Cyber crisis averted: using safety science principles to learn from success

Jeroen Wolbers
Leiden University
j.j.wolbers@fgga.leidenuniv.nl

Tommy van Steen
Leiden University
t.van.steen@fgga.leidenuniv.nl

Cristina Del-Real
Leiden University
c.del.real@fgga.leidenuniv.nl

Bibi van den Berg
Leiden University
b.van.den.berg@fgga.leidenuniv.nl

ABSTRACT

This paper aims to bridge the gap between crisis management and cyber security research. Current research often focuses on learning from failures, neglecting insights from successful interventions. This paper advocates for expanding the focus to include lessons from what works, particularly in preventing and mitigating cyber incidents. Drawing from safety science, we propose a research agenda combining Safety-I (accident analysis) and Safety-II (adaptation) approaches. We highlight three promising research areas: 1) identifying security practices that prevent cyber crises, 2) fostering high-reliability traits in organizations to enhance foresight, 3) and examining the role of near-misses in cyber security. This approach aims to generate valuable insights for both academics and practitioners in cyber security and crisis management.

Keywords

Cyber security, safety, high-reliability organizations, accidents, crisis management.

INTRODUCTION

In an era marked by rapid technological advancement and interconnected systems, the cyber landscape is evolving at an unprecedented pace. Traditionally, the field of crisis management has concentrated on physical incidents, emergencies, and disasters (Boin et al., 2018). Currently, the field is increasingly recognizing the disruptive threats cyber crises pose to our societies (Backman, 2021). We now find ourselves navigating a new frontier where the once-clear boundaries between the physical and digital domains are increasingly blurred. Attacks and incidents in cyberspace can have far-reaching consequences. A cyber crisis can be seen as a transboundary crisis in which exploitation of a vital cyber asset can cause serious damage or disruption to critical societal functions (Prevezianou, 2021: 65). Especially attacks on critical infrastructures can have severe impacts on society, triggering cascading effects that may disrupt supply chains, electricity grids, financial systems, or water supplies.

There are plenty of prominent examples to illustrate the negative consequences of cyber incidents. In 2017, the WannaCry ransomware attack affected hundreds of thousands of computers across 150 countries, disrupting essential services such as healthcare and transportation (Cisco Talos, 2017). In the same year, the NotPetya malware attack targeted businesses worldwide, with some estimates rising the number of affected organizations to more than 2,300 (Kaspersky Daily, 2017). Among them, NotPetya affected 4,000 servers, 45,000 PCs, and 2,500 applications in Maersk, a global shipping giant. The company incurred losses between $250 million and $300 million, while the global economic impact reached an estimated $10 billion to $11 billion (Wired, 2018). The attack demonstrated the fragility of supply chains and third-party software security risks, and showed how digital threats can disrupt real-world logistics. The Colonial Pipeline ransomware attack in May 2021 highlighted the vulnerability of critical infrastructure to cyber threats. This incident, carried out by the hacker group DarkSide, led to the shutdown of 5,500 miles of pipeline, causing fuel shortages and price spikes across the Southeastern United States (Reuters, 2021). Additionally, the 2020 SolarWinds cyber-attack, in which hackers infiltrated numerous U.S. government and corporate networks, resulted in widespread vulnerabilities and triggered a massive national security response (CSO Online, 2021). Collectively, these incidents demonstrate how cyber-attacks can disrupt essential services, erode public confidence, and demand a coordinated response to mitigate their effects.

Despite the growing threat and impact of cyber incidents, they often remain shrouded in mystery within the crisis management domain (Prevezianou, 2021). Conversely, cyber security literature frequently emphasizes prevention and focuses on technical solutions, neglecting the valuable insights that crisis management can provide for responding to cyber incidents (Baskerville et al., 2014). It is imperative to increase the familiarity between these literatures, fostering a more comprehensive understanding that transcends the traditional physical-digital divide.

We do so by heading into a new direction. A key issue that is hampering both literatures is that we tend to focus on learning from failure by analyzing key incidents, attacks, crises and disasters. The names of major cyber crises mentioned earlier likely sound familiar because they capture our attention and inspire scholars to analyze and learn from what went wrong. A similar trend is evident in crisis and disaster management, where catastrophic events such as Hurricane Katrina, 9/11, COVID-19, and Fukushima (Wolbers et al., 2021) have attracted significant scholarly attention. Major catastrophes trigger our imagination and are a pull factor for other disciplines to study crisis management.

However, a critical problem shared by both crisis management and cyber security literature is that a predominant focus on major incidents and failures overlooks valuable lessons from successful interventions that prevent crises from occurring in the first place. We must also learn from what goes right: from successful prevention strategies. The potential for such learning is significant, as it is remarkable how few cyber incidents occur, given the highly digitalized and interconnected world we live in (van Steen et al., 2024).

In order to learn from success and examine effective prevention strategies, we draw parallels from the field of safety science. Similar to cyber and crisis management, safety science has traditionally focused on learning from mistakes to prevent future incidents, which became known as the 'Safety-I approach' (Hollnagel, 2013). However, scholars in this field recognized that focusing solely on incidents disregarded lessons from countless instances where issues were successfully managed or circumvented. As Hollnagel et al. (2015: 20) note, *'the surprise is not that things occasionally go wrong, but that they go right so often'*. Instead, a growing number of scholars emphasizes the importance of learning from how people adapt to handle complexities, respond to challenges, and synchronize activities to achieve shared goals (Provan et al., 2020). This approach, which examines how people actually perform their work to manage variable circumstances, is known as the 'Safety-II approach.' We believe that cyber security and crisis management can gain valuable insights from this shift in safety science.

This paper presents our central argument for adopting the Safety-II approach in the scientific study of cyber crises. This argument forms the foundation of a five-year project: *From 'what went wrong?' to 'what works well?' Using Safety 2 principles to develop new cyber security solutions*, which seeks to answer the question: *What do successful organizations do differently to prevent cyber incidents or mitigate their impact?* The paper articulates the project's core premise, shifting the focus from learning from failures to learning from successes, and outlines a research agenda that integrates insights from safety sciences to advance cyber crisis management strategies.

LEARNING FROM FAILURE

Before discussing the importance of learning from successful operations in detail, it is essential to go back to the foundations of the traditional approach in safety science, which focuses on learning from accidents and failures. This approach, known as the 'Safety-I approach', employs accident analysis to identify causes and contributory factors, aiming to derive lessons and prevent future failures and accidents (Hollnagel, 2013).

The origins of the Safety-I approach can be traced back to industrial systems of the 1970s, which were profoundly influenced by the design principles of Scientific Management (Taylor, 1911). During this period, significant efficiency gains were achieved through time-motion studies, which decomposed tasks and identified optimal methods for their execution. The underlying assumption was that systems could be effectively controlled: processes could be well-designed, procedures comprehensive, people adequately trained, and contingencies fully anticipated by designers (Hollnagel, 2013; Provan et al., 2020). Even as Scientific Management later evolved into Total Quality Management (Hackman & Wageman, 1995), which emphasized continuous improvement over a single "best way" of performing tasks, many of its core assumptions persisted and continued to shape the development of the Safety-I approach.

The Safety-I approach is characterized by three primary principles. The first principle is *decomposability*. Systems can be broken down into individual components, allowing for the identification of root causes when failures occur. In terms of safety, this principle implies that it is possible to pinpoint which specific component within a system caused a malfunction. A second principle is *bimodality*, which assumes that a system is either functioning or malfunctioning, with no intermediate states. The third principle concerns *human error*, which is treated as a potential hazard, akin to mechanical failures or procedural defects. Humans are viewed as system components whose mistakes can contribute to breakdowns (Reason, 1990; 2017). Errors are considered failures within the system, similar to technological or procedural faults (Hollnagel et al., 2015).

Together, these assumptions form the foundation of what is known as the *causality credo*—the belief that adverse outcomes result from identifiable failures with specific causes that can be discovered and addressed. Consequently, the Safety-I approach emphasizes improving safety through meticulous work planning, comprehensive instructions, and targeted training. These assumptions also shape the methodologies used in accident investigations and safety improvements.

There are many well-known examples of accident investigations that illustrate how Safety-I principles have contributed to a better understanding of system failures. One classic example is the 1987 capsizing of the *Herald of Free Enterprise* off the Belgian coast (Rijpma, 2003). The investigation revealed that poor management practices failed to enforce safety procedures, and allowed the ferry to set sail with its bow doors open, causing water to flood the car deck, which lacked proper compartmentalization. Another landmark case is the 1984 gas explosion in Bhopal, India, which remains the world's worst industrial disaster up to date (Weick, 2010). The incident occurred when water entered a tank containing methyl isocyanate during maintenance operations releasing a highly toxic gas cloud, killing thousands and injuring many more. Operators detected gas odors but failed to make sense of the impending disaster because they believed the tank had been shut down six weeks earlier (Weick, 2010). The investigation revealed a series of human and mechanical errors, compounded by poor maintenance and safety practices.

The NASA Challenger and Columbia space shuttle disasters also speak to the imagination of safety scholars (Vaughan, 1997; Boin & Schulman, 2008; Dunbar & Garud, 2009). In both cases, NASA engineers and managers fell victim to the normalization of deviance—a gradual process through which shifting safety norms were increasingly perceived as "acceptable risks." This normalization was reinforced by social and organizational dynamics, such as groupthink and the effects of structural secrecy between organizational levels (Vaughan, 1997). A particularly sobering aspect of these disasters is that NASA failed to internalize critical lessons from the Challenger disaster, a shortcoming driven in part by public and political pressure surrounding space exploration. Ultimately, this failure to learn contributed to the subsequent Columbia disaster, which led to the termination of the space shuttle program.

Similarly, the aviation industry has experienced numerous high-profile accidents that have become pivotal case studies for accident analysis. One notable example is the catastrophic collision between two Boeing 747 jets on the runway in Tenerife, Spain, which still remains the deadliest aviation accident in history (Weick, 1990). Another critical incident is the crash of Air France Flight 447, which plummeted into the Atlantic Ocean after its pilots misinterpreted the autopilot and were caught in a stall (Oliver et al., 2017). Analyses of such accidents have frequently revealed technical system failures compounded by cognitive errors and human factors, highlighting the complex interplay between technology, human interaction, decision-making, and organizational behavior.

In order to understand the logic behind claims about what goes wrong in safety critical operations, it is helpful to engage with the concepts of *Work-as-Imagined* and *Work-as-Done* (Hollnagel et al., 2015). *Work-as-Imagined* represents an idealized view of the formal task environment, outlining what should happen under normal working conditions. It emphasizes the design and prescription of work practices from the 'blunt end' of the organization. This includes carefully defining and formalizing tasks, codifying them into procedures, and implementing monitoring systems to ensure compliance with established norms. *Work-as-Done*, on the other hand, captures what actually happens in practice, how work unfolds over time under complex and often unpredictable conditions. It takes place at the 'sharp end' of the organization and highlights how individuals adapt to irregularities, challenges, and surprises, as well as how they synchronize their activities to achieve their goals. The central challenge lies in guiding these adaptations effectively and ensuring that people are provided with the necessary resources to respond to emerging situations in a productive and safe manner.

In attempts to prevent accidents, safety science professionals implement safety management systems designed to address unacceptable risks by eliminating their causes or strengthening barriers to prevent incidents. As frequently described in the literature, safety management emphasizes standardization and compliance (Hollnagel et al., 2015). Safety management systems and their corresponding safety culture aim to align professionals with organizational safety requirements through centralized control (Provan et al., 2020). These systems prioritize documenting rules and procedures to enhance safety. The underlying assumption is that variability can be managed through checks and balances (Sarvari et al., 2024). Consequently, safety management systems typically combine hazard analysis, controls, monitoring, and standardized performance indicators (Provan et al., 2020). Since these functions are carried out by safety professionals, there is often a strong drive to achieve closure by marking tasks as complete and identifying clear stopping points when processes or actions are deemed unsafe (Provan et al., 2020). This approach frequently reveals a gap between *Work-as-Done* and *Work-as-Imagined*, creating a continuous need to revise procedures and regulations. As a result, today's safety cultures, rooted in the Safety-I approach, place significant emphasis on monitoring and compliance, reinforcing centralized control.

In the Safety-I approach, scholars focus on learning from situations where there was a lack of safety. This accident-centered perspective has dominated the safety science literature, supported by prominent accident theories such as *Normal Accident Theory* by Charles Perrow and *Disaster Incubation Theory* by Barry Turner. Perrow (1984) argues that in certain technological systems, accidents are inevitable due to their interactive complexity and tight coupling. In such systems, failures can interact in unpredictable ways, cascading out of control in ways that cannot be foreseen by operators or designers. Still, the relevance of Perrow's (1984) theory has been questioned by some, as not all industries or organizations operate with systems characterized by such extreme complexity and tight coupling (Hopkins, 1999).

Here *Disaster Incubation Theory* by Turner and Pidgeon (1997) comes in to play, presented in the book *Man-Made Disasters*, that describes how accidents are caused by failures of foresight. Turner argues that, during the period preceding an accident, known as the *incubation period*, warning signals about impending danger are often overlooked or ignored, sometimes for extended periods. According to Turner (1976), this negligence results from socio-psychological and organizational factors, including rigid institutional beliefs, distracting decoy phenomena, neglect of external warnings, difficulties in handling information, and a tendency to downplay potential dangers. Over time, these issues accumulate as latent conditions, which can manifest into a precipitating event, ultimately resulting in disaster.

The two theories present contrasting views on the causes of accidents (Gephart, 1984). Perrow's *Normal Accident Theory* is rooted in technological determinism, positing that accidents are inherent to the structure of complex, tightly coupled systems. Conversely, Turner's *Incubation Theory* emphasizes the unnoticed accumulation of hazards driven by human and organizational factors (Rijpma, 2003). From Perrow's perspective, accidents are unavoidable under specific conditions of technological complexity. In contrast, Turner's perspective suggests that accidents can be prevented through vigilance, awareness, and proactive organizational learning. Turner's view, in particular, opens up opportunities to explore how organizations can prevent accidents and learn from instances of successful safety management.

LEARNING FROM SUCCESS

Over the years, various theories have emerged to explain how industries operating highly complex and hazardous technologies can achieve surprisingly high safety records. This is the breeding ground of High Reliability Theory, Resilience Engineering, and the Safety-II approach (Provan et al., 2020). The core reasoning behind these approaches is that focusing solely on the causes of failures defines safety only by its absence. This perspective leaves little room to assess or understand what constitutes safety in situations where no adverse outcomes have occurred. Consequently, the Safety-I approach struggles to demonstrate whether safety measures are genuinely effective. Moreover, in many complex systems, the principles of functioning are only partially understood, and predictability is limited during both the design and operational phases. As a result, the Safety-I approach has been increasingly challenged by theories that explore how actors achieve reliable performance.

Safety-II theories advocate for a decentralized approach that emphasizes *Work-as-Done*, contrasting with the centralized approach typical of traditional safety management. Within socio-technical systems, variability in performance is considered both normal and necessary. Adjusting performance to changing conditions is essential for maintaining system functionality, highlighting the importance of flexibility and adaptability. As a result, Safety-II places a strong emphasis on how professionals at the 'sharp end' of the organization adapt to variability and dynamically align their actions to balance safety and efficiency (Reason, 1990; Schulman, 1993). Front-line professionals play a critical role in anticipating emerging risks and intervene to prevent incidents, a process known as guided adaptability (Provan et al., 2020).

In practice, this means that front-line workers must often deviate from plans, rules, and procedures to address the emergent nature of complex work environments. Such adaptations frequently lead to incidental deviations from standard procedures and the development of routine workarounds, which are informal adjustments that help organizations remain productive under shifting conditions (Hollnagel, 2013). The *Work-as-Done* perspective is central to key concepts of high-reliability organizations, such as sensitivity to operations, deference to expertise, and commitment to resilience (Schulman, 1993; Roe & Schulman, 2008). These concepts underscore the importance for high reliability organizations to invest in knowledge and quick, critical judgments by its experts.

In this context, safety management efforts within the Safety-II approach rely on sophisticated interactions aimed at translating experiences from *Work-as-Done* into *Work-as-Imagined*. In practice, this means that high-reliability organizations continuously revise and negotiate their formal rule systems, with safety professionals collaboratively updating procedures and establishing tolerances to reflect operational realities (Schulman, 1993). Similarly, Oliver et al. (2017) caution that critical expertise within organizations may erode if the focus is solely on documenting procedures to standardize work and impose stricter limits on professionals. They argue that organizations should allow room for professionals to encounter and learn from errors and violations, rather than

exclusively enforcing compliance. As such, high-reliability organizing depends on a *negotiated order* that consists of an ongoing process of interaction and adjustment among professionals, managers, and authorities to align formal rules with operational practice (Schulman, 1993).

TWO SIDES OF THE SAME COIN

It would be a mistake to disregard the valuable insights gained from the Safety-I approach, including accident analyses, safety culture, and safety management systems. Much knowledge has been gained throughout the years from different types of risk and accident analyses that contributed to a better understanding of system failures. For many organizations, this resulted in safety systems that identify and analyze hazards, implement controls, monitor conformance, and standardize safety cultures (Provan et al., 2020). This is often complemented with inspections and audits resulting in detailed safety reports that contribute to effective quality control systems.

However, it is equally evident that the complexity of today's socio-technical systems surpasses the foundational principles of this approach. In response, Safety-II principles shift the focus to human conduct, innovation, and adaptation as essential strategies for managing variability. Yet, it would be equally misguided to rely solely on an optimistic view of human adaptability. As Oliver et al. (2017) demonstrated in their analysis of the Air France crash, and Dunbar and Garud (2009) in their study of the Columbia Space Shuttle disaster, both organizations and their professionals have limits: adaptation can only go so far. Therefore, understanding the ongoing balancing act between safety and efficiency requires a nuanced perspective, one that acknowledges the value of both Safety-I and Safety-II approaches as two sides of the same coin.

A compelling example to illustrate this point is vividly described in the book "Friendly Fire" by Scott Snook (2000). Snook recounts how a U.S. Air Force squadron operating in Iraq adjusted its squawk code procedures for radar recognition, while an Army Blackhawk helicopter unit remained unaware of these changes. This communication breakdown resulted in a tragic friendly fire incident, where Air Force fighter jets misidentified two U.S. Army helicopters as enemy aircraft. Believing they had engaged two enemy Hinds, the fighter pilots later realized the deadly consequences of their actions. Snook's analysis highlights how 'practical drift', the gradual divergence of actual practices from standard procedures, can lead to catastrophic outcomes. Routine adaptations to procedures, intended to enhance operational efficiency, ultimately contributed to disaster.

Still, also more positive 'Safety-II inspired' analyses of practical drift exist, illustrating that drift is not inherently destructive. Pettersen and Schulman (2019) argue that some degree of practical drift is necessary for organizational adaptation and only results in failure when it is not regulated by expert reliability professionals. According to their analysis, safety-critical issues arise when locally adapted practices fall out of sync with the broader system. In such cases, if reliability professionals are absent or their role is compromised, organizations may experience 'reliability drift', a dangerous misalignment between local adaptations and system-wide safety requirements (Pettersen & Schulman, 2019). Pettersen and Schulman (2019) note that practical drift can also have positive effects. When properly managed, it can foster organizational adaptation to changing environments, enhancing the resilience and situational awareness of safety professionals rather than rendering them vulnerable.

This discussion of practical drift and reliability drift highlights that safety-critical organizations manage multiple risks in parallel. To develop a sophisticated understanding of what contributes to safe operations, it is imperative to integrate both perspectives: while lessons from previous accidents and near-misses are valuable, it is equally important to recognize the efforts made during daily operations to maintain alignment between *work-as-done* and *work-as-imagined*. Together, these perspectives provide a robust conceptual foundation for developing a research agenda that addresses both cyber safety and cyber crises.

TOWARDS A RESEARCH AGENDA

When we return our attention to the subject of cyber crises, we encounter a wealth of research opportunities inspired by the literatures on Safety-I and Safety-II. When organizations seek to advance their cyber security, they often focus on identifying system vulnerabilities, learning from past breaches, or investing in new solutions such as technical tools and awareness campaigns (van Steen et al., 2024). Although attempts to infiltrate systems are frequent, successful breaches are relatively rare. Transposing Safety-II to cyber security, it becomes valuable to investigate situations where organizations face multiple attack vectors but successfully defend against them. For example, a vigilant employee might avert a social engineering attempt, or a technical defense system could neutralize a threat before it causes harm. In line with the transition from a Safety-I to a Safety-II approach, focusing solely on incidents where security measures failed overlooks the numerous instances where proactive actions and adaptive responses effectively prevented or mitigated threats. Despite their importance, these successful defenses remain underexplored in the scientific literature. As a result, valuable insights that could enhance cyber security practices and reveal emerging solutions for strengthening organizational defenses are being overlooked.

Experiences of reliable operations are crucial for understanding how organizations can enhance their cyber security posture. The absence of incidents often reflects effective preventive measures implemented by individuals and teams. However, failing to acknowledge these successes can hinder an organization's ability to defend against future cyber threats. Successful organizations may employ simple yet effective solutions that mitigate multiple risks. These measures might go unnoticed in post-attack analyses because they prevent incidents from occurring in the first place. Additionally, organizations may already have robust preparation and response strategies in place, such as technical defenses and the presence of 'security champions', employees who promote and reinforce team security. Successful organizations might also learn from attackers and develop more sophisticated defense schemes to mitigate attacks, for instance by using a red herring to divert the attackers' attention away from critical system components once inside. Without adopting a Safety-II perspective, these effective practices remain hidden, depriving other organizations of valuable opportunities to learn and improve their defenses. Addressing this gap in cyber security research can significantly strengthen organizational resilience and enhance preparedness against emerging cyber threats.

We see a range of research opportunities that could advance the fields of cyber security and information systems for crisis management. These opportunities align with our central research question: *what do successful organizations do differently to prevent the occurrence of cyber incidents, or reduce their impact?* In the context of cyber crisis management, this project will specifically examine the most effective strategies organizations have in order to prepare for, respond to, and learn from cyber incidents. The underlying assumption is that, while cyber incidents cannot always be prevented, some organizations manage them more successfully than others. This project seeks to identify the factors that enable certain organizations to navigate cyber crises effectively, while others struggle.

The first line of work focuses on describing the security practices implemented by organizations with a high level of cyber security maturity. How are these organizations set up in terms of structures, roles, procedures, security operation centers, and staff training? How does this impact security and detection capabilities? This analysis will provide insight into *Work-as-Done* in cyber security through qualitative field studies in two key organizations: a major Dutch bank, which is part of the national critical infrastructure, and a leading Dutch cyber security company specializing in incident response. Both organizations are continuously exploring new methods to thwart cyber criminals, often employing a range of in-house solutions. Their extensive practical experience and engagement with organizations that have faced cyber-attacks make them highly valuable research partners. Through these collaborations, we will conduct fieldwork with their clients, test our findings in ecologically valid settings, and contribute to spreading success stories across their networks. In addition to our own research agenda, the broader cyber security and crisis management fields may benefit from similar approaches to examine the ways how these organizations deal with performance variability in a context that is not only about safety but especially about providing security, which requires more strategic behavior to changing risks and threats.

A second promising line of research involves exploring how organizations can develop traits of high-reliability organizations to enhance their foresight (Christianson et al., 2011). A key premise of Turner's accident theory is that organizations must prevent warning signals of impending danger from being ignored during a prolonged incubation period. In a cyber security context, this concept could apply to attack vectors that exploit a system vulnerability long before launching an attack, lying dormant until the opportune moment to strike. Vigilance and cognitive awareness are crucial to enhance an organization's ability to monitor its systems and detect weak points in their daily operations. A key aspect to study here is the performance of Security Operation Centers (SOCs) that have an important role in leveraging tactical situational awareness about incidents and possible breaches, through for instance dashboards, collaborative visualizations, and a common operational picture (Mullins et al., 2020; Wolbers & Boersma, 2013). SOCs serve as the frontline in monitoring, detecting, and responding to digital threats. Traditionally, SOCs are structured around a tiered model that prioritizes real-time incident handling: junior analysts filter alerts, while more experienced teams investigate complex threats and proactively hunt for undetected compromises (i.e., threat hunters) (Vielberth et al., 2020). This functional architecture, often supported by specialized roles in malware analysis, digital forensics, and threat intelligence, is primarily designed to detect, contain and mitigate unfolding incidents. However, as cyber crisis governance shifts toward resilience thinking, this model can be expanded to not only respond to incidents, but also to reflect systematically on events that were successfully averted.

A related issue is that high-reliability research often derives its principles from fieldwork on exceptionally safe organizations but tends to generalize findings across vastly different contexts. This research frequently conflates adaptations observed in chaotic settings, such as crisis response (Bigley & Roberts, 2001) or space exploration (Boin & Schulman, 2008), with those in highly routinized and standardized environments, such as nuclear power plants or commercial aviation (Schulman, 1993; Roe & Schulman, 2008; Weick & Roberts, 1993; Oliver et al., 2017). Because the sector's context significantly influences the variability observed in socio-technological systems, different cyber security settings, like critical infrastructure and less safety critical organizations, will

offer distinct opportunities to examine high-reliability traits. It remains uncertain whether we will identify true 'high-reliability organizations' in the cyber security domain or, more likely, encounter organizations striving for reliable performance in complex settings marked by varying levels of vulnerability.

A third promising line of research involves exploring the role of near-misses in cyber security (Bair et al., 2018). While many cyber security solutions are derived from research on incidents, this approach overlooks valuable insights from near-misses and non-incidents–instances where organizations successfully prevented threats through effective cyber security. Focusing solely on incidents means missing opportunities to learn from what organizations are doing right. A best-practice in the field of aviation and medicine is the use of extensive near-miss reporting and analysis systems (Dillon & Tinsley, 2008), which could be translated to a 'cyber safety reporting system' (Bair et al., 2018). This would allow security teams to log events where incidents were averted, whether through early detection, procedural adaptation, or user vigilance. These systems would not only document threat-neutralization, but also protective factors, such as communication, decision-making, and effective tooling (Thoroman et al., 2019). This research should examine successful practices at the individual, team, and organizational level (van Steen et al., 2024). In this way, near-miss analysis could contribute to a shift from Safety-I to Safety-II by capturing how actors flexibly adjust to disturbances, similar to studies of frontline teams (Schakel & Wolbers, 2021). We propose expanding the existing toolbox of cyber security solutions with such instances of successful prevention or mitigation.

CONCLUSION

It is time to lift the veil that keeps cyber crises shrouded in mystery in the field of crisis management. Crises more often have a digital nature that transcends the traditional physical-digital divide. Cyber crises are also known to have cascading effects, especially when critical infrastructures are involved. Still, we should not make the mistake to only focus on accidents and disasters. A focus on learning from failure tends to overlook valuable lessons from successful interventions: we must also learn from what goes right.

The potential for such learning is significant, as it is remarkable how few cyber incidents occur, given the highly digitalized world we live in. To learn from success, we draw parallels from the field of safety science to reflect research possibilities based on the question: *What do successful organizations do differently to prevent cyber incidents or reduce their impact?* Our proposed research agenda fosters insights from both the Safety-I approach focusing on accident analysis, and the Safety-II approach focusing on adaptation and learning. We propose three lines of research that might help in connecting the cyber security and crisis management domain, based on insights from safety science: 1) studying security practices that help avert cyber crises, 2) exploring how organizations can develop traits of high-reliability to enhance their foresight, and 3) exploring the role of near-misses in cyber security. We believe that this research agenda provides a solid ground to explore what insights both domains have to offer the ISCRAM community.

REFERENCES

Bair, J., Bellovin, S. M., Manley, A., Reid, B., & Shostack, A. (2017). That was close: Reward reporting of cybersecurity near misses. *Colorado Technology Law Journal*, *16*(2), 327-364.

Backman, S. (2021). Conceptualizing cyber crises. *Journal of Contingencies and Crisis Management*, *29*(4), 429-438.

Baskerville, R., Spagnoletti, P., & Kim, J. (2014). Incident-centered information security: Managing a strategic balance between prevention and response. *Information & Management*, *51*(1), 138-151.

Bigley, G. A., & Roberts, K. H. (2001). The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal*, *44*(6), 1281-1299.

Boin, A., & Schulman, P. (2008). Assessing NASA's safety culture: the limits and possibilities of high-reliability theory. *Public Administration Review*, *68*(6), 1050-1062.

Boin, A., Hart, P. T., & Kuipers, S. (2018). The crisis approach. *Handbook of Disaster Research*, 23-38.

Christianson, M. K., Sutcliffe, K. M., Miller, M. A., & Iwashyna, T. J. (2011). Becoming a high reliability organization. *Critical Care*, *15*(6), 314-318.

Cisco Talos (2017) *Player 3 Has Entered the Game: Say Hello to 'WannaCry'*. Blog post on 12 May 2017 https://blog.talosintelligence.com/wannacry/, link consulted on 21-02-2025.

CSO Online (2021) The SolarWinds hack timeline: Who knew what, and when? Blog post on 4 June 2021, https://www.csoonline.com/article/570537/the-solarwinds-hack-timeline-who-knew-what-and-when.html, link consulted on 21-02-2025.

Dillon, R. L., & Tinsley, C. H. (2008). How near-misses influence decision making under risk: A missed opportunity for learning. *Management Science*, *54*(8), 1425-1440.

Dunbar, R. L., & Garud, R. (2009). Distributed knowledge and indeterminate meaning: The case of the Columbia shuttle flight. *Organization Studies*, *30*(4), 397-421.

Gephart Jr, R. P. (1984). Making sense of organizationally based environmental disasters. *Journal of Management*, *10*(2), 205-225.

Hackman, J. R., & Wageman, R. (1995). Total quality management: Empirical, conceptual, and practical issues. *Administrative Science Quarterly*, 309-342.

Hollnagel, E. (Ed.). (2013). *Resilience engineering in practice: A guidebook*. Ashgate Publishing, Ltd.

Hollnagel, E., Wears R.L, & Braithwaite, J. (2015). "*From Safety-I to Safety-II: A White Paper*."

Hopkins, A. (1999). The Limits of Normal Accident Theory. *Safety Science*, *32*(2), pp. 93–102.

Farjoun, M. (2005). Organizational Learning and Action in the Midst of safety drift: Revisiting the Space Shuttle Program's recent history. In Starbuc & Farjoun (eds.). *Organization at the limit: Lessons from the Columbia disaster*, 60-80.

Kaspersky Daily (2017) *New Petya / NotPetya / ExPetr ransomware outbreak*. Blog post on 27 June 2017 https://www.kaspersky.com/blog/new-ransomware-epidemics/17314/, link consulted on 21-02-2025.

Mullins, R., Nargi, B., & Fouse, A. (2020). Understanding and enabling tactical situational awareness in a security operations center. In *Advances in Human Factors in Cybersecurity: AHFE 2020 Virtual Conference on Human Factors in Cybersecurity, July 16–20, 2020, USA* (pp. 75-82). Springer International Publishing.

Oliver, N., Calvard, T., & Potočnik, K. (2017). Cognition, technology, and organizational limits: Lessons from the Air France 447 disaster. *Organization Science*, *28*(4), 729-743.

Perrow, C. (1984). *Normal accidents: Living with high risk technologies.* New York, NY: Basic Books.

Pettersen, K. A., & Schulman, P. R. (2019). Drift, adaptation, resilience and reliability: Toward an empirical clarification. *Safety Science*, *117*, 460-468.

Prevezianou, M. F. (2021). Beyond ones and zeros: Conceptualizing cyber crises. *Risk, Hazards & Crisis in Public Policy*, *12*(1), 51-72.

Provan, D. J., Woods, D. D., Dekker, S. W., & Rae, A. J. (2020). Safety II professionals: How resilience engineering can transform safety practice. *Reliability Engineering & System Safety*, *195*, 106740.

Reason, J. T. (1990). *Human error.* Cambridge, UK: Cambridge University Press.

Reason, J. (2017). *The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries*. Routledge, London

Reuters (2021) *Cyber-attack shuts down U.S. fuel pipeline 'jugular,' Biden briefed*. News article on 8 May 2021 https://www.reuters.com/technology/colonial-pipeline-halts-all-pipeline-operations-after-cybersecurity-attack-2021-05-08/, link consulted on 21-02-2025.

Rijpma, J. A. (2003). From deadlock to dead end: The normal accidents-high reliability debate revisited. *Journal of Contingencies and Crisis Management*, *11*(1), 37-45.

Roe, E., & Schulman, P. S. (2008). *High reliability management: Operating on the edge*. Stanford, CA: Stanford Business Books.

Sarvari, H., Edwards, D. J., Rillie, I., & Posillico, J. J. (2024). Building a safer future: Analysis of studies on safety I and safety II in the construction industry. *Safety Science*, *178*, 106621.

Schakel, J. K., & Wolbers, J. (2021). To the edge and beyond: How fast-response organizations adapt in rapidly changing crisis situations. *Human Relations*, *74*(3), 405-436.

Schulman, P. R. (1993). The negotiated order of organizational reliability. *Administration & Society*, *25*(3), 353-372.

Snook, S. A. (2000). *Friendly fire: The accidental shootdown of US Black Hawks over northern Iraq.* New Haven, NJ: Princeton University Press.

Taylor, F. W. (1911). Principles and methods of scientific management. *Journal of Accountancy*, *12*(3), 3.

Thoroman, B., Goode, N., Salmon, P., & Wooley, M. (2019). What went right? An analysis of the protective factors in aviation near misses. *Ergonomics*, *62*(2), 192-203.

Turner, B. A. (1976). The organizational and interorganizational development of disasters. *Administrative Science Quarterly*, 378-397.

Turner, B. A., & Pidgeon, N. F. (1997). *Man-made disasters.* Oxford: Butterworth-Heinemann.

Vaughan, D. (1997). *The Challenger launch decision: Risky technology, culture, and deviance at NASA.* Chicago, IL: University of Chicago Press.

van Steen, T., Del-Real, C., & van den Berg, B. (2024). What Works Well? A Safety-II Approach to Cybersecurity. In *International Conference on Human-Computer Interaction* (pp. 250-262). Cham: Springer Nature Switzerland.

Vielberth, M., Bohm, F., Fichtinger, I., & Pernul, G. (2020). Security Operations Center: A Systematic Study and Open Challenges. *IEEE Access*, 8, 227756–227779.

Weick, K. E. (1990). The vulnerable system: An analysis of the Tenerife air disaster. *Journal of Management, 16*(3), 571-593.

Weick, K. E., & Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly*, *38*(3), 357-381.

Weick, K. E. (2010). Reflections on enacted sensemaking in the Bhopal disaster. *Journal of Management Studies*, *47*(3), 537-550.

Wired (2018) *The Untold Story of NotPetya, the Most Devastating Cyberattack in History*. Blog post on 22 August 2018 https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/, link consulted on 21-02-2025.

Wolbers, J., & Boersma, K. (2013). The common operational picture as collective sensemaking. *Journal of Contingencies and Crisis management*, *21*(4), 186-199.

Wolbers, J., Kuipers, S., & Boin, A. (2021). A systematic review of 20 years of crisis and disaster research: Trends and progress. *Risk, Hazards & Crisis in Public Policy*, *12*(4), 374-392.