognitive interruptions are more subtle, but as for perceptual and motor processes, our resources to handle tasks are limited (Wickens et al., 2013). Also for cognitive tasks, interference occurs when the same resources are needed for cognitive processes (Wickens, 2008). By analyzing the operators tasks in a continuum, from tasks taking a couple of seconds to tasks that demands focus for perhaps twenty minutes, the operators work situation can be improved. It is often easy to shift between tasks that take a couple of seconds, but a forced interruption of the primary task is not optimal. Interruptions during task with high *mental workload* (see below) is harder to recover from than during tasks with low mental workload (Stanton et al., 2013). The literature reveals that there is some confusion of how to measure multitasking performance, but speed and error rates are sometimes used. Two standard questionnaires have been used: the Inventory of Polychronic Values (IPV) and Modified Plychronic Attitude Index 3 (MPAI3) (König & Waller, 2010; Poposki & Oswald, 2010).

3.6. Mental workload

Mental workload is an established concept and there is a large body of literature on the topic (Tsang & Wilson, 1997). The important and underlying theoretical assumption is operators limited capacity to process information (Kahneman, 1973). With greater task difficulty and complexity increasing mental workload is acquired and when demands exceed capacity task performance will decrease. Another known phenomena is that interruptions during task with high *mental workload* are harder to recover from than during tasks with low mental workload (Stanton et al., 2013).

Measuring workload requires reliable and valid metrics but since there are multiple methods it is not always obvious which method to choice. Methods differ in sensitivity and must therefore be matched to the situation in hand (Matthews, Reinerman-Jones, Barber, & Abich, 2015). Some examples of methods to measure mental workload are: 1) primary and secondary task performance (performing two tasks to evaluate spare capacity), 2) psychophysiological measures (e.g. heart rate), and 3) NASA-TLX (subjective ratings of operators mental workload). Often these

measures can be combined and sometimes simplified versions of these methods are used.

4. Summary and conclusions

Previous research has found that IDS-operators play an important role and face several challenging tasks when they 1) monitor systems, 2) analyse events and 3) respond to events. In terms of established concepts from human factors research, the success in all these phases appears to be related to situational awareness, mental workload and multitasking; issues pertaining to automation appear to be relevant for phases of monitoring and analysis; the operator's attention and vigilance appears to be important in the monitoring phase.

Although IDS-operators are known to be important and IDS are important cyber security tools, the extant literature contains few quantitative tests on IDS-operators or the challenges they face. Established measurement techniques of both objective and subjective types could be used to better understand how these concepts relate to, and influence, the efficacy of IDS-operators. For example, eye-trackers would give us valuable information about search pattern and how to focus attention to the right information; self-ratings of mental workload (e.g. from NASA-TLX) could be used to measure mental workload and understand how it relates to performance; the importance of different types of situational awareness could be measured using SAGAT. These, and other established methods from human factors research, ought to be leveraged in further research on IDS-operators. Next we plan to validate the model (Figure 2) in a two-step process. First we will conduct further discussions and interviews with IDS-operators, and second we will run experiments with IDS-operators. In the second part human factors concepts and measures will be used (e.g. SA, multitasking, vigilance, mental workload) to get a better understanding of IDS-operators and their work situation.

5. References

Ben-Asher, N., & Gonzalez, C. (2015). Effects of cyber security knowledge on attack detection. Computers in Human Behavior, 48(0), 51-61.

Endsley, M. (1995a). Measurement of Situation Awareness in Dynamic Systems. Human Factors, 37(1), 65-84.

Endsley, M. (1995b). Toward a Theory of Situational Awareness in Dynamic Systems. Human Factors, 37(1), 32-64.

Franke, U., & Brynielsson, J. (2014). Cyber situational awareness – A systematic review of the literature. Computers & Security, 46(0), 18-31.

Goodall, J., R., Lutters, W., G., & Komplodi, A. (2004). The Work of Intrusion Detection: Rethinking the Role of Security Analysts. Paper presented at the Proceedings of the Tenth Americas Conference on Information Systems, New York, US.

Jian, J.-Y., Bisantz, A. M., Drury, C. G., & Llinas, J. (2000). Foundations for an empirically determined scale of trust in automated systems. Dayton, OH, US.: Wright-Patterson Air Force Base.

Kahneman, D. (1973). Attention and effort. Englewood, NJ: Prentice Hall.

König, C. J., & Waller, M. J. (2010). Time for Reflection: A Critical Examination of Polychronicity. Human Performance, 23(2), 173-190. doi: 10.1080/08959281003621703

Mancuso, V. F., Christensen, J. C., Cowley, J., Finomore, V., Gonzalez, C., & Knott, B. (2014). Human Factors in Cyber Warfare II: Emerging Perspectives. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58(1), 415-418.

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., & Abich, J. (2015). The Psychometrics of Mental Workload: Multiple Measures Are Sensitive but Divergent. Human Factors: The Journal of the Human Factors and Ergonomics Society, 57(1), 125-143.

Poposki, E. M., & Oswald, F. L. (2010). The Multitasking Preference Inventory: Toward an Improved Measure of Individual Differences in Polychronicity. Human Performance, 23(3), 247-264.

Sawyer, B. D., Finomore, V. S., Funke, G. J., Mancuso, V. F., Funke, M. E., Matthews, G., & Warm, J. S. (2014). Cyber Vigilance: Effects of Signal Probability and Event Rate. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58(1), 1771-1775.

Sheridan, T. B., & Verplank, W. (1978). Human and Computer Control of Undersea Teleoperators. Cambridge, MA: Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT. .

Sommestad, T., & Hunstad, A. (2013). Intrusion detection and the role of the system administrator. Information Management & Computer Security, 21(1), 30-40.

Stanton, N., Salmon, P., Rafferty, L., Walker, G., Baber, C., & Jenkins, D. (2013). Human Factors Methods. Surrey, England: Ashgate Publishing Limited.

Stevens-Adams, S., Carbajal, A., Silva, A., Nauer, K., Anderson, B., Reed, T., & Forsythe, C. (2013). Enhanced Training for Cyber Situational Awareness, Foundations of Augmented Cognition (Vol. 8027, pp. 90-99): Springer Berlin Heidelberg.

Thompson, R. S., Rantanen, E. M., & Yurcik, W. (2006). Network Intrusion Detection Cognitive Task Analysis: Textual and Visual Tool Usage and Recommendations. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50(5), 669-673.

Tsang, P., & Wilson, G. F. (1997). Mental workload. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics (Second edition ed.). New York, U.S.: John Wiley & Sons, Inc.

Werlinger, R., Hawkey, K., Muldner, K., Jaferian, P., & Beznosov, K. (2008). The challenges of using an intrusion detection system: is it worth the effort? Paper presented at the Proceedings of the 4th symposium on Usable privacy and security, Pittsburgh, Pennsylvania, USA.

Wickens, C. (2008). Multiple Resources and Mental Workload. Human Factors, 50(3), 449-455.

Wickens, C. (2013). Attention. In D. N. Lee & A. Kirlik (Eds.), The Oxford Handbook of Cognitive Engineering. Oxford, U.K.: Oxford University Press.

Wickens, C., Hollands, J., Banburry, S., & Parasuraman, R. (2013). Engineering Psychology and Human Performance. New York: Pearson Education Inc.

Zimmerman, C. (2014). Ten Strategies of a World-Class Cynersecurity Operations Center. Bedford, MA, U.S.: The Mitre Corporation.