Paralyzed or Compromised: A Case Study of Decisions in Cyber-Physical Systems

Håvard Jakobsen Ofte(✉) and Sokratis Katsikas

Norwegian University of Science and Technology, Teknologiveien 22, 2815 Gjøvik, Norway
{havard.ofte,sokratis.katsikas}@ntnu.no

Abstract. Human operators of Cyber-Physical Systems (CPSs) within Critical Infrastructure (CI) need to protect their systems from cyber-attacks. When CPSs are compromised the operators might be faced with the dilemma of letting the systems be compromised to maintain the operation of CPSs or to paralyze the CPSs to mitigate the attack. How human operators resolve this dilemma was investigated through a case study of the Sunburst attack within the electrical power and manufacturing CI in Norway. Four actors were interviewed regarding the dilemma, including three actors interviewed regarding their handling of the Sunburst case. The interviews with additional incident reports from one of the actors were analyzed inductively to identify how the human operators made decisions in this context. Ten themes were identified and synthesized into a logic model of the decision process. The logic model was then compared to existing theoretical models of Situation Awareness (SA) to assess if SA theory could explain the findings. This study concludes that existing SA models are compatible with the findings. Some parts of the logic model based on the findings provide unique contributions to the understanding of the decisions. One important finding is that the design of the systems related to CPSs must allow adequate mitigation alternatives. The study highlights several implications for practice and further research. Although the findings may not be generalizable beyond the setting of the case, the study contributes to bridging the recognized research gap of empirical studies of the SA of human operators of CPSs.

Keywords: Cyber-physical systems · Security · Situation Awareness · Sunburst attack

1 Introduction

The rise of Information Technology (IT) during the last decades has been characterized by the virtualization, digital processing, and efficient transfer of information. This has made us associate technological development with systems of the logical domain. Technology before the rise of IT, was in contrast first and foremost associated with the mechanization of physical processes. The industrial revolution created technology for mass manufacturing, effective means of transport and enabled societies to rely on large scale systems of energy production and consumption. Looking back through history, we can recognize that the industrial revolution introduced technology for the physical

A. Moallem (Ed.): HCII 2024, LNCS 14729, pp. 134–152, 2024.
https://doi.org/10.1007/978-3-031-61382-1_9

domain and the IT revolution introduced technology for the logical domain. We are now entering a time of two parallel disruptive developments that challenge both of these historical technological paradigms. One is the introduction of technology-based judgement and subjecthood i.e., Artificial Intelligence (AI). The second development is the merging of technology for the physical domain with technology for the logical domain. This last technological amalgamation has been termed Cyber-Physical Systems (CPSs) [1].

CPSs are characterized by the interaction of cyber and physical components like sensors, processors, and actuators in order to adaptively control physical systems through logical processing, often with the ambition of creating autonomous systems [2]. CPS systems are found within several domains like automated driving, automated medical devices, and Critical Infrastructure (CI) like power plants, distribution of electrical power and smart manufacturing [3]. CIs like power plants and industrial manufacturing are by definition important for society and must therefore be ensured to work as intended. In addition, they can pose serious threats to human safety and to the environment if they malfunction [4].

With the integration of Information Technology (IT) and Operational technology (OT) these infrastructures are increasingly becoming CPSs [5]. This poses a new and complex layer of risks [6]. Within IT the risks posed to the technology is often connected to cyber security. When IT is integrated into OT, one invites the possibility of threats against reliable operation and safety as a consequence of cyber security breaches [7]. The research community has pointed out that existing guidelines for cyber security are not sufficient to meet these challenges [8]. Mitigation of cyber intrusion often includes the shutdown or at least logical isolation of IT assets. But when these assets also control critical physical processes like dams or smelters, the option of shutting them down entails something fundamentally different than unavailable IT services [7].

In the crux of this dilemma is the human operator weighing the options and making the crucial decisions. These decisions must often be made fast and without sufficient information [9]. Such situations challenge the operator's awareness of the available information and the suitability of potential actions. When the cyber security suddenly fails and there has been a breach in the CPSs, a decision must be made. Should the systems be shut down preventing further compromise but leaving the physical processes paralyzed? Or should one leave the system compromised and keep physical processes operative? This study investigates this posed dilemma.

Situation awareness (SA) has been a successful theoretical framework within human factors and decision making among human operators in several critical domains [10]. Situation awareness has also been researched within the field of cyber security [11]. Situation awareness is therefore well suited as a theoretical lens for understanding the combined challenge of human operators negotiating between cyber security and operation safety or reliability within CIs. Because little research regarding this challenge currently exists, this study will be conducted as a case-study investigating how these dilemmas are negotiated in the real world today, combined with an analysis regarding how SA can be used as an explanatory model. Based on these goals the following research questions are posed in this study:

- Research Question 1 (RQ1): How do human operators of cyber-physical systems in critical infrastructure decide between continuing to use or stop using systems that might be compromised?
- Research Question 2 (RQ2): How can these decision processes be explained by existing theory of cyber-situation awareness?

The research questions are investigated through a case study of the Sunburst attack [12] within electrical power and manufacturing CI in Norway. Several actors involved in the decisions regarding continued use of potentially compromised systems within CPSs were interviewed. The collected data from the interviews was combined with the incident reports from one of the actors. This combined data was used to describe the case and how decisions regarding the dilemma were made. The results of this investigation were then compared with existing SA models to analyze if SA is suited as an explanatory framework for the human decision making. This study thus contributes to a better understanding of the posed dilemma and how human operators resolve it. This improved understanding raises important practical implications for decision making in this context as well as important implications for further research.

2 Related Work and Background

In this section the concept of CPS is defined and presented in the context of CI. Related work regarding CPS security and safety is also presented. Then research works related to human factors of CPSs are highlighted. Lastly, in this section the theoretical foundation of SA is presented and linked to the context of this case.

2.1 CPSs

The term CPS is attributed to Helen Gill in 2006 and has since been widely adopted. A CPS can be defined as: "a system that can effectively integrate cyber and physical components using the modern sensor, computing and network technologies" [3]. CPSs are quite diverse and found in several domains [2] like industry 4.0 [5], medical CPS devices [13], automated driving [14], and smart grids [15]. It is an important challenge to merge the different fields of knowledge and practice stemming from different parts of CPSs diverse technological ancestry [1]. One central issue of this discussion is the negotiation of different principal guidelines regarding security. Some studies highlight this negotiation by distinguishing between information security for the IT domain and control security for the OT domain [2]. Others point to differences in the security focus related to the CIA triad for OT (availability) and IT (confidentiality) [5]. Yet, others highlight the need for alignment between security (meaning the protection against malicious actors) and safety (meaning protection against failure of physical systems) [16].

There is currently a need for research on how to make decisions for CPSs that involve, at the same time, criteria from different fields of expertise. The common research approach is to deal with physical systems security and cyber security separately, and there are only a few studies that attempt to reconcile the two [7].

One approach to the security of CPSs within CI is the structured approach to designing CPSs provided by the Purdue model [17]. The Purdue model consists of five zones

and six levels of operations that segment the controls and networks within industrial control systems [18]. The Purdue model highlights the need to carefully plan and design the connection between IT services and industrial control systems. The Purdue model thus contributes to the reconciliation of IT and OT security with guidelines on how to design system architecture. The Purdue model is also referenced in several international standards [17].

In this case study the following perspective on CPS security of CI is used: The threats of compromise are posed to the CPS through the IT domain and adversaries can attack the systems using known mechanisms researched in the field of cyber security [2]. Such cyber threats can manifest throughout the architecture of CPSs from the physical layer to the application layer [3]. However, the cyberthreats in this context have an added dimension of potential negative consequences related to the physical operations that characterize CPSs. A cyber threat thus has the potential to make the physical operations malfunction, be unreliable or in a worst-case scenario pose a threat to human or environmental safety [15]. The decisions regarding what potential actions might be most suitable in response to the threat are further complicated by the fact that the mitigations themselves also have potential negative consequences.

Common mitigation strategies in cyber security are isolation or shutdown of IT-assets [19]. Within the context of CPSs in CI such mitigations can themselves pose threats against reliable operation or safety. This is at the heart of the posed dilemma and the topic of this study is thus one example of how CPSs pose new challenges regarding the prioritization of different security principles. In existing standards and guidelines such dilemmas between principles are mentioned, but there are only general guidelines on how to handle the processes of decision. One example is how NIST acknowledges the posed dilemma followed by a recommendation of risk assessment as the method to decide: *"For example, one possible response option is to physically isolate the system under attack. However, this may have a negative impact on the OT and may not be possible without impacting operational performance or safety. A focused risk assessment should be used to determine the response action."* [19].

2.2 Human Operators in CPSs

In this study it is assumed that human operators are the actors that make such decisions. Based on the descriptions of CPSs in the research literature, this is not a given. The definitions of CPSs are agnostic regarding the need for human operators. This issue is heatedly debated in many academic circles. Techno-optimists are excited by the utopian possibilities of completely autonomous systems, whereas alarmists proclaim the imminent end of our civilization if we give up our human control of CPSs to AI [20]. Within the research literature the discussion of the necessity of human operators is often described as keeping the human-in-the-loop [21]. When the human is in-the-loop of CPSs it entails another layer of complexity regarding how the human interacts with the system, and it introduces the risk of human errors [22]. The field of human factors research has analyzed these dynamics and identified solutions for improving human performance within such systems [23]. However, when reviewing the existing human factors research literature regarding the posed dilemma, there are only a few studies that brush the topic [9, 24, 25].

2.3 Situation Awareness

Regarding the human factors of operating critical physical systems, SA is a highly recognized theoretical framework. SA has investigated human performance of operators in several critical sectors like nuclear power plant control, aviation, military operations, and surgical practices [26]. The most recognized definition of SA within human factors is the following by Endsley: *"The perception of the elements in the environment, the comprehension of their meaning, and the projection of their status in the near future"* [27].

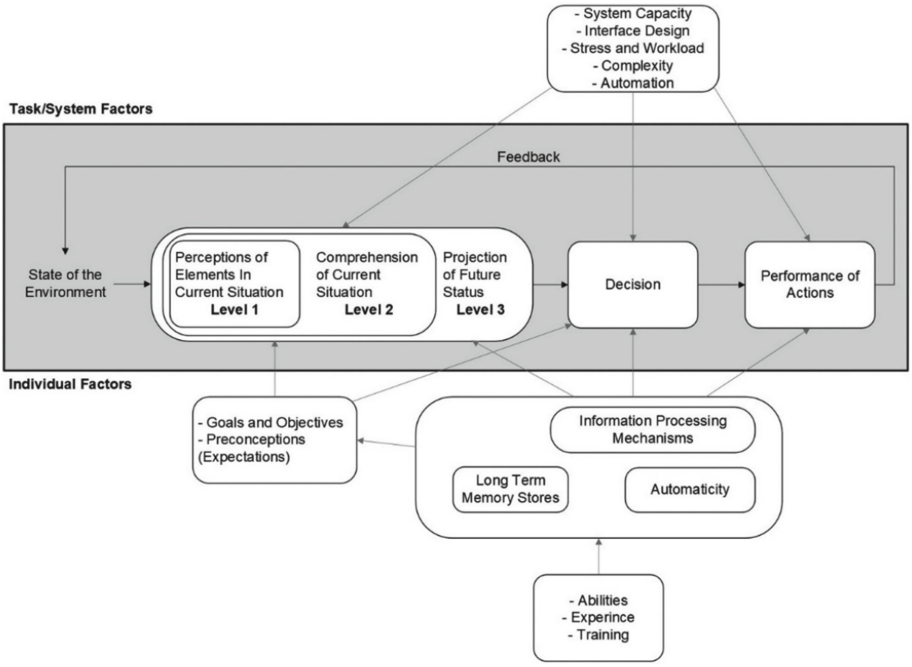

Fig. 1. Endsley's three level SA model [27]

The theoretical model of SA presented in Fig. 1 is based on cognitive psychology. At its core it understands the human operator's SA as three levels of cognition. The three levels are perception, comprehension, and projection. Human operators perceive elements of the situation which is cognitively processed to gain comprehension. Then the human operators project the situation cognitively into the future to gain awareness. The resulting awareness is used to assess the suitability of different actions. The process continues with operators making decisions and performing actions. These actions affect the environment and provide feedback loops for the operator through perception of the environment. The SA process is influenced by external factors related to the task or system as well as individual factors related to the human operator him/herself [28].

SA has also been used as a theoretical framework within the field of cyber security and is then often termed cyber-SA [29]. There is nevertheless a lack of empirical research

regarding cyber-SA [11]. The existing research has dominantly been based on Endsley's theory, and this model is thus the de-facto explanatory model within cyber-SA [9].

This study aims to investigate if the existing SA theory might provide a good explanatory framework for how human operators can make well informed decisions regarding the posed dilemma of CPS security in CI. The use of SA as a theoretical lens for examining these questions is suggested in existing literature [9, 24].

2.4 The Sunburst Attack

In late 2020 it became known that the Texas-based SolarWinds network monitoring and management system had been compromised. The attack on SolarWinds dubbed Sunburst exploited a vulnerability in the update system of the module called Orion. Malicious code embedded in official Orion updates created a backdoor into the system used worldwide by public and private organizations [30]. The backdoor was used for data extraction and in some cases for inserting additional malware into affected systems. The attack seemed aimed at exfiltrating confidential and sensitive data aided by spyware inserted through the backdoor [31]. The Sunburst attack affected at least 18. 000 organizations worldwide, but the exploitation of the backdoor seems to be aimed at US entities. The attack compromised systems within governmental bodies like the Department of Defense, the Department of Commerce, and the Department of Energy, including the National Nuclear Security Administration [30]. When the news of the vulnerability was made public it soon became known that the backdoor had already been present within official update versions of Orion for several months [32].

In December 2020 organizations all over the world got notifications from SolarWinds regarding the vulnerability of Orion. Organizations operating CPSs within CI were suddenly confronted with the possibility of having compromised systems within their operation. Research done in a digital infrastructure preparedness organization in Norway found that the Sunburst attack exposed a lack of predefined responsibilities in such situations [33]. Existing research literature on the Sunburst attack mostly focuses on the compromised high-profile US governmental bodies, and that Microsoft had some of its source code exfiltrated [30]. Yet, this attack case also provides a specific case of the dilemma presented in this paper.

3 Method

The defined research questions were examined through a case study following an established methodology [34]. The process of defining the case involved specifying the relevant criteria for the case. The case had to meet the following criteria: (a) The case had to include compromised IT systems in a setting of CPSs within CI; (b) The case had to include human operators faced with the dilemma presented in the research questions; (c) The human operators had to be in a setting where it would be reasonable to compare with the existing SA theory; (d) The case had to be of such a nature that access to data and participants was possible. Based on these criteria a preliminary interview and an in-depth interview of decision making in such cases were conducted with respondents within Norwegian CI. The preliminary interview included the posing of the criteria and

the research questions and asking if the respondent knew of such a case. The in-depth interview asked about decision making processes in cases following the defined criteria. Based on these two interviews the Sunburst attack was chosen as the case for further investigation. An overview of the attack is presented in Sect. 2.4. This case was chosen because it involved the acknowledged dilemma of continued use of a potentially compromised system for monitoring networks throughout several CIs.

The case study is conducted as a single-case explanatory design, following recommended guidelines [34]. The case was investigated through three in-depth interviews with respondents from three different organizations. Three of the interviews examined this through the Sunburst case. The interviews were conducted as semi-structured interviews in Norwegian following an interview guide developed based on the research questions and the defined case. In addition, the incident reports from respondent 1 (see Table 1) related to the case were reviewed as part of the case database. A description of all participating actors as well as respondents is presented in Table 1. The actors and respondents are anonymized to preserve confidentiality.

Table 1. Respondents and provided data.

Respondent number	Actor description	Description of respondent	Provided data
1a	Security Operations Center (SOC) providing services to customers in Critical Infrastructure	SOC Director, over 10 years of relevant experience	Preliminary interview for identifying case, incident reports on Sunburst case
1b		Part of the SOC incident response team, over 10 years of relevant experience	1 in-depth case interview on decision making in the Sunburst case
2	Large actor within Norwegian critical infrastructure (manufacturing)	Security executive, over 10 years of relevant experience	1 in-depth interview on related decision making
3	National security agency within Norwegian critical infrastructure	Part of the national incident response team. Over 10 years of relevant experience	1 in-depth case interview on decision making in the Sunburst case
4	Large scale actor within Norwegian critical infrastructure (power sector)	Security executive, over 10 years of relevant experience	1 in-depth case interview on decision making in the Sunburst case

The data analysis for this case study followed a four-step process adhering to recommended guidelines [34]: (a) The data was gathered and organized into a database for the case, this included the transcription of interviews; (b) dissembling the data by coding

and categorizing it into meaningful units; (c) reassembling the data, by creating a case description and an inductive explanation of the decision making process of the case; (d) interpreting the data, by comparing and contrasting the findings with the existing theoretical framework of the SA model presented in Fig. 1. The analysis process also used guidelines for the thematic analysis method when coding the meaningful units and aggregating them to themes [35]. In addition, the synthesized results from the inductive analysis of the data (a-c) were presented as a logical model of decision making [34]. The presented findings from the inductive analysis alongside the logic model, answered RQ1. The results from the inductive analysis were then used to investigate RQ2 according to guideline (d) analyzing whether existing SA models could explain the findings.

All respondents participated based on informed consent. As part of the ethical considerations the respondents were asked to approve the information presented about them and their organizations in the paper, to ensure that no sensitive information was disclosed. The quotes included in the paper were translated from Norwegian to English. The respondents were given the opportunity to revise the formulation of their own quotes. The quotes where reformulated through written correspondence with the respective respondents when they found the translated quotes misrepresented their intended meaning.

4 Findings

Based on the criteria for the case and the preliminary interviews the case of Sunburst within Norwegian CI was chosen. This case is presented from the perspective of the different actors in the following subsection. Then a thematic description of the decision-making processes will be given. The themes are then synthesized in a logical model of how the human operators decide on actions in response to compromised CPSs. The findings and the logical model are then compared with existing SA models.

4.1 The Sunburst Case from the Actors' Perspective

The SOC (Respondents 1) was responsible for SolarWinds systems in several customers. The SOC registered an advisory bulletin regarding a vulnerability in SolarWinds Orion after regular office hours on December 13, 2020. In the early morning before office hours of December 14, the SOC receives a mail from SolarWinds warning about the Sunburst attack. The SOC does not immediately recognize the severity of the attack, but this is gradually understood during the morning of December 14, while the SOC also communicates with the national security agency establishing a common understanding of the attack. Customers within CI using SolarWinds are notified throughout the morning by the SOC, by forwarding the mail from SolarWinds. In the afternoon of December 14, the SOC escalates the incident and activates its response team and procedures to the fullest extent. There were publicly available descriptions of how to verify if a system was compromised. SOC operators used these descriptions to verify if customers were compromised; meanwhile SOC specialists verified the available description of the malware. Throughout the following days the SOC conducted intense incident response activity involving security measures like isolation, patching and communication with

stakeholders. This was done as a prioritized effort based on the criticality of compromised systems. On December 22 the SOC concludes its incident response and only follows up with customers regarding patching and incident evaluations.

The national security agency (Respondent 3) was made aware of the attack approximately at the same time as Respondent 1. Their first response was to verify the reality of the attack. This was done throughout December 14 in close communication with Respondent 1, among others. Warnings to CI owners within the power sector were issued, both in written form and in online briefings. The warnings and briefings were done iteratively and contained increasing levels of details regarding how to identify whether systems were compromised and how to mitigate. When patches from SolarWinds were released, Respondent 3 also distributed advisory statements regarding patching all relevant systems. In many aspects the role of Respondent 3 throughout the attack was to provide all stakeholders within the Norwegian power sector with verified and updated information during the attack. They also served as a contact point between other stakeholders.

The owner of CI within the electrical power sector (Respondent 4) received notice of the attack on the morning of December 14. They quickly confirmed that their Orion systems were affected. Based on log files they recognized that they were affected before the message from Respondent 3 arrived. The incident response routines were activated throughout the organization. The affected systems were located so that critical CPSs could be reached through them. The organization quickly segmented their networks in an effort to neutralize the attack. This was based on the information that was publicly available. After some time, it was decided to shut the compromised systems off within the most critical parts of their infrastructure. This was done after forensic material from the systems was secured. The unavailable systems resulted in reduced ability to monitor networks within these defined parts of their infrastructure, causing lowered ability to proactively operate their networks. When patches were made available, the organization decided to not patch existing systems within the most critical parts of the infrastructure, but to completely replace the affected systems.

Respondent 2 was not interviewed regarding the Sunburst case. The in-depth interview with Respondent 2 was aimed at decision making processes in related situations.

4.2 Themes Identified in the Analysis

The following is a presentation of all the themes that were identified through the analysis of the interview transcripts and the incident report. The themes are presented with a general description exemplified by respondent quotes.

Understanding Asset Topology and Functions. Several of the respondents describe how important it is that the topology and functions of the CPSs and connected systems are documented and well understood: *"You need to understand your own infrastructure. You need to comprehend what you have internally, and you need to take responsibility for it." Respondent 3.* The understanding of the functionality and topology of the assets is a prerequisite for assessing and performing mitigations: *"When we have insight into how the network is set up, how the infrastructure is, it's easier for us to evaluate quickly. For many of the customers this was the case. For some of the other customers we only*

knew that they had the system, but we didn't know much about their network. In those cases, it was essentially our responsibility just to inform and give general guidance to them because we had no insight into their environment." Respondent 1b. We see how the SOC consider their ability to respond to attacks to be dependent on their level of insight into customers systems.

System Design. The theme of system design of CPSs was a common theme throughout the interviews. All the respondents highlighted that in order to respond adequately to threats and attacks the architecture of the CPSs had to be designed in a manner that enabled mitigation alternatives. When asked what design principles they used, the respondents pointed to the Purdue model: *"Are you familiar with the Purdue model? It involves organizing your architecture into distinct levels. You segment the system into smaller closed zones. If any issues have propagated within that closed system and haven't extended beyond it, you've effectively minimized the extent of potential damage." Respondent 1b.* It was further highlighted that this approach to system design was demanding to implement in large scale existing infrastructure and that the process is dependent upon the understanding of asset topology and functions: *"It demands a great deal of transformation technically to do it. I think it's wise to have an overall plan with annual goals. You need an overview of all your assets based on quality documentation. It is possible, but it demands a lot to get it done." Respondent 3.*

Knowledge and Skills. The respondents point to the need for a wide range of knowledge and skills to respond to attacks aimed at CPSs in CI. They highlight that the adequate response often involves a high degree of insight into the physical processes that are controlled, alongside expertise knowledge regarding the threat or attack at hand. These need to be combined to gain an understanding of the potential impact of the attack: *"We had a rather good mix of personnel. Some with security expertise, others with networking knowledge, customer insights, and SolarWinds knowledge. In our emergency response team, we had a diverse blend of individuals who understood the different technology involved in the entire system. This allowed us to engage in meaningful dialogues where experts from their respective fields could assess how the attack would impact things from their viewpoint." Respondent 1b.* The knowledge and skills available during an incident like the Sunburst attack will also influence how the response is organized regarding roles and responsibilities.

Roles and Responsibilities. The respondents highlighted the need for predefined and clear roles and responsibilities during incidents like Sunburst. The roles that need to be defined include roles connected to IT security, continued operation of industrial systems and people with defined roles regarding business impact: *"It is clear the importance of having both operational representatives and business stakeholders present in such discussions or meetings. After all, they are the ones who shoulder the risk." Respondent 4.* Several of the respondents also point out the importance of having one responsible leader for incident response with the ability and authority to make difficult decisions and maintain control over the other roles: *"If you don't have the authoritarian figure who takes responsibility, delegates, and maintains some control, it's easy for things to veer off track. Then some people end up working in parallel and have overlapping roles." Respondent 1b.* The respondents did not base the roles on any pre-defined standards; on

the other hand, they highlighted that the roles were developed and optimized through training and exercises.

Training and Exercises. All respondents pointed to training and exercises as an essential preparation for mitigating attacks and handling security incidents. They also trained specifically for CPS-related scenarios: *"This is part of what we do, it's quite significant. The sites have their contingency plans or continuity plans. Naturally, we practice scenarios where we have to deactivate parts of the control system. This may be based on physical scenarios like floods or fires. And many of the sites naturally have plans to staff up, to compensate for the lack of system support." Respondent 2.* Another aspect of the training and exercises is that these are used to train on effective communication during an incident: *"When they have exercises, they practice on organizing the response and the chain of command. Because it's crucial in situations like emergencies so you don't spend a lot of time figuring out who should take responsibility or whom to contact. Knowing this is essential." Respondent 1b.* The respondents also talk about the importance of training for realistic incidents to be mentally prepared when incidents do occur.

Shared Understanding. When asked about what determines the effectiveness of incident response in CPSs the respondents highlight the importance of gaining a shared understanding of the situation. As described, the combination of different expertise and roles demands coordination. This coordination is successful when decisions are based on the best available information from several perspectives: *"If someone provides information that connects with other information, suddenly you see the bigger picture. It is important that people in a response team share information regardless if they think it is important or not. Information might prove important when received by someone with a different expertise or combined with information from others. We all sat together and worked in that way." Respondent 1b.* The sharing of different understandings was especially important when assessing the potential impact of incidents.

Assessing Impact. The first thing to assess during an attack is the nature of the attack and the potential impact it will have in the infrastructure that is attacked. The respondents explain that the quality of the information regarding an attack is key regarding effective mitigation. In the case of the Sunburst attack there was quite detailed information available when the attack was made known: *"When the news was released, it came with a full report on the vulnerability and how you could exploit it. It was a security company that had discovered the vulnerability, so you got a very comprehensive write-up. That way, you knew if the system was vulnerable and what indicators to look for." Respondent 1b.* In other attacks such information is not available and this adds to the complexity of responding.

Assessing Mitigation Alternatives. The respondents explained how the assessment of different possible mitigations was at the center of the response decisions. These assessments were reliant upon information gathered through the processes described in previously presented themes. The assessment consists of identifying the potential mitigations and then assess them regarding their effectiveness and what the mitigations will cost in terms of loss of functionality: *"One knows that if you take the system offline or make any changes to it, it creates significant challenges in terms of operations. From a security perspective, one might not fully understand this or have an overview of the*

consequences it will have. So, that is perhaps the difficult part here as well. From the security side, this threat is viewed as highly dangerous. However, the operations side may not perceive it that way." Respondent 4. The respondents also highlight that these assessments can cause prolonged response times because of uncertainty: *"The biggest challenge lies in explaining the situation and reaching a decision. Trying to take some action based on a concrete picture, so you're not just standing there waiting. Waiting, analyzing, and discussing. That's the real challenge—to maintain progress in handling it." Respondent 4.* The respondents conclude that decisive action is necessary to avoid this issue.

Decisive Action. The respondents explain that one will never have met one's need for information before decisions must be made regarding mitigating attacks against CIs. This according to the respondents must be met with decisive actions from a leader with authority: *"What is important, is that you have someone—a leader—who takes clear charge and has both the authority and the courage to cut through. They must genuinely drive decision-making and accept that decisions made can be wrong. Because you can always gather more information, and not be able to move forward—that is maybe the worst thing you can do." Respondent 4.* The respondents also made very clear that in cases of compromised CPSs in critical infrastructure the decisions are first and foremost aimed at the safe operation of infrastructure; IT security is a secondary priority: *"The ones making the major decisions directly impacting the plant operation must be based locally, because you need to fully understand what the consequences are. It is particularly important for the safety of personnel. You cannot have anyone manipulating the systems unless those that are responsible for safe operations are fully aware of the ramifications." Respondent 2.* This also highlights the importance of communication between stakeholders during incidents.

Communication with Stakeholders. One important aspect throughout the case is the communication between different stakeholders. All respondents communicated with stakeholders throughout the Sunburst attack. The different actors had different roles regarding such communication. The SOC focused on communicating the updated information to their customers: *"We had very good and close dialogue with the customers. For some we provided information about indicators they could look for themselves. Others we assisted in looking for the indicators. The level of assistance was based on a prioritization of criticality. Still, we maintained a close dialogue with all the customers throughout the incident." Respondent 1b.* The national security agency had a different role: they gave briefings for the whole power industry regarding the situation with verified information: *"Our first priority is to uncover, what is actually true regarding the attack? Does this match? We were in early contact with the local support for the software. We tried to coordinate information from briefing to briefing. We answered questions like: when is the next opportunity for updated information, when is the briefing; What do we do in the meantime? And we tried to get in contact with the supplier, which is not always easy because they're in a tough spot." Respondent 3.*

4.3 Logic Model of Decision Making

The identified themes from the interviews represent issues that the respondents find relevant and important regarding how to make decisions between continuing to use or stop using potentially compromised systems in CIs. To better explain the process of making such decisions, the themes are synthesized into a logical model. The model has two major parts; the first is the preparations that are done before an attack (left side) and the second is the response to the attack (right side). The preparations are split in two types, namely the technical aspects (top) and the human aspects (bottom). Within each of the aspect types the respective preparations correspond to the identified themes. The different parts of the preparations within each type are dependent upon each other in a step-by-step fashion shown by arrows. The two types of preparations affect each other. Likewise, the response to an incident is split into the two stages of shared understanding (top) and making and communicating decisions (bottom). The two stages consist of respective parts of the response process that corresponds to the identified themes. The stage of shared understanding (top, right) also corresponds to an identified theme. The dependence between parts of each stage is also shown by arrows. The stages are shown to potentially repeat cyclically by the arrows going both ways between them. The model also shows with arrows (from left to right) how different parts of the response process are dependent upon different parts of the preparations.

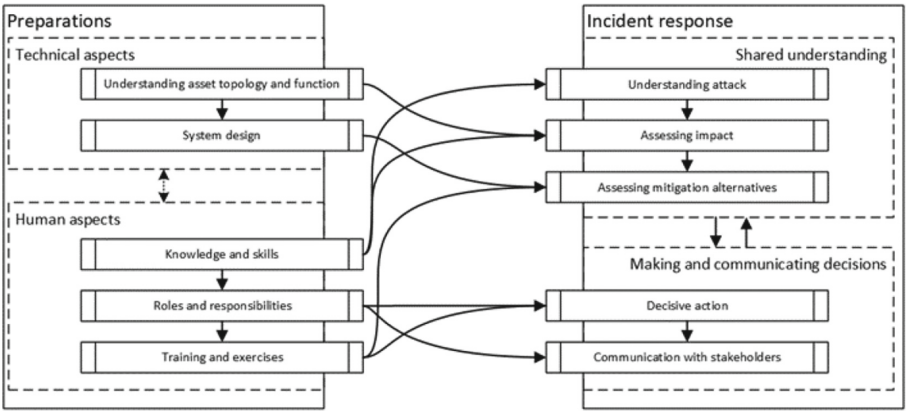

Fig. 2. Logic model of decision-making in CPSs

The model shows how decisions are made regarding to use or not to use systems that might be compromised. The decisions are based upon a shared understanding of the attack, its impact, and an assessment of the available mitigation alternatives. This shared understanding is dependent upon the technical understanding of asset topology and function, as well as on a system design that allows for adequate mitigation alternatives. The human operators use their combined knowledge to gain a shared understanding during an incident, and this process demands training and exercise to carry out effectively in collaboration. To implement the decisions, the operators need to take decisive action and communicate it effectively. The decisiveness is dependent on well-defined roles and

responsibilities honed through training and exercises. The actions need to be communicated to stakeholders and this communication also provides information regarding the often-needed iteration of a shared understanding.

5 Discussion

The model presented in Fig. 2 presents the findings from the inductive analysis of the case material. When we compare this to Endsley's model of SA [27] in Fig. 1, we see some clear similarities. The identified stage of shared understanding consists of three parts: understanding the attack, assessing impact, and assessing mitigation alternatives. If we compare this to the SA model, we can recognize that the understanding is achieved in levels in both models. The three parts of shared understanding in Fig. 2 do not completely match the SA model. One could argue that both understanding the attack and assessing impact are mostly related to comprehension (level 2) in the SA model. Assessing mitigation alternatives is highly related to projection (level 3) in the SA model. Level 3 in the SA model may also be overlapping with parts of assessing impact in the Logical model. The stage of making and communicating decisions in Fig. 2 overlaps to a large degree with decision and performance of action in the SA model. The similarity is especially clear if we consider the communication with stakeholders as overlapping with the feedback loop of the SA model.

When we consider the preparations, we see a strong similarity with the factors shown to influence the SA in Endsley's model. The technical aspects of the preparations highly overlap with the task/system factors in the SA model. The human aspects in Fig. 2 overlaps with the individual factors of the SA model. We can see that knowledge and skills and training and exercises in Fig. 2 are highly similar to abilities, experience, and training in Fig. 1. One could further argue that roles and responsibilities in Fig. 2 overlaps with goals and objectives and preconceptions in Fig. 1. In total the models are highly similar; this strengthens the argument that existing SA theory can explain the decision-making processes related to the dilemma posed in this study.

There are nevertheless some differences that are important to highlight. In the Endsley's SA model there is a clear individual perspective. The model explains how individual operators gain and use SA. The logic model presented in this study does not have an individual perspective, but rather an organizational perspective on decision making. This can to some extent explain the lack of themes connected to the cognitive processes like information processing mechanisms in the individual factors of the SA model. In addition, level 1 of the SA model is only partially overlapping with the logic model. The last point can arguably be explained by how participants were asked about their decisions in the Sunburst case. This might have reduced the respondents' focus on the monitoring of elements in the situation prior to the recognition of the attack.

One possible alternative explanation from existing SA theory that might resolve these discrepancies are SA models for groups [36]. One candidate is the Team SA model synthesized by Salmon et al., [37]. This model for example includes common/shared picture as part of the SA process. Still, the Team SA model is presented at a higher abstraction level, so it only indicates many of the details present in the logic model of Fig. 2. Another alternative explanation is the Distributed SA model [38]. This model

allows for a stronger connection between the technical aspects and shared understanding in Fig. 2. On the other hand, the Distributed SA model is less specific than Endsley's model regarding the process of decision making and action. Within the research literature investigating cyber-SA, Endsley's model is almost exclusively referenced regardless of the level of analysis i.e., even group and system level cyber-SA research use the individual model of Endsley as explanatory model [9]. There is thus a gap regarding theoretical SA models that are developed for explaining the type of decision-making processes investigated in this study. This leaves Endsley's model as the most fitting existing theoretical explanatory model.

The logic model in Fig. 2 is presenting specific dependencies between preparations and incident response regarding the posed dilemma that Endsley's model is too general to encompass. This is a unique contribution regarding understanding of how these dilemmas are resolved by the process of decision making. This case study shows how specific knowledge and skills are needed to understand attacks against CPSs and to assess their potential impacts. The knowledge and skills needed for these tasks are seldom found in individual operators alone. This is because the knowledge of IT-based attacks and the functionality of CPSs are two distinct expertise areas. Actors within CI should therefore organize and train on the collaboration between experts from these different fields of expertise.

Additionally, the importance of system design is something not captured by existing SA models. The respondents clearly stated that the mitigation alternatives one has during an attack on CPSs are totally dependent upon the design of one's systems. If the systems are designed in a way that allows for fast and secure segmentation in accordance with the guidelines of the Purdue model [17], the response has a far greater array of potential mitigation alternatives. The segregation or segmentation of networks are often far less invasive than shutdowns or manual control options. This is one of the most important focus areas for the respondents, because existing CI that are gradually converted to CPSs are often not adhering to these design principles. This leaves human operators with few mitigation alternatives that they know will be effective. The dilemma is in these cases very realistic where the decision is between letting the system be compromised or paralyzing the system through shutdowns or manual operation. When the CPSs are designed properly one can with confidence isolate parts of the systems and let other parts operate largely unaffected.

It is important to recognize that though the findings in this study provide unique contributions, these are only based on a quite restrictive case of operators of CPSs in CI. The study only investigated the case of the Sunburst attack. Had other attacks or cases been investigated, other findings might have resulted. Further the relative low number of respondents makes the findings less generalizable. This is made even more relevant given that the respondents all came from Norwegian CI actors, which is a specific context that may have affected the findings [39]. Despite these limitations the study presents empirically-based findings within a recognized research gap [40]. There are few studies examining incident response from the human operators' side, and even fewer studying the human SA in this setting [11]. This study provides clear recommendations for practice including the need for specific preparations regarding attacks against CPSs in CI as well as a more tailored presentation of the decision process itself in comparison with existing

SA theory. This study also implicates the need for further research regarding SA in such settings. A complete Goal-directed Task analysis of SA could complement and validate the findings of this study and provide a more generalizable model of decision making within this setting too [41].

6 Conclusion

This case study investigated the dilemma of human operators of CPSs in CI regarding continuing to use or stop using compromised systems to maintain CPS operations. This was investigated through a case study of the Sunburst attack [12] in Norwegian CI within the electrical power and the manufacturing sectors. An inductive analysis of interviews with four different actors and the incident reports from one of them answered the first research question (RQ1): How do human operators of cyber-physical systems in critical infrastructure decide between continuing to use or stop using systems that might be compromised? The inductive analysis identified ten distinct themes that explained how such decisions were made in the considered case. The findings were synthesized and presented in Fig. 2 as a logical model of the decision process. A deductive comparison of the logical model with existing theoretical SA models answered the second research question (RQ2): How can these decision processes be explained by existing theory of cyber-SA? The analysis compared Endsley's individual SA model [27] with the logical model and found many similarities and some discrepancies. Other existing models of SA were discussed as alternative explanations, but none of them explained the logical model better than Endsley's model.

The findings provide a unique contribution that explains the decision-making process of human operators of CPSs. The study thus contributes to bridging a recognized research gap regarding human operators of CPSs within CI. Although the explanations provided in this case study have limited generalizability, they provide clear indications for further research as well as implications for practice. The study shows how preparations of technical and human aspects directly affect the operators' ability to respond adequately to attacks against the CPSs. One important example is that the design of the systems provides the operator with mitigation alternatives. If the technical systems are designed in the right way, the operator may well be able to paralyze only a minor part of the CI to mitigate compromised systems. The possibility to adapt mitigation against attacks on CPSs seems like the most promising way forward. Then the potential harm of necessary mitigations against attacks will be minimized. This would arguably be even more important if we hand the control of CPSs in large scale infrastructure of critical importance over to AI in the future.

Acknowledgments. This work was supported by the Research Council of Norway (Norges Forskningsråd) under Project number 333900 "*Situation awareness in virtual security operations centers*" and Project number 310105 "*Norwegian Centre for Cyber Security in Critical Sectors (NORCICS)*".

Disclosure of Interests. The first author was employed in a research position with Respondent 1 (See Table 1.) at the time of this study. The authors had no other known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Baheti, R., Gill, H.: Cyber-physical systems. Impact Control Technol. **12**, 161–166 (2011)
2. Ashibani, Y., Mahmoud, Q.H.: Cyber physical systems security: analysis, challenges and solutions. Comput. Secur. **68**, 81–97 (2017). https://doi.org/10.1016/j.cose.2017.04.005
3. Alguliyev, R., Imamverdiyev, Y., Sukhostat, L.: Cyber-physical systems and their security issues. Comput. Ind. **100**, 212–223 (2018). https://doi.org/10.1016/j.compind.2018.04.017
4. Yaacoub, J.-P.A., Salman, O., Noura, H.N., Kaaniche, N., Chehab, A., Malli, M.: Cyber-physical systems security: limitations, issues and future trends. Microprocess. Microsyst. **77** (2020). https://doi.org/10.1016/j.micpro.2020.103201
5. Kayan, H., Nunes, M., Rana, O., Burnap, P., Perera, C.: Cybersecurity of industrial cyber-physical systems: a review. ACM Comput. Surv. (CSUR) **54**, 1–35 (2022). https://doi.org/10.1145/3510410
6. Lezzi, M., Lazoi, M., Corallo, A.: Cybersecurity for industry 4.0 in the current literature: a reference framework. Comput. Ind. **103**, 97–110 (2018). https://doi.org/10.1016/j.compind.2018.09.004
7. El-Kady, A.H., Halim, S., El-Halwagi, M.M., Khan, F.: Analysis of safety and security challenges and opportunities related to cyber-physical systems. Process. Saf. Environ. Prot. **173**, 384–413 (2023). https://doi.org/10.1016/j.psep.2023.03.012
8. Akbarzadeh, A., Katsikas, S.: Unified IT&OT modeling for cybersecurity analysis of cyber-physical systems. IEEE Open J. Ind. Electron. Soc. **3**, 318–328 (2022). https://doi.org/10.1109/ojies.2022.3178834
9. Ofte, H.J., Katsikas, S.: Understanding situation awareness in SOCs, a systematic literature review. Comput. Secur., 103069 (2022). https://doi.org/10.1016/j.cose.2022.103069
10. Stanton, N.A., Salmon, P.M., Walker, G.H., Salas, E., Hancock, P.A.: State-of-science: situation awareness in individuals, teams and systems. Ergonomics **60**, 449–466 (2017). https://doi.org/10.1080/00140139.2017.1278796
11. Gutzwiller, R., Dykstra, J., Payne, B.: Gaps and opportunities in situational awareness for cybersecurity. Digit. Threats Res. Pract. **1** (2020). https://doi.org/10.1145/3384471
12. Willett, M.: Lessons of the SolarWinds hack. Survival **63**, 7–26 (2021). https://doi.org/10.1080/00396338.2021.1906001
13. Dey, N., Ashour, A.S., Shi, F., Fong, S.J., Tavares, J.M.R.: Medical cyber-physical systems: a survey. J. Med. Syst. **42**, 1–13 (2018). https://doi.org/10.1007/s10916-018-0921-x
14. Kim, K., Kim, J.S., Jeong, S., Park, J.-H., Kim, H.K.: Cybersecurity for autonomous vehicles: review of attacks and defense. Comput. Secur. **103**, 102150 (2021). https://doi.org/10.1016/j.cose.2020.102150
15. Yohanandhan, R.V., Elavarasan, R.M., Manoharan, P., Mihet-Popa, L.: Cyber-physical power system (CPPS): a review on modeling, simulation, and analysis with cyber security applications. IEEE Access **8**, 151019–151064 (2020). https://doi.org/10.1109/access.2020.3016826
16. Aven, T.: A unified framework for risk and vulnerability analysis covering both safety and security. Reliab. Eng. Syst. Saf. **92**, 745–754 (2007). https://doi.org/10.1016/j.ress.2006.03.008
17. Boyes, H., Hallaq, B., Cunningham, J., Watson, T.: The industrial internet of things (IIoT): an analysis framework. Comput. Ind. **101**, 1–12 (2018). https://doi.org/10.1016/j.compind.2018.04.015

18. Obregon, L.: Secure architecture for industrial control systems. SANS Institute, White Paper (2015)
19. Stouffer, K., et al.: Guide to operational technology (OT) security. NIST Special Publication, 800-882, Rev. 803 (2023). https://doi.org/10.6028/NIST.SP.800-82r3
20. Turchin, A., Denkenberger, D.: Classification of global catastrophic risks connected with artificial intelligence. AI Soc. **35**, 147–163 (2020). https://doi.org/10.1007/s00146-018-0845-5
21. Nunes, D.S., Zhang, P., Silva, J.S.: A survey on human-in-the-loop applications towards an internet of all. IEEE Commun. Surv. Tutorials **17**, 944–965 (2015). https://doi.org/10.1109/comst.2015.2398816
22. Jirgl, M., Bradac, Z., Fiedler, P.: Human-in-the-loop issue in context of the cyber-physical systems. IFAC-PapersOnLine **51**, 225–230 (2018). https://doi.org/10.1016/j.ifacol.2018.07.158
23. Kadir, B.A., Broberg, O., da Conceicao, C.S.: Current research and future perspectives on human factors and ergonomics in industry 4.0. Comput. Ind. Eng. **137**, 106004 (2019). https://doi.org/10.1016/j.cie.2019.106004
24. Carreras Guzman, N.H., Wied, M., Kozine, I., Lundteigen, M.A.: Conceptualizing the key features of cyber-physical systems in a multi-layered representation for safety and security analysis. Syst. Eng. **23**, 189–210 (2020). https://doi.org/10.1002/sys.21509
25. Pinto, R., Gonçalves, G., Tovar, E., Delsing, J.: Attack detection in cyber-physical production systems using the deterministic dendritic cell algorithm. In: 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1552–1559. IEEE (2020). https://doi.org/10.1109/etfa46521.2020.9212021
26. Endsley, M.R., Garland, D.J.: Theoretical underpinnings of situation awareness: a critical review. Situation Awareness Anal. Meas. **1**, 3–21 (2000)
27. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors **37**, 32–64 (1995). https://doi.org/10.1518/001872095779049543
28. Endsley, M.R.: Designing for Situation Awareness: An Approach to User-Centered Design. CRC Press (2016). https://doi.org/10.1201/9780203485088
29. Jajodia, S., Liu, P., Swarup, V., Wang, C.: Cyber Situational Awareness. Springer, New York (2009). https://doi.org/10.1007/978-1-4419-0140-8
30. Alkhadra, R., Abuzaid, J., AlShammari, M., Mohammad, N.: Solar winds hack: in-depth analysis and countermeasures. In: 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1–7 (2021). https://doi.org/10.1109/ICCCNT51525.2021.9579611
31. Coco, A., Dias, T., van Benthem, T.: Illegal: the SolarWinds hack under international law. Eur. J. Int. Law **33**, 1275–1286 (2022). https://doi.org/10.1093/ejil/chac063
32. Martínez, J., Durán, J.M.: Software supply chain attacks, a threat to global cybersecurity: SolarWinds' case study. Int. J. Saf. Secur. Eng. **11**, 537–545 (2021). https://doi.org/10.18280/ijsse.110505
33. Aakre, S., Aarland, M.: Når en høypålitelig organisasjon blir utsatt for en normalulykke. Praktisk økonomi finans **39**, 34–47 (2023). https://doi.org/10.18261/pof.39.1.4
34. Yin, R.K., Campbell, D.T.: Case Study Research and Applications: Design and Methods. SAGE Publications, Inc., Thousand Oaks, California (2018)
35. Braun, V., Clarke, V.: Thematic Analysis. American Psychological Association (2012)
36. Kaber, D.B., Endsley, M.R.: Team situation awareness for process control safety and performance. Process. Saf. Prog. **17**, 43–48 (1998). https://doi.org/10.1002/prs.680170110
37. Salmon, P.M., et al.: What really is going on? Review of situation awareness models for individuals and teams. Theor. Issues Ergon. Sci. **9**, 297–323 (2008). https://doi.org/10.1080/14639220701561775

38. Stanton, N.A., et al.: Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. Ergonomics **49**, 1288–1311 (2006). https://doi.org/10.1080/00140130600612762
39. Gjesvik, L.: Comparing Cyber Security. Critical Infrastructure Protection in Norway, the UK and Finland. NUPI Report (2019)
40. Gil, M., Albert, M., Fons, J., Pelechano, V.: Engineering human-in-the-loop interactions in cyber-physical systems. Inf. Softw. Technol. **126**, 106349 (2020). https://doi.org/10.1016/j.infsof.2020.106349
41. Endsley, M.R., Connors, E.S.: Foundation and challenges. In: Cyber Defense and Situational Awareness, pp. 7–27 (2014). https://doi.org/10.1007/978-3-319-11391-3_2