A Taxonomy of Cyber Awareness Questions for the User-Centered Design of Cyber Situation Awareness*

Celeste Lyn Paul and Kirsten Whitley

Department of Defense, United States
clpaul@tycho.ncsc.mil, visual@tycho.ncsc.mil

Abstract. This paper offers insights to how cyber security analysts establish and maintain situation awareness of a large computer network. Through a series of interviews, observations, and a card sorting activity, we examined the questions analysts asked themselves during a network event. We present the results of our work as a taxonomy of cyber awareness questions that represents a mental model of situation awareness in cyber security analysts.

Keywords: Computer security, situation awareness, user-centered design.

1 Introduction

This paper presents a taxonomy of cyber awareness questions derived from a series of user-centered research activities that can be used to inform the design and development of cyber situation awareness technology. One of the most important responsibilities of a cyber security analyst is to watch over and protect his network from harm. Maintaining situation awareness of the wide variety of events that occur and massive amounts of data generated is one of many analytic challenges. Situation awareness technology aims to reduce the data overload burden placed on the analyst. Good situation awareness technology requires good design, and good design requires a good understanding of the user and a focus on the user during the design process.

In the case of a cyber security analyst, the practice of good user-centered design is focused on his security-related work processes on a large computer network. One way of understanding how a cyber security analyst accomplishes situation awareness on a large computer network is to study the questions he may ask himself during the course of a network event. Studying the relationships between these questions will lead to a better understanding of the analysts' mental model of cyber situation awareness. A mental model of cyber situation awareness is a valuable tool in the user-centered design of cyber-related technology and to researchers who cannot always study cyber security analysts in the field.

L. Marinos and I. Askoxylakis (Eds.): HAS/HCII 2013, LNCS 8030, pp. 145–154, 2013.

2 Background

There is a growing body of work within the cyber security field that is focused on understanding the work processes of cyber security analysts. For example, the results of a series of interviews with a wide variety of cyber security analysts by Werlinger et al. [8] described three stages of computer network incident response activities: preparation; anomaly detection; and, anomaly analysis. Work by Goodall et al. [5] discussed the work process for network intrusion detection analysts in four task stages: monitoring the network for events; triaging an event; analysis of an event; and, response to an event. Thompson et al. [7] expanded Goodall et al.'s work to include a pre-processing stage before the monitoring stage that involves intrusion detection system preparation, as well as expanding the triage stage to include activities for determining the cause of an event and deciding if the event should be escalated to analysis. Although individually these studies describe different phases of activities within the cyber analytic work process, together they infer a general cyber analysis work process model: preparation, monitoring, detection, analysis, and response to network events.

Few researchers have specifically focused on situation awareness during the cyber analysis work process. Situation awareness is a state of knowledge within the context of a dynamic system, often with three stages: perception, comprehension, and projection [4]. Work by D'Amico et al. [2] examined the analytic questions of intrusion detection analysts to understand how they fused complex data during different stages of situation awareness. They developed a model of situation awareness that extended and overlapped with the model for the cyber analysis work process: event detection (monitoring and detection); situation assessment (analysis); and, threat assessment (response). Later research by D'Amico et al. [3] added role-based work processes that corresponded to their model of situation awareness, such as: triage analysis; escalation analysis; correlation analysis; threat analysis; incident response; and, forensic analysis.

However, there is still a general lack of information on cyber security analysts, their work processes, and how they establish and maintain situation awareness. Conducting empirical and ethnographic research with cyber security analysts is often difficult. There are a number of challenges to involving cyber security analysts in research, such as establishing contact with cyber security analysts who have the time to participate in research and are willing to share potentially sensitive information related to their jobs [1]. Additionally, the role of a cyber security analyst is difficult to define and ranges from a system administrator, intrusion detection analyst, to an incident responder. Cyber security analysts may take on the same, different, or overlapping responsibilities depending on the scope of the job role or size of the organization [3]. As computer networks become larger and more complex, understanding how cyber security analysts manage the large amounts of information generated by these networks and maintain awareness of the increasing number of events on these networks will be critical to future technology design and development.

3 Methodology

A combination of ethnographic research methods were used to understand the mental model of cyber security analysts responsible for a large network. First, interviews and observations were conducted to gain an understanding of analysts' work environment. Then, a card sorting activity was conducted to understand analysts' conceptual models of situation awareness on a network. Analysts in our study were primarily responsible for intrusion detection and not incident response.

3.1 Interviews and Observations

Interviews were conducted with six cyber security experts. Participants had at least one year of previous or current experience working in support of a network operations center as well as additional experience in cyber security. The interviews were open-ended with no structured topics except for the overall purpose of the interview. Participants were asked to discuss their experiences in cyber security and within the operations center. Participants were asked to talk freely, and were only interrupted with follow-up and clarification questions. If the topic did not come up during the initial discussion, participants were prompted to discuss their experiences with cyber situation awareness and the types of high level orientation questions they ask themselves during a new or ongoing event. Interviews lasted between 45 minutes to 1.5 hours. To supplement the interviews, approximately 25 additional hours of observations of a round-the-clock network operations center were conducted. This included general observations of analyst work during normal operations, attending operations center meetings, and observing two training exercises. Participant interruptions were minimal and participants were available to answer questions and discuss their activities. Observation sessions lasted between one and four hours each.

3.2 Card Sorting Activity

Card sorting is a knowledge elicitation method that helps people describe relationships between and hierarchy among concepts [6]. An open card sorting study was conducted with 12 cyber security analysts using 44 cyber situation awareness questions. Participants had at least one year of previous or current experience working in a round-the-clock network operations center and were primarily responsible for network intrusion detection. Participants were not responsible for incident response.

Cyber Awareness Questions. A list of questions was derived from the interview and observation data. These were questions analysts reported asking themselves to establish and maintain awareness of new and ongoing network events. The informality and similarity between questions was not edited to preserve any nuance that existed in question phrasing. Table 1 provides a list of the cyber awareness questions derived from interviews and observations of cyber security analysts and used in the card sorting study.

Procedure. Participants sorted the 44 cyber awareness questions into groups that best reflected their understanding of the questions. Once the questions were sorted into groups, participants labeled each group with a descriptive word or phrase. At the end of the activity the study moderator debriefed the participants' work by asking them to explain how they sorted the questions and why. Card sorting sessions lasted between 45 minutes and one hour.

Table 1. Cyber awareness questions used in card sorting study

1. Are there more or less bad guys attacking my network than normal?	23. What did the bad guys take?
2. Can I see the attack I know is happening?	24. What do I do about the attack?
3. Does the attack have a negative effect on other business operations?	25. What do I not see happening on my network?
4. Does this attack matter?	26. What does my network look like to the bad guys?
5. Have I seen an attack like this before?	27. What does my network look like?
6. How did the bad guys get into my network?	28. What does the attack look like?
7. How is my network being attacked?	29. What does the event on my network mean?
8. How is my network different from last week?	30. What happened on the network last night?
9. How serious is the attack?	31. What is different on my network from last week?
10. How successful was the attack?	32. What is happening on my network now?
11. Is anything different happening on my network than normal?	33. What is happening with my network?
12. Is anything interesting happening on my network?	34. What is normal for my network?
13. Is it a good day on the network?	35. What is not normal for my network?
14. Is my network configured correctly?	36. What is the most important event happening on my network?
15. Is my network healthy?	37. What is the status of my network?
16. Is something bad happening on the network?	38. What malware have been detected on my network?
17. Is something happening on the network?	39. What systems are up or down on my network?
18. Is the event on my network good, bad, or just different?	40. Where are the bad guys attacking from?
19. Is there more or less traffic on my network than normal?	41. Where on my network am I being attacked?
20. Is this a new attack I have not seen before?	42. Who is attacking my network?
21. What are the bad guys doing on my network?	43. Why is my network being attacked?
22. What did the bad guys do?	44. Why are computers on my network not available?

3.3 Analysis

An analysis of the top question pairs based on descriptive statistics and question co-occurrence provided insights to the most critical cyber situation awareness questions. Co-occurrence was calculated as the number of participants who sorted two questions together independent of the group the questions were sorted into during the card sorting activity. Graph visualization of question co-occurrence was then used to

analyze clusters of questions. Graph features such as, network weight, clusters, and bridges were used to identify topic areas. Content analysis of the clusters provided insight to topic areas that shape an analyst mental model for cyber situation awareness. Knowledge from the interviews and observations provided additional context and was integrated into the interpretation and understanding of the results.

3.4 Limitations

Cyber security analysts are often difficult to involve in research [1]. Only a limited number of cyber security analysts were available to participate in this study. Card sorting studies can be run with a large number of participants using quantitative analysis methods or a small number of participants using qualitative analysis methods [6]. We chose to conduct a small qualitative card sorting study because of the benefits of in-depth qualitative analysis and the challenges recruiting cyber security analysts. To compensate for a smaller study, we triangulated our results with graph visualization analysis and the results from observations and interviews.

4 Results

4.1 Top Question Pairs

There were 144 card pairs with 50% (6/12 participants) co-occurrence representing 98% (43/44) of the questions in the study. There were 21 question pairs with 75% (9/12 participants) co-occurrence representing 52% (23/44) of the questions in the study. Overall, there was good representation of all the questions in the study within the highest co-occurrence pairs. Table 2 provides a list of the cyber awareness question pairs with 75% co-occurrence. Additionally, we found three types of relationships between the highest co-occurrence question pairs that we define as: question similarity, question sets, and question order. The question types were derived from qualitative analysis of the question relationships.

Similarity. The first type of question pair relationship was based on *similarity* (A is the same as B) in which two questions are asking the same thing. These questions are essentially the same, just asked differently depending on the situation:

"*How is my network different from last week?*"
"*What is different on my network from last week?*" (9/12 participants)

Set. The second type of question pair relationship was a *logical set* (A and B are the same type) in which questions are distinctly different but related in purpose or goal:

"*Is anything different happening on my network than normal?*"
"*Is anything interesting happening on my network?*" (10/12 participants)

Table 2. Top 75% co-occurrence (9/12 or more participants) cyber awareness question pairs

CO	Question	Question
92%	Is anything interesting happening on my network?	Is something bad happening on the network?
92%	What did the bad guys take?	How successful was the attack?
83%	What happened on the network last night?	What is different on my network from last week?
83%	Is something happening on the network?	Is anything interesting happening on my network?
83%	Is anything interesting happening on my network?	Is anything different happening on my network than normal?
83%	What does the attack look like?	Have I seen an attack like this before?
75%	How is my network different from last week?	What is different on my network from last week?
75%	Is anything different happening on my network than normal?	Is something happening on the network?
75%	Is anything different happening on my network than normal?	Is something bad happening on the network?
75%	Is anything different happening on my network than normal?	What is happening with my network?
75%	Is anything different happening on my network than normal?	Is there more or less traffic on my network than normal?
75%	Is something happening on the network?	Is something bad happening on the network?
75%	What is happening with my network?	What do I not see happening on my network?
75%	Is it a good day on the network?	Is my network healthy?
75%	Is it a good day on the network?	What is the status of my network?
75%	What is the status of my network?	What is normal for my network?
75%	What is the status of my network?	What systems are up or down on my network?
75%	What does the attack look like?	Who is attacking my network?
75%	Have I seen an attack like this before?	Who is attacking my network?
75%	Does this attack matter?	How serious is the attack?
75%	What did the bad guys do?	What did the bad guys take?

While the framework of this question pair is very similar, e.g., "Is anything … on my network?", the use of "different" and "interesting" make the questions distinct. Based on the knowledge gained from the interviews and observations, "different" is not always "interesting" but both are equally important and asked.

Order. The third type of question pair relationship was a *logical order* (A comes before B) in which a question was a logical follow-up or a requirement to the previous question. For example, the order of these questions implies an analytic process, including the priority or requirement to answer certain questions before others:

"*What does the attack look like?*"
"*Have I seen an attack like this before?*" (10/12 participants)

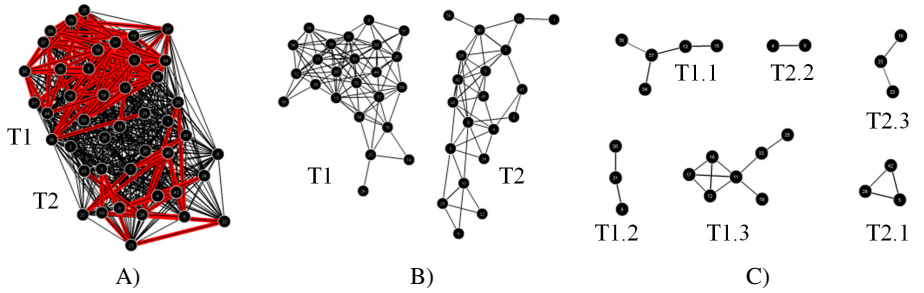

Fig. 1. Graph visualizations of question co-occurrence with potential topics. A) Base network with 50% (6/12 participants) co-occurrence highlighted revealing two topic areas; B) 50% co-occurrence network with two main topic areas; C) 75% (9/12 participants) co-occurrence network revealing six sub-topic areas.

4.2 Graph Visualization and Content Analysis

The most interesting graph visualizations were those that were expressed by the highest number of participants. Figure 1 shows graph visualizations for all question pairs (Fig.1-A), 50% co-occurrence (Fig.1-B) that represented questions paired by at least half of the participants in the study, and 75% co-occurrence (Fig.1-C) that represented questions paired by a majority of participants.

An overlay of the 50% co-occurrence question pairs on the base network visualization showed two question co-occurrence clusters, potentially revealing two main topic areas (Fig.1-A). A visualization of the 50% co-occurrence question pairs (Fig.1-B) showed the two clusters found in the base network (Fig.1-A) as well as graph features such as sub-clusters and bridges that identify possible sub-clusters. A visualization of the 75% co-occurrence question pairs (Fig.1-C) showed six small clusters that are a sub-set of the two 50% co-occurrence clusters (Fig.1-B).

Content analysis of the questions in the two clusters (Fig.1-B) revealed potential topics in Event Detection (T1) and Event Orientation (T2). Further content analysis of the six clusters from the 75%+ co-occurrence question pairs (Fig.1-C) revealed potential sub-topics such as Network Baseline (T1.1), Change Detection (T1.2), Network Activity (T1.3), Event Identification (T2.1), Mission Impact (T2.2), and Damage Assessment (T2.3).

Further analysis of different levels of co-occurrence visualization disambiguated the relationships between question pairs that were not clearly from one of the six 75%+ co-occurrence clusters. For example, several additional questions can be classified in one of the six topics by examining the visualization for the 67% co-occurrence question pairs (8/12 participants) and 58% co-occurrence question pairs (7/12 participants). These additional questions are included in Table 3 taxonomy of cyber awareness questions.

5 Taxonomy of Cyber Awareness Questions

The presented taxonomy of cyber awareness questions offers insights into different stages of cyber situation awareness (Table 3). The categories were derived from the results of the study and previous work in cyber situation awareness. The questions were organized into categories based on their co-occurrence score from our study, ranging from 58% (7/12 participants) to 92% (11/12 participants) co-occurrence.

Event Detection. This category contains questions that analysts ask prior and during initial awareness of a network event. In Event Detection, these questions roughly align with the perception phase of situation awareness.

Network Baseline. A network baseline is a model or snapshot of the network when it functioning in a "normal" state, in which "normal" is often the best approximation of healthy, acceptable operation. Comparison to their mental baseline was a common way analysts in this study articulated how their analytic needs precede cyber events.

Change Detection. Change detection is the ability to compare states of the network to identify differences and trends. The concept differs only slightly from network Baseline in that, here, analysts focus on the comparison between two network states.

Network Activity. Network activity reflects a shift from "normal" to "not normal" network activity that acts as a cue for the analyst to narrow his attention for in-depth analysis. These questions relate closest to the situation awareness concept of perception as well as allude to the transition between Change Detection and Event Identification.

Event Orientation. This category contains questions that analysts ask and are most closely aligned with the comprehension stage of situation analysis. In Event Orientation, analysts are working to maximize insight into an identified cyber event.

Identification. Identification is the recognition that a subset of network activity warrants analytic attention. This category is the detailed analysis of an event to identify who, what, when, where, and why and attack is happening and to possibly link the activity to familiar threats.

Mission Impact. Mission impact is analysis to prioritize the importance of an identified threat. Analysts must judge the severity of the threat to business operations, such as personnel necessary to respond to the threat, to help determine how to distribute limited resources for investigating and responding to the threat.

Damage Assessment. Damage assessment is analysis to inform a response to an identified threat. These questions differ somewhat from Mission Impact; here, the goal is to understand the full effects of the attack on the internal network.

Table 3. Taxonomy of Cyber Awareness Questions for Cyber Situation Awareness

Event Detection	**Event Orientation**
Network Baseline	*Event Identification*
• Is it a good day on the network? • Is my network configured correctly? • Is my network healthy? • What does my network look like? • What is happening on the network now? • What is normal for my network? • What is not normal for my network? • What is the status of my network? • What systems are up or down on my network?	• Have I seen an attack like this before? • Is this a new attack I have not seen before? • How is my network being attacked? • What are the bad guys doing on my network? • What does the attack look like? • Where on my network am I being attacked? • Who is attacking my network? • Where are they bad guys attacking from? • Why is my network being attacked?
Change Detection	*Mission Impact*
• How is my network different from last week? • What happened on the network last night? • What is different on my network from last week?	• Does this attack matter? • How serious is the attack? • What do I do about the attack?
Network Activity	*Damage Assessment*
• Is anything different happening on my network than normal? • Is anything interesting happening on my network? • Is something bad happening on the network? • Is something happening on the network? • Is the event on my network good, bad, or just different? • Is there more or less traffic on my network than normal? • What do I not see happening on my network? • What does the event on my network mean? • What is happening with my network? • What is the most important event happening on my network? • Why are computers on my network not available?	• Does the attack have a negative effect on other business operations? • How successful was the attack? • What did the bad guys do? • What did the bad guys take?

6 Conclusion

In this paper we utilized user-centered and ethnographic research methods to explore and understand the mental model of cyber security analysts responsible for a large network. Our results lead to the contribution of a taxonomy of cyber awareness questions that describes a set of questions analysts ask themselves while they establish and maintain situation awareness during a network event.

This taxonomy provides valuable information about the cyber security analyst and will support the user-centered design and development of cyber situation awareness technology. For example, the taxonomy could be used during the design of cyber situation awareness visualization. Good design is especially important for large-scale visualizations that display large amounts of data. This taxonomy of cyber awareness questions would help inform the design of visualizations that would help analysts better establish and maintain situation awareness of a large computer network.

However, this study only addresses part of the picture. Our taxonomy does not include questions related to incident response while other models of cyber situation awareness do. The cyber security analysts in our study were specialized and only responsible for intrusion detection related activities as opposed to other research that studied generalists (e.g., [8]) or specific types of cyber security analysts (e.g., [2, 5, 7]). This may explain the lack of incident response topic area and questions in our taxonomy. Additional work in this cyber situation awareness will contribute additional questions and topic areas to the taxonomy.

References

1. Botta, D., Werlinger, R., Gagné, A., Beznosov, K., Iverson, L., Fels, S., Fisher, B.: Towards Understanding IT Security Professionals and Their Tools. In: ACM Symposium on Usable Privacy and Security, pp. 100–111 (2007)
2. D'Amico, A., Whitley, K., Tesone, D., O'Brien, B., Roth, E.: Achieving Cyber Defense Situational Awareness: A Cognitive Task Analysis of Information Assurance Analysts. In: Human Factors and Ergonomics Society Annual Meeting, pp. 229–233 (2005)
3. D'Amico, A., Whitley, K.: The Real Work of Computer Network Defense Analysts. In: Symposium on Visualizations for Computer Security, pp. 19–37 (2007)
4. Endsley, M.R.: Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors 37(1), 32–64 (1995)
5. Goodall, J.R., Lutters, W.G., Komlodi, A.: Developing expertise for network intrusion detection. Information Technology & People 22(2), 92–108 (2009)
6. Hudson, W.: Card Sorting. In: Soegaard, M., Dam, R. (eds.) The Encyclopedia of Human-Computer Interaction, 2nd edn. The Interaction Design Foundation, Aarhus (2013)
7. Thompson, R.S., Rantanen, E.M., Yurcik, W.: Network Intrusion Detection Cognitive Task Analysis: Textual and Visual Tool Usage and Recommendations. In: Human Factors and Ergonomics Society Annual Meeting, pp. 669–673 (2006)
8. Werlinger, R., Muldner, K., Kawkey, K., Beznosov, K.: Preparation, detection, and analysis: the diagnostic work of IT security incident response. Information Management & Computer Security 18(1), 26–42 (2010)