MIS Quarterly

RESEARCH ARTICLE

HIGH RELIABILITY IN DIGITAL ORGANIZING: MINDLESSNESS, THE FRAME PROBLEM, AND DIGITAL OPERATIONS[1]

Antti Salovaara
Department of Computer Science, University of Helsinki, FI-00014 Helsinki, FINLAND, and
Department of Information and Service Economy, School of Business, Aalto University, FI-00076 Aalto, FINLAND
{antti.salovaara@alumni.helsinki.fi}

Kalle Lyytinen
The Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106-7235, U.S.A., and
Department of Information and Service Economy, School of Business, Aalto University, FI-00076 Aalto, FINLAND
{kalle@case.edu}

Esko Penttinen
Department of Information and Service Economy, School of Business, Aalto University, FI-00076 Aalto, FINLAND
{esko.penttinen@aalto.fi}

Organizations facing high risks and operating in purely digital domains, such as computer security and many financial services, must meet two, contradictory goals: they need to identify digital threats at scale and speed while also avoiding errors that result from automated processing. Research on high-reliability organizations has identified multiple challenges in reaching these goals simultaneously, because automation often renders organizations' operations "mindless" and unable to cope gracefully with changing, complex situations characteristic of high-risk domains. In digital operations, a special challenge arises from the "frame problem" connected with the inability of algorithms to adapt to environments not identified within their developers' initial cognitive frames. An exploratory, theory-generating case study was conducted within a computer security company (F-Secure) to examine how organizations acting in digital domains may achieve high reliability by mitigating the frame problem. This article examines digital organizing of the epistemic and pragmatic features of operations, along with arrangements of these features that respond to the frame problem. Collective mindfulness is identified as emerging in such a sociotechnical setting via a carefully layered, systemic constellation of (human) mindful and (digital) mindless operations while the organization's core operations remain digital and algorithmic. The findings point to heretofore unexplored reliability challenges associated with digital organizing, alongside several relevant ways to overcome and/or mitigate them.

Keywords: High-reliability organization (HRO), digital operations, digital HRO, mindfulness, malware protection, frame problem

[1]Jonathan Wareham was the accepting senior editor for this paper. Panos Constantinides served as the associate editor.

The appendices for this paper are located in the "Online Supplements" section of *MIS Quarterly*'s website (https://misq.org).

Introduction

High-reliability organizations (HROs) are exceptional socio-technical systems that can operate in a nearly error-free manner in environments wherein most other organizations constantly run the risk of major accidents (Bigley and Roberts 2001; Weick and Roberts 1993; Weick and Sutcliffe 2001). Thus far, HRO studies have focused on time-critical, tightly coupled command-and-control operations in complex settings such as aircraft carriers (Rochlin et al. 1987), air-traffic control, and nuclear power plants (e.g., Bourrier 1996; Marcus 1995). Interest in HROs utilizing information technology (IT) has grown recently (Butler and Gray 2006; Dernbecher and Beck 2017), with research delving into IT's role in construction (Carlo et al. 2012) and military operations (Grabowski and Roberts 1999). While some of these domains feature *digital operations* (i.e., autonomous work by IT handling digital inputs and outputs), studies of IT-related operations have addressed primarily operations wherein humans act as independent tool users and decision makers. Moreover, we know of only a handful of studies examining *digital HROs*: organizations conducting "pure" digital operations. Among these organizations, whose core operations are IT-based without humans being active in the operational decision-making loop, are entities working with Internet payments, high-frequency trading, and blockchain-based cryptocurrencies (Van Den Eede et al. 2006).

The quintessential digital HRO is an organization that builds services and solutions to protect IT systems from threats of malicious software, or *computer malware*. High reliability is rendered vital by malware's serious threats to fault-free digital operations. It deploys unexpected, heterogeneous "attack mechanisms" to spy, destroy, or block access to data (e.g., for ransom purposes), hijack computer resources, open "backdoors" for other malware, and otherwise act at a distance. Second, being of digital nature, malware is notoriously difficult to identify. Examining the conditions under which it is operating, it can modify itself and evade detection by "hiding" inside other digital content. Also, as a digital object, it can spread with enormous scale and speed. For example, denial-of-service attacks can hijack tens of thousands of computers, mobilizing and then releasing them suddenly. The WannaCry attack of May 2017 illustrated this poignantly by infiltrating more than 200,000 computers, in 150 countries, among them machines of 61 trusts and hospitals in the United Kingdom's National Health Service, then encrypting stored data and demanding payment for decryption.[2] Because virtually all organizations in the industrialized world depend on IT systems and digital operations, malware's speed, volume, complexity, and geographical dispersion forces protection organizations to rely extensively on algorithms and digital operations in their work to protect against the threats it poses.

Established HRO theories have not yet responded to digital operations' recent upsurge. Why is it important to develop theory that considers the effects of IT-based core operations in high-reliability domains? While research has shown traditional HROs to achieve high reliability through a cognitive orientation called *collective mindfulness*, defined as heedfully anticipating surprises and prioritizing safety over efficiency (Weick and Roberts 1993; Weick et al. 1999), the principal challenge digital HROs face is that IT-based operations are antithetical to such forms of mindfulness and, hence, arguably increase the risk of failure (e.g., Butler and Gray 2006; Grabowski and Roberts 1999; Valorinta 2009). The central issue for digital HROs is how to overcome this conundrum and achieve high reliability: the organization must continue relying on algorithmic, automated operations while simultaneously being mindful—anticipating surprises and prioritizing safety.

This paper addresses the unique role of algorithms—unambiguous and serially executable instruction sequences (Lewis and Papadimitriou 1998)—in shaping digital HROs' operations. We pay particular attention to algorithms as a key element underlying threats to digital high reliability. Scholars of artificial intelligence (AI) hold that mindfulness in algorithmic processing is impossible because algorithms are constrained by the "frame problem": the theoretical impossibility of preparing computational agents for environment states for which they lack associated rules (Dennett 1984; McCarthy and Hayes 1969). In simple terms, algorithms cannot reliably detect and act on events they have not been designed to handle. The frame problem acts counter to mindfulness, which assumes anticipation of surprises in the environment. We argue, however, that such adverse effects with regard to high reliability are not insurmountable: we theorize on high reliability in digital organizing as a broader socio-technical design problem that can be solved through balance between digital and human-based operations.

Anti-malware companies' significant success in protecting against malware suggests that the frame problem can be circumvented and high reliability achieved in organizations that perform fully digital operations. This prompts us to consider how such organizations effectively organize and successfully deal with the paradoxical situation wherein IT-based digital operations are deployed to promote high reliability. Here, we will discuss this conundrum through an exploratory, theory-generating case study addressing two related research questions.

[2] "Ransomware Cyber-Attack Threat Escalating–Europol," BBC News, Technology, May 14, 2017 (http://www.bbc.com/news/technology-39913630).

RQ1: Given that the frame problem imposes limits to mindfulness, does this limitation shape digital organizing in high-reliability operations such as malware protection?

RQ2: If so, how does such a digital HRO combat the frame problem and organize its operations as a collectively mindful sociotechnical system?

We examine operation conditions within which highly successful malware-protection company F-Secure (http://www.f-secure.com/) achieves high reliability and how it organizes its operations such that the frame problem is overcome. Concerning RQ1, we demonstrate that the frame problem indeed informs F-Secure's efforts to address its concerns about high reliability, and we then address RQ2 by examining how this case company systematically organizes its human-based and digital operations. We present a layered sociotechnical structure wherein mindful human-based context-sensitive processing monitors and improves the mindless algorithmic processing at its core while still actively anticipating new, unexpected threats. We find that two sets of operations—epistemic and pragmatic—mutually reinforce the human and digital operations and produce collective mindfulness as an emergent property of the sociotechnical system.

A Review of the Theory

Two complementary approaches have informed organizational analyses of high-reliability organizing thus far. The first, normal accident theory (NAT), builds on Perrow's (1984) seminal research on the causes of the near failure of the Three Mile Island nuclear power plant in 1979. Perrow laid out key arguments pertaining to high reliability for organizations acting on complex systems, from infrastructural elements (power plants, dams, power grids, etc.) to chemical factories and military operations. He posited that, irrespective of effective management, good training, and nearly error-free individual operations, a system characterized by tightly coupled components and complex interactions inevitably experiences a "normal" accident: when small concurrent failures in multiple components coincide in unexpected ways, there will inevitably be escalation to the entire system's breakdown (Oliver et al. 2017; Reason 2000). High reliability is reached at the system level only if the coupled operations remain within certain relatively narrow variance limits. Otherwise, these systems are less able to recover. For example, extreme weather may push operations beyond their limits, with the ensuing events culminating in system failure (Perrow 1984). To reduce likelihood of failure, Perrow recommended that, whenever possible, complex systems be organized with a loose coupling that reduces interactive complexity.

The second approach, the theory of high-reliability organizations and organizing, complements NAT. Its goal is to overcome NAT's inherent pessimism (rooted in the systemic study of operations and their couplings) by drawing inspiration instead from the observation that grand system failures are relatively rare and that many organizations whose operations exhibit high interaction complexity can actually operate in a nearly error-free manner. To account for this observation, HRO scholars have sought to complement NAT's systemic analyses with research into organization-internal cognition and its arrangement for higher reliability (Sagan 1993). This work has largely involved inductive ethnographic study of near-error-free organizing in typical complex interactive settings: nuclear power plants, airplane cockpits, etc. (e.g., Bigley and Roberts 2001; Roberts 1990; Roberts et al. 1994; Vogus and Welbourne 2003; Waller and Roberts 2003; Weick and Roberts 1993; Weick and Sutcliffe 2001). These studies delineate between HROs and traditional efficiency-driven organizations by dominant cognitive orientation: HROs focus on and learn from operations' failures rather than successes and prioritize reliability over efficiency (Weick et al. 1999). The unique *cognitive* mindset that guides HROs to handle normal, inevitable threats and accidents is an orientation manifested in five system-level characteristics:

- *Preoccupation with failure*: HROs maintain emphasis on the possibility of normal failures. In treating near failures as presenting important lessons, they focus on making sense of events and situations that either are rare or have never actually arisen.

- *Reluctance to simplify interpretations:* HROs maintain healthy skepticism about obvious interpretations in their sense-making, to minimize "blind spots" when examining failures and their pathways. In contrast, they promote actors' heedful attention to minute anomalies and early warning signals.

- *Under-specification of structures*: HROs form multifunctional "garbage can" processes (Cohen et al. 1972) involving a broad spectrum of experts who can fluidly assume responsibility for alternative facets of problem-solving as necessary for generating flexible, dynamic responses to a failure's escalation.

- *Commitment to resilience*: HROs cope with surprises "in the moment" by swiftly generating untried approaches, persisting in "bricolage" (Bigley and Roberts 2001; Lévi-Strauss 1966), and learning from prior experience that often involved only a few samples.

Table 1. Characteristics of Digital Operations

Characteristic	Definition	Implications for Malware Protection
Exactness	Digital operations share the same homogeneous representation: binaries expressed with discrete strings of 1s and 0s (Tilson et al. 2010; Yoo et al. 2010). Digital systems have exact states, and their processes are deterministic.	Even small, one-bit variations can change a digital system's behavior or cause its state to denote dramatically different things. Exactness enables replicable intrusion: if the given software is on several computers, the same intrusion works for them all.
Editability	Digitally represented operations can be modified easily: their physical instantiations in digital memory chips offer easy manipulation (Kallinikos et al. 2013).	Malware can modify itself and the surrounding digital content while retaining the original operations' essential characteristics. Hence, the modified objects are seen *qua* the originals while carrying parasitic content and promoting alternative behaviors.
Programmability	Digital operations rely on von Neumann architecture: the same medium can be used to store the data and the rules (algorithm). Because algorithms themselves are data, digital objects are inherently capable of modifying their own state and reprogramming their behaviors and other objects' (Kallinikos et al. 2013; Yoo et al. 2010).	Malware can self-mutate and appear in numerous guises. It also may change the behavior of other digital objects.
Transferability	Digital system states can be transferred and replicated at low effort and cost. Transfer from one system to another can be performed such that the digital operations remain unchanged (Kallinikos 2012).	Malware infections can spread at speed globally. Once a malevolent agent finds a vulnerability in one software system, it can be quickly distributed and the vulnerability exploited globally.

- *Sensitivity to operations*: HROs develop a holistic view of their operations and environment (Vogus and Welbourne 2003), inviting actors to assess potential interdependencies among events constantly, for more prudent (effective) decision on consequent actions (Roberts and Rousseau 1989).

While both schools strive to understand how organizations can behave reliably in high-risk domains, they differ in their view of accidents' sources and the potential for mitigation. While NAT presents conditions for a system collapse that springs from generic and systemic characteristics of complex interactive systems (how their operations are coupled) (Perrow 2010), HRO scholars do not approach failures in nearly deterministic terms as being produced in an inevitable fashion by intractable, latent systemic interactions. The HRO approach has an agentic orientation, addressing every near failure or accident in such complex settings as having reasons that can be identified, interpreted, accounted for, and potentially anticipated by local agents (Weick et al. 1999). Malware protection is just such a context: identifiable sources for failures truly exist—malware and its operations can, in principle, always be analyzed and detected, if the agent is "smart" enough to identify the rules for their detection from the digital data. Hence, our analysis of digital organizing draws primarily from HRO theory but we draw on NAT in focusing on the unique nature of digital systems and related operations.

Digital Operations and the Frame Problem

Malware protection is naturally amenable to HRO-based analysis, in calling for anticipation of surprises and for a preoccupation with failure. Yet the operations' digital nature demands additional clarification that constitute boundary conditions for robust theory-building like in the case of NAT. Table 1 synthesizes four characteristics unique to digital operations—exactness, transferability, editability, and programmability—as identified in recent work in innovation theory, information systems, and institutional analyses of technology (Kallinikos et al. 2013; Tilson et al 2010; Yoo et al. 2010). We posit that these characteristics jointly can explain why digital operations are vulnerable to threats in a different way than human-based operations.

The first characteristic involves digital content's and processes' *exactness*. Human-based control rests on approximate measurements of operations. These neither require nor apply absolute precision, and minor variations in a system's inputs seldom cause dramatic changes in outputs. Digital opera-

tions' exactness, in contrast, renders them deterministic in their consequences and therefore brittle. Occasionally, even the tiniest change at a particular point in digital representation generates a fundamentally different behavior within a digital system and its outputs. A flipped sign bit in a digitally encoded number, for example, inverts the value and possibly the related operation's behavior too.[3]

Next, digital operations are *editable* and *programmable*. Modifications entail only changes in representations encoded in electromagnetic memory arrays, making them malleable, continuously expandable, and open to new interpretation and accommodation. Finally, digital representations have nearly limitless *transferability*: digital operations and data can be transferred and replicated at low cost and with vast reach.

Another feature that distinguishes digital from human-based operations is the *frame problem* following from the former's algorithmic foundation. This concept has its origins in computer science, primarily AI research (McCarthy and Hayes 1969; see also Dennett 1984), which from its inception has had an aspiration to emulate human intelligence by expanding the boundaries of rule-based intelligence. The frame problem formulation highlights the inherent limitations to computational agents' competent action in a dynamic environment for which they lack *a priori* articulated rules (McCarthy and Hayes 1969).[4] In colloquial terms, these agents cannot think "outside the box" or beyond the rules. The frame problem introduces a challenge to digital operations' reliability in ambiguous and complex settings such as detecting potential malware threats. An algorithm's capacity to identify and act upon events and states has a fixed upper bound dictated by the current frame manifested in its rules. In this sense, digital operations are mindless and play by the book: they are fixated on "a single perspective without awareness that things could be otherwise" (Weick et al. 1999, p. 38). Although algorithms' sophistication and frame can, in theory, be expanded *ad infinitum* by their augmentation with deeper and more expansive knowledge of the world (Dennett 1984) and with related categories and rules, the boundaries imposed by the frame problem remain, however large the frame and the rule set are made over time.

The rising power of machine learning[5] (e.g., Brynjolfsson and McAfee 2014) does not solve the frame problem either. The widely discussed deep learning techniques are indeed successful in generalizing solution patterns for limited tasks (such as interpreting X-ray images), solving "closed-world" problems (e.g., playing the game Go), and generating new outputs for restricted tasks (as in chip design or game visualizations) by learning continuously from data (presented as experience). In all such cases, however, the algorithms need large bodies of input data to learn the solution: the problem's scope remains constrained by the goal functions and the data available (e.g., Domingos 2012). Indeed, after decades of effort, research remains unable to tackle the challenge of "strong AI": developing algorithms that independently identify novel concepts and infer laws from unexpected situations for governing their future behaviors—a situation crucially needed in systems with "collective mindfulness" that demand reluctance to simplify interpretations. Unless strong AI somehow becomes reality, computers' representation capacity is going to remain limited to what programmers supply in one form or another; they are ultimately bounded by the representational frame[6] around what they can do (Dennett 1984; Shanahan 2016). As Wittgenstein (1922) would say, the limits of an algorithm's operations on its symbols are also the limits of its world.

In this regard, malware protection presents a particularly hard case in constantly calling for overcoming the frame problem in resolving of security threats. Malware developers make smart use of the aforementioned four characteristics of digital operations by purposefully programming their software to alter and replicate itself continuously and remain unrecognizable by protection algorithms. Hence, for the HRO to remain reliable and effective, there must be constant strivings for the malware-protection software to "go beyond" its existing frame.

An effect similar to the frame problem, although milder, is visible in human operations. Routinized human-based operations, often called standard operating procedures (see Swanson and Ramiller 2004; Vogus and Welbourne 2003), are akin to algorithm-following to reduce variance and increase reliability. Something similar to the frame problem limits standard operating procedures and increases the threat

[3]Several classic cases illustrate how such small errors can, for instance, cause rockets to follow the wrong trajectory and crash.

[4]So-called frame axioms—specifications of behaviors generally feasible in any environment—have been explored to overcome this limitation of algorithms, but agents ultimately get trapped in infinite regression of irrelevant conclusions because they start off incapable of singling out relevant conclusions from among mostly irrelevant ones (Shanahan 2016).

[5]These include also second-generation neural networks, genetic algorithms, swarm intelligence, and deep learning.

[6]To alleviate this problem, scholars have suggested the embodied agent: a system that discards internal representations and proverbially "use[s] the world as its own model" (Brooks 1995, p. 139). Some argue that such systems can cope gracefully with complex worlds (Dreyfus 1972, 1979; Winograd and Flores 1986); however, such approaches are not relevant for our purposes, because here all materials and operations (the world) are already digital and disembodied by nature.

of their mindlessness. Individuals and groups may fixate on a single perspective (Weick et al. 1999), fail to anticipate unexpected events (Oliver et al. 2017; Perrow 1984, pp. 318-321), grow overconfident (Kahneman 2011), become trapped in groupthink (Janis 1972), and dismiss/resist new and conflicting information (Dane 2010; Ericsson 2009; Fiske and Taylor 1991; Kahneman 2011). The frame problem's human equivalent is termed the *entrenchment problem* (Dane 2010). Although showing similarities, the two differ in one important respect: digital operations are unavoidably deterministic and bounded by the frame problem, while humans are capable, at least in principle, of context-sensitive processing and often can manifest mindful behavior even in standardized practices if the right cue or person is present. The participants have potential to anticipate, build alternative cognitive models for situations, and engage in context-shifting, bricolage, and improvisation. Also, such mindfulness—reluctance to simplify interpretations—in human operations can be fostered via exposure to increasingly dynamic environments and tasks (Dane 2010). Thus humans can be made to learn to become mindful. However, there is an upper limit to the extent of human-based operations' mindfulness in high-risk environments. In comparison to digital operations' scale and speed, human memory capacity and processing speed are highly limited. Therefore, the overarching challenge for greater reliability of digital operations is in organizing the overall sociotechnical system, composed of humans and digital operations and involving both mindful and mindless forms, in such a way that the combination reaches higher levels of collective mindfulness.

Mindfulness and Mindlessness in IT-Based Operations

Per our discussion above, digital HROs must constantly achieve balance of the fast and scalable but frame-problem-ridden (i.e., mindless) digital operations with the slow and local but mindful human-based operations. Threats to reliability in IT-based operations are well known: IT use renders organizations vulnerable to simplification, playing by the book, and shallow awareness (e.g., Butler and Gray 2006; Grabowski and Roberts 1999; Valorinta 2009). Somewhat counterintuitively, IT systems' high reliability due to their algorithmic nature may decrease reliability at the organization level, because IT use promotes routinization and mindlessness. Thus, *unreliable* IT systems may do more to prompt mindful behaviors, by reminding the organization's members to remain attentive to system outcomes (Butler and Gray 2006).

The critical concern of finding a balance between mindfulness and mindlessness has been prominent in literature on automation in high-risk environments (Butler and Gray 2006; Farjoun 2010; Sonenshein 2016). Consensus is emerging on the necessity of finding positive complementarities between the two modes of operation (Bigley and Roberts 2001; Cameron 1986; Carlo et al. 2012; Farjoun 2010; Levinthal and Rerup 2006; Louis and Sutton 1991; Rochlin 1993; Whetten and Cameron 1994). Indeed, Weick et al.'s (1999, p. 53) initial conceptualization of mindfulness strongly hints at the need for constantly balancing requirements for considering details versus the whole and for playing by the book versus beyond it. In such complementary configurations, IT acts as the glue allowing mindfulness-promoting cooperation across fluid organizational structures (Grabowski and Roberts 1999). Ramiller and Swanson (2009) suggest similarly that critically focused IT-use evaluations help organizations become aware of undesired mindless effects of IT's use. Finally, Carlo et al. (2004, 2012) discuss applications of digital three-dimensional models and related tools in complex construction projects as having to support both mindless and mindful operations. Balance is not static but a process, since mindful and mindless uses emerge contextually in alternating fashion and never simultaneously. Yet few empirical operations-level analyses exist on how mindfulness and mindlessness get effectively balanced within digital operations. Thatcher et al. (2018), for example, focus solely on individual-level mindfulness. Dernbecher and Beck's (2017) review on mindfulness in information systems research, in turn, describes only a few studies bringing together mindfulness, high reliability, failure-prone contexts, and emphasis on digital operations. Most studies of mindfulness highlighted other matters: they compared individual-level and collective mindfulness related to IT or discussed mindfulness in relation to information systems development or IT-based innovation. Only around 10 articles (approximately 20% of Dernbecher and Beck's data corpus) addressed organization-level reliability and operations. In most cases, the research context involved not high risks and reliability but relationships of mindfulness and decision support (Van de Walle and Turoff 2008), business risks (Simha and Kishore 2011), product development (Merminod et al. 2008), requirement elicitation (Sammon et al. 2014), or offshoring control (Pu and Kishore 2006). On the other hand, although Weick and Sutcliffe (2006) discuss high risks and organizational attention in detail, they do not link these to IT-based operations or the challenge posed by the frame problem. Finally, while Van Den Eede et al. (2006) apply HRO principles to financial services and associated IT operations, even their extensive analyses do not take into account the challenge inherent to algorithmic operations.

One reason for the scarcity of operation-level analyses for mindfulness–mindlessness balance in digital operations may lie in the HRO theory's cognitive orientation. Weick et al.'s

Table 2. Three Dimensions for Analyzing High-Reliability Digital Operations

Feature	Feature Type	
Nature of operation	*Human-based*: approximate, error-prone, limited by memory capacity and processing speed, of varying precision, context-sensitive	*Digital*: exact, transferable, editable and programmable via expression of binary data
Nature of cognition	*Mindful*: heedful, with anticipation of surprises and prioritization of safety in operations, unconstrained by the frame problem	*Mindless*: constrained by the frame problem via algorithm-use or reliance on highly structured routines
Purpose	*Epistemic*: interpreting and analyzing information	*Pragmatic*: performing decision-making and acting

(1999) framework of five HRO characteristics depicts how an organization is expected to "think" as it orients itself effectively toward threats. At the same time, such a focus leaves largely unaddressed what organizations actually do or how their thinking and doing interact. To better account for the nature of such operations with regard to digital organizing, we adopt Kirsh and Maglio's (1994) classification of actions in distributed cognitive systems. We deem their taxonomy relevant for considering systemic connections between digital and human operations in creating an effective sociotechnical design, because nearly all digital operations in an organizational setting involve multiple agents (human and algorithm-based), with distinctive stakes and skills, where operations constitute a distributed cognitive system (Boland et al. 1994).[7] Kirsh and Maglio's classification highlights the presence of two types of operations in cognitive systems: (1) *epistemic operations* gathering information and interpreting it for subsequent decision making and (2) *pragmatic operations* making decisions and acting on them. Such classification meshes well with environments of high-reliability organizing, wherein learning about a complex, dynamic environment and operating or acting on it take place simultaneously. In settings of this sort, human actors anticipate the operations' consequences and check the validity of such inferences through environmental information by building, consulting, and reviewing underlying cognitive models.

A Framework for Analyzing HROs

We can now summarize our framework for analyzing digital HROs and their operations. It brings together the three dimensions presented above: the nature of the operations (human-based or digital), the sensitivity to and anticipation of surprises (mindful or mindless), and the purpose of the operations making up the organization's activity flow (epistemic or pragmatic). The first two of these are closely related, because mindlessness is intimately bound up with digital operations via the frame problem. The third dimension, although relatively independent, allows one to analyze the systemic relationships between mindful and mindless operations and to identify their distinct, contradictory goals. With this dimension, largely ignored in past studies of HROs, we expand the analysis into *sociotechnical* inquiry.

Overall, the proposed framework directs attention to the complementary systemic relationships between the two shaded regions in Table 2. We turn our attention next to how these relationships are created and maintained as a critical condition for high reliability in digital operations. The regions identify the two distinct sets of opposing operations in our study context: human operations and digital ones. The area on the left in the table refers to mindful human-based operations with capacity for context-sensitive processing, imagination, and bricolage, but also with potential to fail miserably through the entrenchment problem. The region on the right covers digital operations characterized by algorithmic processing and mindlessness stemming from the frame problem. This tensional framing invites investigation of how the mindful and mindless operations and their organized interactions may display complementarity and produce collective mindfulness as an emergent property of a sociotechnical system.

The Case Study: Reliable Malware Detection

We are now ready to address our two research questions. To examine whether (and, if so, how much) the frame problem

[7]Operations have been extensively studied in organization scholarship, especially in relation to human factors and naturalistic decision making (e.g., Malone and Crowston 1994; Roth et al. 2006; Vicente and Rasmussen 1990). Division into epistemic and pragmatic operations is particularly suitable for studies drawing on research into distributed cognition (Hutchins 1995), because they share theoretical underpinnings with the work of Kirsh and Maglio.

shapes high-reliability operations in digital settings (RQ1) and, if the answer is "yes," how such a digital HRO can be sociotechnically organized to address the frame problem (RQ2), we conducted an exploratory case study by focusing our observation and analysis of "a phenomenon previously inaccessible to social science inquiry" (Yin 2009, p. 48). The study setting of a leading malware-protection company extended to both company-internal operations and the external ones within a broader software ecosystem of vendors (including other malware-protection companies), customers (individuals and organizations alike), and malware-creators. Befitting the study's exploratory nature, the method was inductive; grounded in contextual, qualitative data; and comparative (Eisenhardt 1989). As is deemed normally fitting for "exploratory revelatory" case studies (Eisenhardt 1989; Yin 2009), we constantly iterated between data-collection and analysis, with theoretical triangulation and exclusion of alternative explanations.

The Study Setting

F-Secure is a malware-protection company headquartered in Helsinki.[8] Globally, the company has approximately 950 employees, in 20 offices, making it Europe's largest malware-protection company. F-Secure has offered security services for 25 years and become widely acknowledged as one of the industry's technology leaders. Security-software evaluator AV-Test rated F-Secure's protection the best on the market for five years in the 2010s (see http://www.av-test.org/). In addition to providing security solutions that can be installed for desktop computers, tablets, and smartphones across multiple operating systems, the company provides security services offered by Internet service providers. For both, it produces and operates its own software-security platform, developed and maintained internally.

In a typical day, F-Secure identifies and acts on more than a million software threats (e.g., computer visits to unsafe websites or retrievals of malicious content). To handle this volume, it relies on digital operations and algorithms in its detection and protection processes. Despite the demand for quick response, at scale, which creates impetus to adopt efficiency logic, F-Secure maintains an orientation typical of HROs in that its attention is mainly on reliability and avoidance of failures. In consequence, as we will show, F-Secure organizes its operations in a manner consistent with HROs' traits.

[8]With F-Secure's permission, we have not anonymized the company (see also Salovaara et al. 2015). With so few software security companies on the global stage, it would have been easy to determine the case company's identity anyway.

The company's software operates on three levels (see Appendix A). First, reputation-inspection examines the Internet addresses that the computer attempts to visit, and it blocks access to compromised ones. Next, data arriving from any Internet site will undergo malware-sample-based detection. Finally, behavior-monitoring continuously inspects the computer's internal processes for signs of malicious activity. The logic of these three "phases" of operation draws on algorithmic instructions held in a rule-based "detection engine" running on end-user computers. It must be updated regularly as new threats are identified. Much of the intelligence for this is obtained from "upstream" reports that the end-user installations send to F-Secure's servers. Analysts examine the reported malware threats, using a range of software tools that help to identify them accurately and reveal their operation logic. As necessary, the analysts modify the rules within the detection engine. End users' software typically downloads a newly revised version of the rules several times each day.

F-Secure's operations are organized into a two-unit system aimed at constantly improving the detection rules and the engine's operation (see Figure 1). The *reactive response unit* operates in continuous shifts to gather reports on recent threats, communicate with customers, and make time-critical corrections to detection rules when conditions demand immediate response. Operations within shifts allow for multiple levels of escalation, depending on the threat's urgency and scope. Analysts on the "front-end" team interact with end users encountering "trouble" with F-Secure products who either cannot access content blocked by protection mechanisms or suspect that a malware infection is present. If this team's analysts cannot solve the problem within 1 to 2 hours, they pass it on to a higher-tier response team for intensive examination. That team can escalate the problem further, and it may finally reach the unit's "back-end" expert team. F-Secure differs from most organizations that use escalation to expedite decision making in that it does so to relieve time pressure at the front end and encourage more reliable decisions as situations grow more problematic. The difference is rooted in F-Secure's orientation to reliability: if the back-end team cannot resolve the issue in a few hours, a time that is expected to suffice for a malware sample's classification and timely updating of the detection engine, it directs the issue to the other unit.

The task of the second unit, the *proactive research unit*, is to react to complex escalations from its sister unit. Also, its analysts independently hunt for new threats (in a form of counter-espionage) and perform new types of analyses to reveal system vulnerabilities (for example, with the aid of the unit's "honeypots," vulnerable servers that attract malware attacks). An important practice of this unit is sharing malware

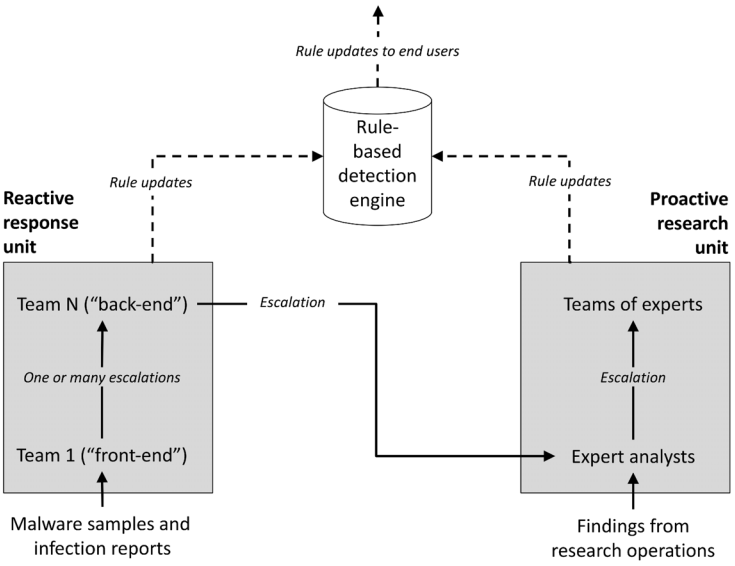

Figure 1. F-Secure's Organization Structure for Malware-Protection Flow

samples with F-Secure's competitors. Despite fierce industry competition, malware-protection companies share samples for normative and ethics reasons. To this end, the unit participates in malware-focused communities and networks. The two-part structure helps F-Secure adapt to the rapidly changing landscape of malware protection, with analysts in both units contributing rules and rule modifications to the algorithmic malware detection and removal carried out in F-Secure's software.

Collection and Analysis of the Data

Previous studies of digital HROs (e.g., Butler and Gray 2006; Ramiller and Swanson 2009) have focused primarily on conditions of reliability on the collective, organizational, level. We directed our attention to operations at the finer granularity organization instead. The level of detail in our empirical analysis was similar to the "activities" level in the APQC framework (see http://www.apqc.org/pcf/) that is used as a hierarchical process-mapping method in operation improvement. The analysis scope encompassed all activities involved in detecting and addressing potential malware threats. These activities are what we have denoted above as *operations*, human-based and digital alike, all of them "concerned with managing inputs (or resources) through transformation processes to deliver outputs" (Rowbotham et al. 2007, p. 2).

For examining the frame problem's impact on high reliability (i.e., answering RQ1), our first step was to assess the presence of the five characteristics of HROs in F-Secure's operations. We sought to identify each as an emergent property and examine whether there was influence by or dependence on the frame problem. To address, in turn, the ways the organization uses algorithms in its digital operations to achieve high reliability (RQ2), we reexamined the data on operations (and their features and connections) to understand how they jointly build a system that generates collective mindfulness.

We conducted four overlapping, iterative phases of data collection and analysis, wherein earlier phases informed subsequent ones (see Table 3). Phase 1 covered semi-structured interviews with two senior managers of F-Secure's malware operations and three senior analysts[9] in the response and research units. These discussions led to the initial insight and conceptual frame for studying F-Secure as an HRO operating in the digital domain. From Phase 1 onward, this frame guided repeated sampling of additional informants to enable critical evaluation of the validity and reliability of inferences from the data obtained in earlier phases. We received significant help from F-Secure managers who at each stage identified the most knowledgeable informants and provided access to them. In Phase 2, we focused on features of operations, seeking to ascertain whether F-Secure's behaviors were consistent with HROs' characteristics. We also started to gather details surrounding the primary unit of analysis (i.e.,

[9]Per F-Secure's request, we do not reveal any information about informants (job role, field of expertise, work site, gender, etc.), because of the attendant risk of compromising their identity to malware creators.

Table 3. The Four Phases of Data Collection

Theme of the Data Collection	Method Used	Duration (Min)	Informant and Work Location	Years at F-Secure	Focus of Outcomes
1. Operations overall	Interview Interview Interview Interview Interview	85 66 55 60 61	Director, HQ Senior Researcher, HQ Service-Owner, HQ Service-Owner, HQ Director, HQ	10 14 7 7 11	F-Secure as an HRO Digital malware-protection architecture (see Figure A1) Organizational structure (see Figure 1)
2. F-Secure as digital HRO	Interview Interview Interview Interview Interview	60 61 56 57 60	Director, HQ Service Owner, HQ Service Owner, HQ Team Leader, Offshore 1 Team Leader, Offshore 1	10 7 7 9 7	Similarities and differences between traditional HROs and F-Secure's operations Operations in malware protection
3. Workflow and communication within F-Secure	Interviews, e-mail, tickets in issue-tracking system	57 85 63 103 38	Service-Owner, HQ Senior Researcher, Offshore 2 Service Owner, HQ Team Leader, Offshore 1 Senior Researcher, HQ	7 5 7 9 14	Vignettes of difficult problem-solving cases Analysis and communication tools
4. Verification of analytical interpretations	Participant observation	270 296	Service-Owner Researcher	7 2	How tools are used in practice Verification of findings from saturated data

operations) that are instrumental to successful protection against malware. Phase 3 focused on gathering and analyzing data on workflow and the division of labor between human-based and digital operations. In this connection, we also developed concrete vignettes of severe malware threats and the associated sets of operations in F-Secure's escalation processes. In the final phase, we conducted participant observation to triangulate the interview-based findings: We shadowed two analysts, for half a day each, who were active in the reputation and detection stages. One of the authors sat beside the informant and took notes, inquiring about the work whenever something needed clarification and when the moment seemed appropriate for an interruption. The visit was followed by immediate transcription of the notes into clearer, less abbreviated accounts of what had been observed. As for the interviews in Phases 1–3, the protocols (reproduced in Appendix B) were formulated in line with earlier HRO research and with insights that emerged in the course of our analysis. The 169 pages of interview transcripts included 13 in-depth stories of difficult problems, with supplemental archival data for 8 of these. The participant observation produced 17 pages of text fleshed out from 44 pages of hand-written notes.

We applied open thematic analysis (Boyatzis 1998) to examine, with the aid of NVivo software, how malware threats were detected, how the detection processes proceeded, and how these processes varied. Our final coding system is shown in Appendix C. We started by searching for thematic indicators of HRO characteristics and mindful/mindless operations. Accordingly, two of the authors coded the data on the basis of the five HRO characteristics. Additional codes were generated for properties of the operational environment and the various roles of technology and algorithmic processing in F-Secure's operations. We triangulated our findings across phases by interviewing informants with potentially contrasting viewpoints (the offshore-team leaders in Phase 2), then with a new type of data (archival data in Phase 3), and finally with the first-person observation (in Phase 4). This addressed all the inferential steps for answering RQ1, on the effect of the frame problem in high-reliability operations.

To address RQ2, Kirsh and Maglio's (1994) classification offered a starting point for coding (see Table 2). We focused on articulating how the operations supported and depended on each other, alongside how they, as a system, together enabled emergence of collective mindfulness. This analysis aimed for better describing and understanding each operation's role in the larger organizational setting. Per Kirsh and Maglio, we classified each operation as either epistemic or pragmatic and used this classification next to theorize about the reasons for the emergent labor division and how responsibilities were allocated between human-based and digital operations. Using indicators from interview data, we mapped all input–output

relationships between operations. For each epistemic operation, we identified the pragmatic operations informed by it (this helped to tie each epistemic operation to one or more pragmatic operation), and for each pragmatic operation we identified digital operations and data on which it operated, along with the nature and scope of its influence on the malware-protection process. Thus, we mapped all epistemic operations that involved gathering information on protection failures to the pragmatic operations wherein analysts modified the digitally encoded malware-detection rules. After this mapping, we arranged the operations into larger blocks: if two epistemic operations were connected to the same pragmatic operation, we grouped them together, and two pragmatic operations acting on the same digital operations or data were assigned to the same block. This helped us abstract operations into groups of higher-level functions in relation to reliability-directed processes, thereby leading us to identify prerequisites for emergence of collective mindfulness in digital operations, the manifestations of operations' mutual dependence, and each operation's role in promoting/enabling higher reliability within F-Secure.

Findings

The Frame Problem in High-Reliability Digital Operations

We commenced our analysis by critically examining the salience of the frame problem in the observed HRO characteristics. A synthesis of the outcomes from our thematic coding is provided in Table 4, columns 1 and 2 (for details, see Appendix C). While none of the operations listed in itself is necessary and sufficient for identifying F-Secure as an HRO, the large number of these operations and their features attest to the company's nature as a digital HRO. The third column presents the primary evidence for answering RQ1 in the affirmative: it summarizes the frame problem's effect in terms of each of the five characteristics. We next review the table's contents in depth, with reference to illustrative cases.

Preoccupation with Failure

Per Weick et al. (1999), the main trait of an HRO is chronic preoccupation with failure, with associated efforts to learn from close calls. This was indeed F-Secure's *modus operandi*. Informants repeatedly described fearing that F-Secure's operations would not be aware of certain threats, be able to recognize them, or be able to protect from them. They associated any such failure with serious consequences for the organization:

> Analyst: *Certain pieces of malware cannot be detected* [through] *normal, routine work, so we need to look at them deeper … because these malware-writers are really creative. When they try to create a new version, they make sure that we have a hard time.*
>
> Researcher: *How is it different, the variant* [i.e., the new version]*?*
>
> Analyst: *You cannot easily see whether the file exists or not, because it's hidden in the rootkit.*

This exemplifies the digital nature of F-Secure's reliability issues also. Computers as physical systems are seldom damaged by malware.[10] Instead of physically compromising the systems, which users usually can easily perceive, malware affects the correctness of logic operations of software run on the computer. Digital operations' editability and programmability permit camouflaging. This leads to constant uncertainty about malware's presence, as hinted at in the extract. The analysts are perpetually preoccupied with the possibility of a threat and a related vulnerability having gone unnoticed and already causing havoc. Investigation of such a possibility has its parallels in traditional HROs' difficulties with anticipating particular complex interactions between systems' tightly interlinked components. An equally important source of preoccupation is the possibility of semantic errors: the detection system may operate from incorrect rules on account of analysts' misplaced or otherwise erroneous rule modifications.

The two forms of preoccupation are intertwined in any concrete "malware-hunting" practice, since the analyst's ability to recognize threats is impaired by both incorrectly specified detection rules and hidden operations of malware. Analysts repeatedly emphasized the difficulty of devising correct detection rules: ones that detect actual errors without being so general as to allow false positives. An inappropriate generality level could result in safe content being categorized as malicious (a false-positive error) or, equally, failure to detect malware and hence to keep it from penetrating the computer (a false negative).

> Analyst: [A false positive means] *we detect a safe Web site as malicious.*
>
> Researcher*: So there's a safe Web site and you rate it as malicious?*

[10] An exception is Stuxnet, which infiltrated an Iranian uranium-processing plant and other industrial facilities in 2010, damaging their centrifuge systems by accelerating the rotor speed beyond physical tolerances.

Table 4. F-Secure Operations Mapped to HROs' Characteristics

HRO Characteristic	Related Operations at F-Secure	The Operations' Relationship to the Frame Problem
1. Preoccupation with failure	1.1 Pretest new rules 1.2 Create "soft" rules 1.3 Solicit customer reports of problems* 1.4 Conduct threat-hunting 1.5 Gather samples via honeypots 1.6 Perform algorithm analysis 1.7 Share samples with competitors 1.8 Analyze samples within a closed network	Operation 1.1 prevented analysts from entering erroneous/conflicting rules for the detection engine and thereby weakening its ability to represent the digital states truthfully. Operations 1.2–1.7 gathered information about new errors in digital operations, thus helping analysts identify when the protection software could not interpret its input and respond correctly. Operation 1.8 prevented unknown errors from entering F-Secure's digital infrastructure.
2. Reluctance to simplify interpretations	2.1 Conduct sample-hunting 2.2 Replicate malware's behavior 2.3 Cross-validate verdicts, using various detection engines 2.4 Perform root-cause and *post mortem* analyses after difficult cases*	Operations 2.1–2.3 gathered information about digital states to be interpreted and acted on in new ways. Operation 2.4 improved operations' attentiveness to critical events and failures. Together, these operations prevented analysts from becoming mindless and operating under the same frame as the underlying digital operations.
3. Sensitivity to operations	3.1 Monitor logs manually 3.2 Deploy automatic log-monitoring 3.3 Solicit customer reports of problems*	These operations helped analysts identify when the protection software could not correctly interpret its input.
4. Commitment to resilience	4.1 Form *ad hoc* problem-solving teams 4.2 Perform root-cause and *post mortem* analyses after difficult cases* 4.3 Code lessons learned into digital operations	These operations helped F-Secure normalize situations wherein a frame underlying the digital operations had "failed."
5. Under-specification of structures	5.1 Modify detection rules 5.2 Patch to address false positives 5.3 Implement single-file detections 5.4 Manage failures through escalation	Operations 5.1–5.3 restored digital operations' ability to address threats when the frame problem had kept them from correctly interpreting inputs. Operation 5.4 enabled flexibly correcting the situation and recovering in the trickiest cases that did not fit the current frame.

*Falls under several characteristics

> Analyst: *Yes, and the effect is that it's blocked in our products. Worst case: you block google.com and the product tells the customers that it's malicious. For example, last year,* [customers] *reported … many false positives, a bunch of binary or executable files. And it was* [due to] *URL-blocking, so I was worried, as it is under my domain. We found out that the blocking was caused mostly by old rules that we had then. These rules were rating a URL as malicious if it met certain criteria, and they were clearly too aggressive. On the other hand, they were also capturing quite a lot of truly malicious sites. So we had a dilemma.*
>
> Researcher: *So you'll end up having false negatives?*
>
> Analyst: *Potentially, yes …. So in order to figure out what we needed to do, we gathered all the URLs that were covered by the rule in question. We analyzed the URLs and found out that they were mainly from China or Russia and not very popular URLs with our user base. While we were still trying to create a more long-term fix, we had to decide whether we should remove the rule and* [make a] *sacrifice temporarily, which means leaving the protection to the rest of the product layers, since our product has three phases of protection anyway.*

The need to hone the most effective rules without false positives exemplifies the ongoing balancing required wherein the necessity of a timely reaction and that of thorough understanding impose opposing demands for the task. When

discussing the fifth characteristic, HROs' under-specification of structures, we will describe how the analysts balanced the two by designing detection rules incrementally through false-positive-related patching (5.2) and single-file detection (5.3).

The preoccupation with failure is evidenced by F-Secure's organizational structure as well. The workflow for malware detection (see Figure 1) has been designed purposefully to mitigate failures in two complementary ways. The response unit responds reactively to "known unknowns": software vulnerabilities identified from past samples and thereby classified into known malware families. Second, the research unit, in handling escalations, uses several distinct methods (operations 1.4, 1.5, and 1.6 in Table 4) to hunt proactively for "unknown unknowns." These comprise threat vectors that are totally new to F-Secure. Another way of learning about them is via the malware-protection companies' sample-sharing (1.7), enabled by the transferability of digital data and operations. Along similar lines, F-Secure involves customers (1.3) in the detection activity by providing multiple means of reporting problems and submitting samples, with a promise of responding rapidly to them.

Finally, F-Secure's safety precautions reflect the preoccupation with failure. Prospective detection rules are pretested (1.1), manually and automatically, before getting implemented in running software. Sometimes analysts publish rules in "soft mode" (1.2), also known as a "dry run": they do not invoke decisions but only report data. Also, analysts perform analyses in an isolated "red" computer network (1.8) to avoid infections within F-Secure (see column 2 in Table 4). These examples attest to the severity of the frame problem's effects. Since the algorithms cannot detect or anticipate errors in their own operation, the whole organization must be arranged around seeking and gathering data on the possible presence of such errors. There is an accompanying suspicion that honing new rules could introduce further errors due to complex interactions of system states/rules or limited understanding of the domain (the primary source of the frame problem).

Reluctance to Simplify Interpretations

Reluctance to simplify reflects the necessity of maintaining divergent points of view and exercising healthy skepticism of one's interpretations (Weick et al. 1999). At F-Secure, this principle's importance is elevated on account of the content's malevolence and the properties of the digital operations. Given malware's editability and programmability, its overt appearance may not reflect its true semantics: while apparently unchanged in meaning (e.g., a digital image seeming the same as before), a piece of data may contain hidden payloads. Because malware is tricky to interpret, F-Secure's analysts must engage in sample-hunting (2.1) to verify what a data sample might hide, or they might have to "run" a sample and replicate its behaviors (2.2) to ascertain maliciousness. Through participant observation, we learned that malware is recognizable in most cases for subsequent attempts to download more data. Sometimes successive payloads differ, depending on temporal or other conditions. We witnessed one malware sample download further samples and launch distinct infections that varied with the environment in which the client's software was running. Hence, to double-check the interpretations, F-Secure's protection software contains not just F-Secure's proprietary detection engine but also an embedded version of a partner's alternative protection software[11] (2.3). Although this may seem surprising, it is a common and accepted practice in the industry.

> Researcher: *How do you make the third-party scanning engine work in practice?*
>
> Analyst*: The thing is that it is an independent firm … a subcontractor at the same time. So they get samples independently from us. We have one of their scanning engines and their database in our product.*

A third example of the use of multiple interpretations is root-cause, or *post-mortem*, analysis (2.4). Such analyses, which are obligatory after a significant escalation, are applied to identify and comprehend reasons for possible errors in F-Secure's analyses and to learn from them.

> *We would have post mortems and gather the information about the case, whether there was a process failure that led to it or what we could have done to ensure it does not happen again or reduce the risk that it happens, what happened, and … determine our future modifications. Maybe the process needs to be tightened, or maybe we need to add another testing or add a new item to the checkbox of items that must be checked.*

Root-cause analyses demonstrate generalization from a single failure (Weick et al. 1999), a principle of HROs whereby every instance of failure is viewed as pointing to a possible deterministic error in the larger system. If a failure can occur once in a digital system, it will happen again, because digital operations' "exact" nature invariably produces the same output for the given input. Through root-cause analysis, the sources of error can be more easily identified and rooted out.

[11] See https://community.f-secure.com/t5/Business/What-engines-does-F-Secure-use/td-p/68282.

The focus on maintaining diversity of perspectives in F-Secure's operations is well motivated by awareness of the frame problem. Algorithmic processing aids in rapid and expedient analysis of samples, but the frame renders interpretations rigid and leaves a void in detecting some possible threats. F-Secure strives to break free of this limitation by increasing the variety of interpretations of the software's behaviors beyond those that digital detection operations would generate on their own.

Sensitivity to Operations

Sensitivity to operations refers to an actor's ability to construct and maintain a detailed picture of the operations and related threats in real time (Vogus and Welbourne 2003; Weick et al. 1999). This is a true challenge in digital, automated environments. The frame problem renders digital operations severely limited: they detect only specific events and states, and they can maintain a picture of only the phenomena they were developed to observe and detect. This challenge persists even when a human analyst is added to the loop, since the analyst's situational awareness depends mainly on the information generated and mediated by the digital system's operations. Because the states and their interdependencies may be highly complex, the analyst is tasked with maintaining a detailed understanding of what is going on in real time. For example, if unfamiliar aggressive malware spreads without any visible behavioral manifestation, no current digital operation can recognize it and mediate relevant information. To detect such unnoticed events or threats, F-Secure uses a customer-service help desk and a Web form (3.3).

> Researcher: *So the work of the response team is quite time-sensitive?*
>
> Analyst: *Yes, exactly. It is time-critical because we have to respond to the cases within the time stated in the SLA* [service level agreement]. *Internally, if it is a High case, we have to solve it in two hours; if it is a Critical case, we have to respond within 30 minutes; and in other cases we have to respond within eight hours or 24 hours.*
>
> [This time]*, the detection hit wrong kinds of files, which resulted in false alarms.* [It was] *nothing massive but still produced false alarms for around 70–80 users. False positives. Eighty users is a lot in our standards. We take our service level seriously.*

Web forms offer a powerful tool for maintaining sensitivity, for three reasons. First, because malware protection is a critical service for many customers, they have greater incentive to provide high-quality information about possible intrusions. Second, every incoming piece of information may be an early signal of a possible epidemic threat, and the escalation flow ensures that F-Secure is prepared to act on each one rapidly and flexibly. Third, ease of transferability leaves few obstacles to malware's spread, so it is likely to surface in geographically dispersed locations. Situation-awareness necessitates a globally accessible, shared reporting tool. Web forms meet this condition as well.

Other awareness-increasing mechanisms rely on event-logging. Tools for manual monitoring of logs (3.1) give means for querying and visualizing possible situations upstream. Our participant observation illuminated how this mechanism helped analysts prioritize their tasks and attend to urgent and potentially more threatening cases. These tools also supply initial diagnostic information on the samples that analysts are examining. Another means of simplified graphical representation facilitating ongoing sensitivity to operations is F-Secure's real-time map that displays the protection software's malware detections on end-users' computers and lists the most frequent malware infections of the last 24 hours (see http://worldmap3.f-secure.com/).

> *So this* [pointing to the map on the informant's laptop] *is where the top 10 goes, and if it goes here to #1 and you don't expect it to be there or you just created a detection and suddenly it spikes there, it would be very curious: is there—what* [do] *you call this?—is there a massive infection happening, or is this a false positive, which usually* [behaves] *like that? It's either a very massive infection or a false positive.*

Automatic log-monitoring (3.2), in turn, uses programs that continuously "watch" the stream of incoming data, for identifying anomalies. When any are found, analysts in the proactive research unit receive an e-mail alerting to a new observation. Such monitoring aids in recognizing situations wherein the digital protection may have started interpreting inputs incorrectly. The customer help desk and monitoring serve as reactive means of identifying possible failures via diagnostic information on present malware-protection behaviors and performance. The frame problem prevents the detection engine from noticing its errors; these added operations increase analysts' ability to recognize and handle potential threats in a more timely manner.

Commitment to Resilience

Commitment to resilience involves striving for graceful coping with surprises, learning from mistakes, and "bouncing back" (Wildavsky 1991). This trait of HROs ties in with the frame problem in that it draws attention to situations wherein inherent weaknesses of digital operations have already led to failures. Commitment to resilience is evident throughout F-Secure's resolution of escalation cases (4.1). It copes with surprises by using flexible, case-by-case allocation of the scarce analyst resources. In the event of escalation, team members abandon other tasks and work together until the problem is solved.

> *It does not depend on what team you belong to, as long as you can help here, ad hoc …. There's a simple process for handling this. For example, me, I know which kinds of experts we need who may help. But that kind of attention is very expensive: we have to stop all the things we are currently doing.*

The toughest ("level-A") escalations prompt more dramatic changes in F-Secure's processes.

> *If that* [escalation] *happens, we have defined in our processes that we must perform a root-cause analysis, and we have also always done so. That is, once the situation is under control, we analyze the reason for the problem and make short-term and long-term plans to make sure that these things will not happen again.*

As the quote suggests, root-cause analyses (4.2) form commitment to resilience, via accumulated knowledge of how to avoid future failures. At F-Secure, the reprogrammability and editability of digital operations enables bouncing back in a manner not possible for traditional HROs: rectification of a detected error can be directly coded into the detection engine with little effort in a new rule (4.3). In such cases, the edit/repair acts as both a recipe for the remedy and the remedy itself. Such resolution is impossible in a physical domain, where a procedure's instructions (its rules) and execution (its physical implementation) remain only loosely coupled. In physical arenas, change in the operations requires new physical components, communications, and training, all of which induce slowness. Thanks to digital operations' editability, transferability, and programmability, F-Secure corrects its errors almost instantly.

> Analyst*: What typically happens here with escalations is that we take a look at the problem. Then we learn from it, and then that kind of escalations does not happen anymore.*
>
> Researcher: *How about in the response unit? How do they learn on the basis of what has been figured out higher in the organization?*
>
> Analyst*: They don't learn. I mean they don't need to.*
>
> Researcher*: So the problem is never actualized again?*
>
> Analyst*: Yes.*

Under-Specification of Structures

Under-specification of structures refers to a tendency for HROs' systems to avoid unnecessarily codified, rigid procedures that result in by-the-book actions completed "on autopilot" (Weick et al. 1999). At F-Secure, problem-solving escalations (5.4) (see Figure 1) form the most direct way to avoid rigidity. Openness to escalations acts as a safety valve in the normal workflow. Exceedingly difficult cases eventually reach the research unit, creating *ad hoc* escalation teams. Nearly all such escalations stem from the representation-related limitations of digital operations; that is, they can be traced to the frame problem. Underlying every escalation is the analysts' awareness that algorithm-based protection cannot properly assess the current state of digital operations.

> Researcher: *What are the reasons they escalate these cases?*
>
> Analyst: *Today, malware infections are motivated by money or information theft. That's why the guys who are writing the malware … are very creative in circumventing the already established solution to known infections. So they invent ways to infect a system or user in a different way. So the known ones that are being served by the server-side rule engine are already circumvented by this. Therefore, it gets escalated to us.*

F-Secure fights such rigidity also with a set of operations specific to digital operations. First, rectifying operations requires changes only to their representations: to instructions. In this context, editability and programmability in digital operations guarantee that analysts can swiftly modify the algorithm by entering new rules (5.1) that overcome the defects found in earlier operations.

> *Instead of writing code to automate the treatment of incoming samples, we have a rule-based mechanism, which allows us to change our workflow in a flexible manner if needed, even within a single day. And we do that often if there is something wrong with the workflow. I simply write new rules, test it, and push it to production. This allows us to do one year's development in one day.*

Second, more flexible, incremental editing and reprogramming are introduced on occasion. Akin to semi-structures (Weick et al. 1999), these include false-positive handling patches (5.2) and single-file detections (5.3). To patch for false positives, analysts moderate overly aggressive behavior whereby rules keep users from accessing safe content, while single-file detections involve applying narrower rules to address false negatives (undetected malware). Both can be rapidly uploaded to users' computers, where they "buy time" while F-Secure attempts to craft a more comprehensive and fully effective set of rules. Analysts describe single-file detection as a response to common "zero-day vulnerabilities" (silent discoveries of security holes in software releases). The vendors of the vulnerable software usually fill such holes too slowly, so F-Secure needs to add corresponding detection to its engine.

> *This case was a Flash exploit; there have been a lot of these zero-days .... In this case, it was a vulnerability in the Flash player and Adobe did not have an update for it yet .... So what I did was that I sent a message to our* [response-unit] *shift and asked them to add a detection* [rule] *for this sample. That was a single-file detection, so ... if we were lucky, the bad guys would use the same sample* [for all users] *and our users would be safe for a while. The purpose was to play* [for] *a bit more time for us.*

In summary, rule updates (5.1–5.3) add flexibility to the "competence" of digital operations when the frame problem prevents such operations from correctly interpreting incoming samples. Escalations (5.4), in turn, help to introduce corrections flexibly in time to address the most difficult cases.

In summary, our analysis demonstrated the frame problem's intimate connection to digital operations' hindrances to high reliability. We also demonstrated F-Secure's need to organize its human-based and digital operations so as to address the frame problem (RQ1) and showed that F-Secure meets all the criteria for an HRO, by identifying several operations instrumental to producing each trait of HROs (in Table 4). It is clear also that, in response to the frame problem and the resulting mindlessness of digital operations, the overall focus in F-Secure's organizational design is on overcoming these limitations and reducing the effects of "fast but dumb" (and hence untrustworthy) algorithms. So far, however, our analysis has been flat, failing to describe how these operations mesh and thereby create a system characterized by both mindlessness and mindfulness, one that still produces high reliability. We turn to this question about emergence next, examining the organization of the operations and their interconnections.

Organizing Digital Operations for Collective Mindfulness

Any HRO's ability to use digital operations in part to support collective mindfulness depends ultimately on how it organizes its overall system of operations holistically (Butler and Gray 2006). Regrettably, such structuring principles for semiautonomous operations have remained opaque and largely unexplored in HRO research. We do not yet know how a HRO can successfully organize its overall "stack" of operations for high reliability in a digital domain plagued by the frame problem.

Below, we will consider how the interdependent organization of epistemic and pragmatic operations contributes to emergence of collective mindfulness. We find value in Kirsh and Maglio's (1994) epistemic–pragmatic dimension for directing us to operations that produce information and to operations that influence the environment in response to that information. It also helps us consider the connections between the two. As we deployed this classification, however, we became aware of three higher-level operations that fell into neither class. We excluded these operations, which covered general ways in which other operations could be carried out or adjusted to context, from further analysis. These were analyses of samples within an isolated network (1.8), *ad hoc* problem-solving teams (4.1), and escalation-based failure management (5.4). Alongside the 17 operations remaining, we added the "default" core operation within F-Secure's malware-protection system: automated, algorithm-based (rule-encoded) detection of malware. We had omitted this from earlier analysis because it forms the *mindless core* of F-Secure's digital operations and is performed solely by computers. Deployed as much for efficiency (being fast, efficient, and without downtime) as for reliability (not making errors if properly programmed), the core operation forms the most fundamental, pragmatic operation in the company's (digital) activities. Ultimately, it is the sole operation that detects and removes pieces of malware.

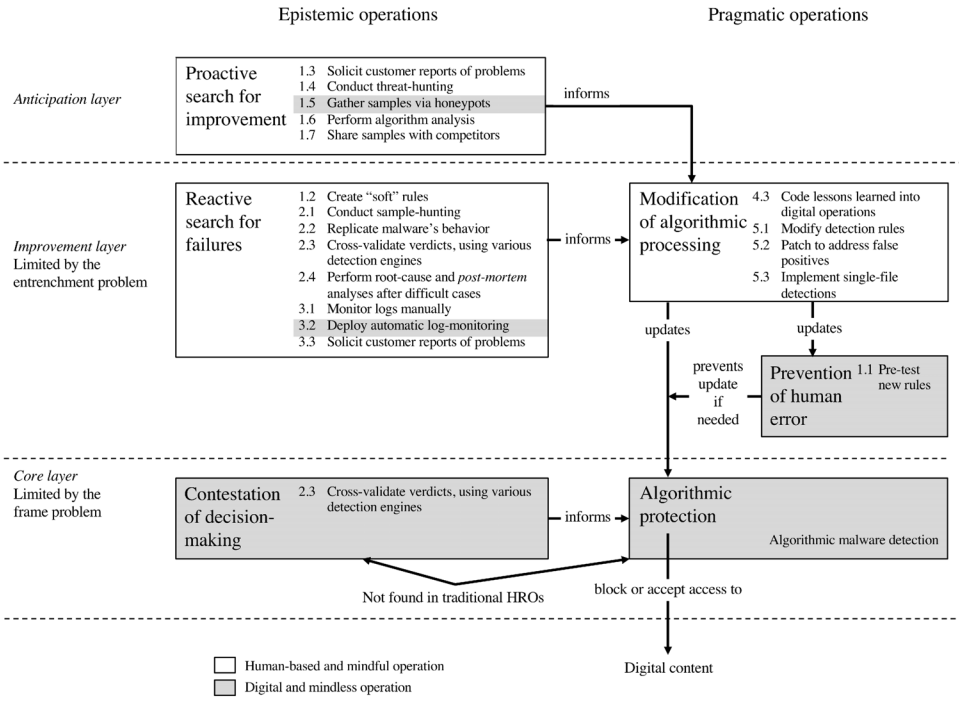

Figure 2. F-Secure's Organization of Operation Blocks Forming a Three-Layer Structure

Our method of analysis enabled clustering operations (with their various interactions) into blocks. Figure 2 offers a graphical summary. We will follow three steps in discussing this analysis. In the first step, we describe how the 3+3 blocks of operations in Figure 2 give F-Secure a set of mechanisms to address the frame problem. Next, we discuss how the structuring of the blocks forms a foundation for the emergence of collective mindfulness. Finally, we use Figure 2 to theorize about mutual reinforcement of responsibilities between human and digital operations and their role in the emergence of collective mindfulness. This aids in delving further into how traditional and digital HROs differ.

Mitigating the Frame Problem by Layering Operations

Each layer in Figure 2 serves distinct organizational and cognitive purposes. This helps to crystallize the insight that human elements in digital organizing exist primarily to monitor the digital operations' reliability and to revise them as necessary. The figure highlights the threat of the entrenchment problem also (Dane 2010): human operators are always susceptible to their own "frame problem" if seeking to correct errors only reactively. Accordingly, F-Secure's three-layer organization serves to mitigate both the frame problem and the entrenchment problem. Per Figure 2, the digital core and its operations must be constantly monitored by the analysts, whose reactive repairs, in turn, must be monitored and improved via constant proactive searches for unknown threats. The layers are formed around core, improvement, and anticipation functions.

F-Secure's *core layer* is composed of two blocks of operations. The algorithmic-protection block consists of pragmatic operations and is entirely digital. It runs physically on client computers and in the cloud. Beside it resides the layer's epistemic block, which we refer to here as contestation of algorithmic decision-making. Here, a third-party detection engine (operation 2.3 in Table 4) provides complementary information for aiding in evaluating algorithmic-protection-process decisions. The core layer's operations (shown in gray in Figure 2) are all mindless: they are algorithmic and constrained by the frame problem.

The *improvement layer* continuously repairs the core layer's operations to mitigate its unavoidable errors and related failures. To achieve high reliability, F-Secure cannot let the core run amok, so it engages in continuous mindful surveillance of the core. At the second layer we find both epistemic and pragmatic operations, jointly heeding potential threats in the core's operations. Analysts modify the detection rules in

a pragmatic, human-based mindful block denoted as modification of algorithmic processing. Its epistemic counterpart—reactive search for failures—informs and guides modifications by focusing on "known unknown" threats. This search is reactive and attentive to weaknesses already identified in the core.

> [A rule-based engine] *still needs, of course, an analyst guiding the machine, fixing the rules.*
>
> *These guys* [in the response unit]*, they may also fix the rules. Or they will just, like, do immediate fixes and let lab analysts fix the final rules.*
>
> *The report was manual. The fix that we did was manual, I mean he fixed the code. The response team's intermediate fix of that one URL was manual* [after which it was escalated to the research unit]*, and pretty much the changing of the rules that I did was manual.*

This layer's epistemic block is similar to the pragmatic one. It is primarily human-based and mindful (as illustrated by the white background in the figure). In contrast, the layer's third block, prevention of human error, performs automated testing of new rules before the analysts' modifications take effect at the core layer. These tests identify and draw attention to lapses and potentially conflicting rules, thereby addressing human error rooted in difficulties in working with the algorithmic, complicated, and highly precision-demanding digital operations. After all, small variations may bring huge changes in behavior.

While these mindful human operations at the improvement layer prevent F-Secure from falling prey to the frame problem, they can still be hampered by human entrenchment, which threatens to make F-Secure's operations blind to possible "black swans" (Taleb 2007)—surprises whose signals are not foreseeable or visible in the relevant digital operations. A case in point would be a malware type with a truly novel exploitation pattern: countermeasures to entrenchment, such as seeking information outside one's own task domain (Dane 2010), are unavailable and would not succeed at the improvement layer, because its operations are reactive, backward-looking, and focused on addressing visible threats on a short time horizon.

The third layer, the *anticipation layer,* counters threats of entrenchment. It consists of epistemic efforts to turn "unknown unknowns" (black swans) into "known unknowns." This layer's operations are outward-facing, to anticipate threats proactively and reveal needs to update rules or restructure operations.

> *We have active individuals who monitor the threat landscape and follow Twitter and blogs, listening to what is going on.*
>
> *Some of us have a role for doing threat-hunting, in principle to follow what happens in the world on this malware scene. Follow all the possible sources that we have, noticing that now there is a zero-day infection or this kind of proof of concept* [or that] *someone has identified a vulnerability. Usually that leads to a moment at which some malware-creators start exploiting this vulnerability and deploy malware using it. Our goal is that we already have protections for it.*

In effect, the anticipation layer supplies the improvement layer with alternative, fresh viewpoints and resists short-sightedness born out of entrenchment arising from its reactive approach to the threats at hand.

This three-layered organization of operations makes F-Secure capable of addressing all the aforementioned problems that can compromise high reliability in digital operations. With this structure, F-Secure anticipates black swans that are likely to compromise detection performance in the long term. It also addresses the mutually conflicting needs for high performance and high reliability. F-Secure achieves efficiency and scale through algorithmic processing at the mindless core layer; it gains rapid and mindful adaptation at the improvement layer via continuous, swift, and correct rule-engine changes; and, finally, it engages in mindful anticipation of unknown threats at the anticipation layer—manifesting preoccupation with failure and reluctance to simplify operations.

Emergence of Collective Mindfulness

While F-Secure thus addresses the frame-problem-imposed limits to high reliability through careful and systemic layering of its operations, this does not yet account for how such an organization contributes to collective mindfulness as a balancing act between the mindful and mindless operations. We posit that this systemic property emerges from specific couplings between operations within and (especially) between layers. Therefore, we carry our analysis further by connecting each operation block to each HRO characteristic, proceeding from our coding and analysis of the operations' structure. Figure 2's depiction of the six blocks and their arrangement assists in understanding how the various operations enmesh in an organization-wide arrangement contributing to the emergence of collective mindfulness.

In the figure, we can observe tight couplings between HRO characteristics and the epistemic or pragmatic operations. Activities concerned with preoccupation with failure (coded as 1.x), reluctance to simplify interpretations (2.x), and sensitivity to operations (3.x) are epistemic, while commitment to resilience (4.x) and under-specification of structures (5.x) are exclusively pragmatic. The only exception to this division is pretesting of new rules (1.1). While related primarily to the preoccupation with failure, it is pragmatic in that it prevents analysts from introducing errors. Overall, the tight coupling between operation types and HRO traits suggests a division of responsibilities between human-based and digital operations within digital HROs. We note also that, *while the epistemic operations can be either human-based or digital, the pragmatic ones must be human-based outside the digital core, to overcome the frame problem*.

This principle precisely articulates the difference between digital and traditional HROs. In traditional (sociotechnical) HROs, humans operate directly on physical artifacts through pragmatic operations, some perhaps digitally mediated. In contrast, the core operations defining a digital HRO's principal input–output operations *vis-à-vis* its environment are fully digital in nature. Human-based pragmatic operations influence these processes indirectly, through modifications to algorithms that guide core-layer operations or to environments where the algorithms run. Accordingly, we posit that *the core and its algorithmic mindless pragmatic operations form the essential characteristic of digital HROs*. In other respects, the arrangement in Figure 2 does not differ from the operations of the HROs that were recently reviewed by Dernbecher and Beck (2017). For example, most HROs apply pragmatic algorithm-based safety mechanisms to overcome the challenges of potential human errors.

We suggest also that, overall, the three-layered architecture of operations abstracted from F-Secure is generalizable to most HROs. To conclude, we find that collective mindfulness as an emergent system-level property arises from a systemic confluence of all HRO characteristics and their grounding in the systemic organization of related operations. This implies that collective mindfulness is impossible without the presence of mindless operations and a connection to them (Carlo et al. 2012; Farjoun 2010; Levinthal and Rerup 2006). It is precisely the specific form of interleaving and systemic coupling between epistemic and pragmatic operations—human-based and digital—that leads to collective mindfulness emerging.

Discussion

Since they are driven by reliability over efficiency, all HROs display the five defining characteristics presented. However, there are differences. The company in our case study operates on digital inputs and outputs, and its core operations must rely on efficient algorithmic processing. F-Secure necessarily suffers from the frame problem and related threats to high reliability: at any given time, it is potentially prone to fatal errors since its view of the environment remains strictly bounded. This stands in stark contrast against traditional HROs, wherein mindlessness arises mainly from cognitive limitations in human operations and from operational inaccuracies snowballing to fatal outcomes through complex interactions of the technological system (e.g., Perrow 1984; Weick and Sutcliffe 2001). Algorithms, with their exactness and high-speed performance, overcome such cognitive limitations, so their role in safety-critical operations is growing. So are the scope and extent of digital operations, with such operations forming the heart of detecting money laundering, maintaining scalable cloud services, operating national banks, managing cryptocurrencies, controlling transportation systems (soon to include autonomous cars and air-traffic control), and many Internet of Things operations. Yet, our study is the first to investigate how high reliability can be achieved and maintained in such a special class of HRO. Our research also addresses a paradox of these HROs: on one hand, they rely on algorithms to achieve efficiency *and* high reliability, while, on the other, these algorithms remain the primary source of their mindless behavior. This paradox can be understood only in terms of three facets of digital operations: the unique properties of digital technologies and related operations (see Table 1); the presence of the frame problem and its potential negative implications for digital operations' reliability; and the criticality of weaving human-based, mindful epistemic and pragmatic operations in with the algorithm-based digital operations in addressing the frame problem.

By identifying the complementary roles of mindful and mindless operations in digital HROs, our work makes several contributions to HRO theory and studies of such organizing. We have expanded HRO theory by identifying new challenges in organizing digital HROs for high reliability. We showed that the frame problem is unavoidable in such settings and that its mitigation calls for new means to create collective mindfulness. All five key traits of F-Secure as an HRO were affected by the frame problem, inviting careful sociotechnical orchestration of operations that builds collective mindfulness. This was enabled by a systemic coupling of human and digital operations across several layers wherein pragmatic decisions were the responsibility of humans in the upper layers while epistemic operations could be either human or digital, as dictated by the nature of the relevant information processing. In particular, we proposed that a happy marriage between mindful and mindless operations is a precondition to generating collective mindfulness at the system level.

Table 5. Roles of Digital Processing in Achieving High Reliability

Nature of Digital Operation	Role for High Reliability
Epistemic	Inform operations
Pragmatic (control operations)	Prevent physical operations if, for example, danger of grave human error exists
Pragmatic (execute operations)	Influence digital inputs directly without humans in the action loop (this demands mindful monitoring and updates, both reactive and proactive)

Second, we have advanced scholarship by attending to details of operations and their structuring. Although many earlier HRO studies involved in-depth *in situ* collection and analysis of data, their main contribution lay in apprehending and inductively generalizing to emergent organization-wide cognitive properties (e.g., Grabowski and Roberts 1999; Roberts 1990; Weick et al. 1999). While exceptions exist—Bigley and Roberts (2001) examined a hierarchy of operations in emergency response, Valorinta (2009) identified the role of IT-based operations but only in *non*-high-reliability logistics, and Carlo et al. (2012) considered digital operations' role in construction operations (although without a systematic model of the structuring)—our focus remains unique. We have examined digital and human operations in awareness of their distinctive nature, systemic couplings, and interactions between algorithmic and human-based processing. In this regard, we see ample opportunities for additional fine-grained operations-level analyses examining how information-conveying tools can be used and appropriated in alternative couplings, in relation to each of the traits of HROs (e.g., Wahlström et al. 2011).

Third, we have offered new insight with regard to the frame problem's significance for reliability of organizations that depend on digital operations and organizing. If an organization's core (digital) operations employ algorithmic processing and those operations could have hazardous outcomes, it must somehow address and reduce negative consequences of the frame problem. Surprisingly, we did not find any prior treatment of the frame problem or of its importance in connection with systemic computing risks and business-continuity studies. The few papers on the role of IT in organizational reliability and sustainability of operations neither acknowledge nor address the frame problem (e.g., Butler and Gray 2006; Carlo et al. 2012). They focus primarily on reliability threats in organizations' IT use in terms of its cognitive orientation, thereby relegating the frame-problem-addressing structuring principles to the background.

In light of our analysis, we can now summarize digital operations' roles in terms of three types: an epistemic role, a pragmatic role in control of operations, and a pragmatic role in executing operations (see Table 5). The first two are common to all HROs, while the third is specific to digital HROs. This framework can aid in the search for potential positive effects of digital operations in high-reliability settings. When such roles are recognized and orchestrated by means of a layered organization that incorporates human-based operations, overcoming the frame problem and achieving collective mindfulness becomes possible as the human "sits in the backseat."

We note several limitations in the study. This is exploratory research carried out with a single organization, so the findings do not automatically extend to other settings. The results call for careful validation via analyses across a broader range of digital operations. Money-laundering detection, large-scale cloud-based services, central banks' operations, and air-traffic control offer good initial opportunities. Such operations are constantly exposed to unknown threats, intentional (hacking, theft, vandalism, etc.) and unintentional (from the weather and accidents). We should stress too that we encountered some data collection restrictions. We were not allowed to report informants' gender, specific roles in the organization, or identity, details that may be relevant to interpreting the context and data. Also, for competition and security reasons, F-Secure did not share performance metrics or details on the full scale of its operations. Additionally, security concerns precluded recording video during our participant observation, so we were dependent solely on hand-written notes. However, these limitations do not decrease the quality of the data we did collect or the analyses: all inferences presented follow from reported and recorded data alone. For example, we reached empirical and theoretical saturation fairly easily in collecting and analyzing interview data, and the participant-observation evidence sufficed for triangulation of the findings and evidence. Still, broader sampling of employees, over extended periods (in an ethnographic sense), that involves various times of day and a wider spectrum of task follow-up could have led to more accurate and finer-grained data, informing increasingly revealing analyses of high-reliability practices. In particular, we missed the window of rare black swans in terms of situations that culminate in extreme escalation (in the near-miss spirit). Instead, understanding the role of such situations required relying on informants' recall ability.

Future research needs to expand on the work in several directions. We note three here: examining differences between human-based and algorithmic processing; a well-fleshed-out synthesis of HRO and NAT theories; and connections between the frame problem and particular characteristics of the digital operations.

First, more extensive analysis of the differences between digital operations' digital and human-based processing is in order. While we took the frame problem as a basic premise to inform the empirical study, the frame problem is likely manifested in multiple, setting-influenced ways. Whenever possible, further research should give an assessment of the pertinent context-specific cognitive differences between humans' work and algorithm-based processing, alongside particular manifestations of the frame problem. Also, our study invites additional, related research on the generalizability of the organizing principles we found for digital operations, so as to validate whether the three-layer structure identified and the classing of operations as epistemic or pragmatic are applicable to a wider range of settings. Scholars should expand the work into richer contextually and temporally aware analyses (e.g., focused on how various operations and their couplings develop and specialize over time). Analysis of more varied forms of couplings is important, because our findings from F-Secure's use of a centralized system to store, coordinate, and manage detection rules may not hold for all digital HRO settings.

Second, while NAT and HRO theory have been articulated as mutually exclusive views of organization in high-risk environments (e.g., Leveson et al. 2009), neither on its own offers a theoretical frame that sensitizes to unique threats stemming from digital operations and data. Given the importance of digital systems and operations in contemporary society, considering the two schools of thought alongside each other is insufficient: they could be integrated for fuller theoretical insight by analyzing the nature of digital operations more carefully. F-Secure's operations, for example, exhibit multiple features consistent with characteristics dealt with in NAT. For instance, F-Secure accepts that rules cannot be designed "once and for all"; they remain provisional, awaiting revision. It also acknowledges the possibility of complex interactions between rules. Aware that rule-specification errors and unexpected interactions could occur, F-Secure strives to stay attentive in this regard, so as to be ready for action in case these become reality. This closely approaches NAT-type perspectives, taking escalation to fatal error as inevitable whenever a specific constellation of system parameters, however unlikely, is present (Perrow 1984). Yet a difference remains: F-Secure does not approach risks in the probabilistic terms NAT suggests. It does *not* consider failures to arise from highly unlikely external conditions and combinations of these. Its industry is key in this regard: conditions here are far from coincidental—they have an intentional origin. This points to two research opportunities. The first involves synthesis: investigating, accordingly, how the two, opposing theoretical views can be integrated in future studies of digital operations (not least because digital HROs may not fit either perfectly). Second, the nature and role of intentional threats deserves more attention, especially for grappling with today's growing specters of carefully orchestrated threats. Where threats are not accidental but opportunistic, one would expect more frequent potential escalations. Also, malware-creators' intentions may lead to a broad spectrum of threats vastly different from the sort found in other digital domains, such as copyright infringement, forgery, or digital identity theft.

Finally, we have provided the first presentation of conceptual relations among high reliability, the frame problem, and properties of digital operations (exactness, transferability, editability, and programmability). In so doing, we have suggested a model for a digital (sociotechnical) system's internal organizing for high reliability. Certainly, other properties of digital operations and data also might influence reliability, and other forms of organizing may be instrumental in ensuring high reliability. Hence, we would welcome more advanced theorizing on the links among digital operations, their traits, and high reliability. For example, machine-learning techniques may introduce entirely new sets of pitfalls to high reliability, exemplified in recent accidents involving autonomous cars and bicyclists. These streams of inquiry could signal establishment of a research field focused on digital high-reliability organizing.

Acknowledgments

Antti Salovaara received financial support for this work from Academy of Finland (grants 259281 and 298879). We are eternally thankful to the F-Secure managers, in particular Kimmo Kasslin who generously provided access to informants, and employees who shared their time, enthusiasm, and knowledge, without which this study would have been impossible. We are grateful also to Jannis Kallinikos, Richard Boland, the senior editor, associate editor, and three anonymous reviewers whose constructive criticism and support made the manuscript crisper and more focused, reflecting its true potential. Naturally, all errors remain ours—in the spirit of human entrenchment.

References

Bigley, G. A., and Roberts, K. H. 2001. "The Incident Command System: High-Reliability Organizing for Complex and Volatile Task Environments," *Academy of Management Journal* (44:6), pp. 1281-1299.

Boland, R. J., Tenkasi, R. V., and Te'eni, D. 1994. "Designing Information Technology to Support Distributed Cognition," *Organization Science* (5:3), pp. 456-475.

Bourrier, M. 1996. "Organizing Maintenance Work at Two American Nuclear Power Plants," *Journal of Contingencies and Crisis Management* (4), pp. 104-112.

Boyatzis, R. E. 1998. *Transforming Qualitative Information: Thematic Analysis and Code Development*, Thousand Oaks, CA: SAGE Publications.

Brooks, R. A. 1991. "Intelligence without Representation," *Artificial Intelligence* (47:1-3), pp. 139-159.

Brynjolfsson, E. , and McAfee, A. 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, New York: W. W. Norton & Company.

Butler, B. S., and Gray, P. H. 2006. "Reliability, Mindfulness and Information Systems," *MIS Quarterly* (30:2), pp. 211-224.

Cameron, K. S. 1986. "Effectiveness as Paradox: Consensus and Conflict in Conceptions of Organizational Effectiveness," *Management Science* (32:5), pp. 539-553.

Carlo J., Lyytinen, K., and Boland, R. 2004. "Systemic Risk, IT Artifacts, and High Reliability Organizations: A Case of Constructing a Radical Architecture," in *Proceedings of the 25th International Conference on Information Systems*, paper 56.

Carlo, J., Lyytinen, K., and Boland, R. 2012. "Dialectics of Collective Minding: Creating Radical Architecture with Information Technology," *MIS Quarterly* (36:4), pp. 1081-1108.

Cohen, M. D., March, J. D., and Olsen, J. P. 1972. "A Garbage Can Model of Organizational Choice," *Administrative Science Quarterly* (17:1), pp. 1-25.

Dane, E. 2010. "Reconsidering the Trade-Off Between Expertise and Flexibility: A Cognitive Entrenchment Perspective," *Academy of Management Review* (35:4), pp. 579-603.

Dennett, D. 1984. "Cognitive Wheels: The Frame Problem of AI," in *Minds, Machines And Evolution*, C. Hookway (ed.), Cambridge, UK: Cambridge University Press, pp. 129-152.

Dernbecher, S., and Beck, R. 2017. "The Concept of Mindfulness in Information Systems Research: A Multi-Dimensional Analysis," *European Journal of Information Systems* (26:2), pp. 121-142.

Domingos, P. 2012. "A Few Useful Things to Know about Machine Learning," *Communications of the ACM* (55:10), pp. 78-87.

Dreyfus, H. L. 1972. *What Computers Can't Do: A Critique of Artificial Reason*, Cambridge, MA: The MIT Press.

Dreyfus, H. L. 1979. *What Computers Still Can't Do: A Critique of Artificial Reason*, Cambridge, MA: The MIT Press.

Eisenhardt, K. M. 1989. "Building Theories from Case Study Research," *Academy of Management Review* (14:4), pp. 532-550.

Ericsson, K. A. 2009. "Enhancing the Development of Professional Performance: Implications from the Study of Deliberate Practice," in *Development of Professional Expertise: Toward Measurement of Expert Performance and Design of Optimal Learning Environments*, K. A. Ericsson (ed.), Cambridge, UK: Cambridge University Press, pp. 405-431.

Farjoun, M. 2010. "Beyond Dualism: Stability and Change as a Duality," Academy of *Management Review* (35:2), pp. 202-225.

Fiske, S. T., and Taylor, S. E. 1991. *Social Cognition* (2nd ed.), New York: McGraw-Hill.

Grabowski, M., and Roberts, K. H. 1999. "Risk Mitigation in Virtual Organizations," *Organization Science* (10:6), pp. 704-721.

Hutchins, E. 1995. *Cognition in the Wild*, Cambridge, MA: The MIT Press.

Janis, I. L. 1972. *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes*, Oxford, UK: Houghton Mifflin.

Kahneman, D. 2011. *Thinking, Fast and Slow*, New York: Farrar, Straus & Giroux.

Kallinikos, J. 2012. "Form, Function, and Matter: Crossing the Border of Materiality," in *Materiality and Organizing: Social Interaction in a Technological World*, P. M. Leonardi, B. A. Nardi, and J. Kallinikos (eds.), Oxford, UK: Oxford University Press, pp. 68-87.

Kallinikos, J., Aaltonen, A., and Marton, A. 2013. "The Ambivalent Ontology of Digital Artifacts," *MIS Quarterly* (37:2), pp. 357-370.

Kirsh, D., and Maglio, P. 1994. "On Distinguishing Epistemic from Pragmatic Action," *Cognitive Science* (18:4), pp. 513-549.

Leveson, N., Dulac, N., Marais, K., and Carroll, J. 2009. "Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems," *Organization Studies* (30:2-3), pp. 227-249.

Lévi-Strauss, C. 1966. *The Savage Mind*, Chicago, IL: University of Chicago Press.

Levinthal, D., and Rerup, C. 2006. "Crossing an Apparent Chasm: Bridging Mindful and Less-Mindful Perspectives on Organizational Learning," *Organization Science* (17:4), pp. 502-513.

Lewis, H. R., and Papadimitriou, C. H. 1998. *Elements of the Theory of Computation*, Upper Saddle River, NJ: Prentice-Hall.

Louis, M. R., and Sutton, R. I. 1991. "Switching Cognitive Gears: From Habits of Mind to Active Thinking," *Human Relations* (44:1), pp. 55-76.

Malone, T. W., and Crowston, K. 1994. "The Interdisciplinary Study of Coordination," *ACM Computing Surveys* (26:1), pp. 87-119.

Marcus, A. 1995. "Managing with Danger," *Industrial and Environmental Crisis Quarterly* (9), pp. 139-152.

McCarthy, J., and Hayes, P. J. 1969. "Some Philosophical Problems from the Standpoint of Artificial Intelligence," in *Machine Intelligence: No. 4*, B. N. Meltzer and D. Michie (eds.), Edinburgh, UK: Edinburgh University Press, pp. 463-502.

Merminod, V., Rowe, F., and Watts, S. 2008. "Product Lifecycle Management, Knowledge Integration and Reliability in New Product Co-Development: A Case Study Between Europe and China," in *Proceedings of the 29th International Conference on Information Systems,* paper 67.

Oliver, N., Calvard, T., and Potočnik, K. 2017. "Cognition, Technology, and Organizational Limits: Lessons from the Air France 447 Disaster," *Organization Science* (28:4), pp. 729-743.

Perrow, C. 1984. *Normal Accidents: Living with High-Risk Technologies*, New York: Basic Books.

Perrow, C. 2010. "The Meltdown Was Not an Accident," in *Markets on Trial: The Economic Sociology of the U.S. Financial Crisis: Part A,* M. Lounsbury and P. M. Hirsch (eds.), Bingley, UK: Emerald Group Publishing Limited, pp. 309-330.

Pu, J., and Kishore, R. 2006. "Assuring IT Services Quality through High-Reliability Risk Management in Offshore Business Process Outsourcing," in *Proceedings of the 27th International Conference on Information Systems*, paper 79.

Ramiller, N. C., and Swanson, B. E. 2009. "Mindfulness Routines for Innovating with Information Technology," *Journal of Decision Systems* (18:1), pp. 13-26.

Reason, J. 2000. "Human Error: Models and Management," *British Medical Journal* (320:7237), pp. 768-770.

Roberts, K. H. 1990. "Some Characteristics of One Type of High Reliability Organization," *Organization Science* (1:2), pp. 160-176.

Roberts, K. H., and Rousseau, D. M. 1989. "Research in Nearly Failure-Free, High-Reliability Organizations: Having the Bubble," *IEEE Transactions* on *Engineering Management* (36:2), pp. 132-139.

Roberts, K. H., Stout, S. K., and Halpern, J. J. 1994. "Decision Dynamics in Two High Reliability Military Organizations," *Management Science* (40:5), pp. 614-624.

Rochlin, G. 1993. "Defining 'High Reliability' Organizations in Practice: A Taxonomic Prologue," in *New Challenges to Understanding Organizations*, K. H. Roberts (ed.), New York: Macmillan, pp. 11-32.

Rochlin, G. I., La Porte, T. R., and Roberts, K. H. 1987. "The Self-Designing High-Reliability Organization: Aircraft Carrier Flight Operations at Sea," *Naval War College Review* (40), pp. 76-90.

Roth, E. M., Multer, J., and Raslear, T. 2006. "Shared Situation Awareness as a Contributor to High Reliability Performance in Railroad Operations," *Organization Studies* (27:7), pp. 967-987.

Rowbotham, F., Galloway, L., and Azhashemi, M. 2007. *Operations Management in Context* (2nd ed.), Oxford, UK: Butterworth-Heinemann.

Sagan, S. D. 1993. *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons*, Princeton, NJ: Princeton University Press.

Salovaara, A., Lyytinen, K., and Penttinen, E. 2015. "Flexibility Vs. Structure: How to Manage Reliably Continuously Emerging Threats in Malware Protection," in *Proceedings of the 48th Hawaii International Conference on System Sciences*, Los Alamitos, CA: IEEE Computer Society Press.

Sammon, D., Nagle, T., and McAvoy, J. 2014. "Analysing ISD Performance Using Narrative Networks, Routines and Mindfulness," *Information and Software Technology* (56), pp. 465-467.

Shanahan, M. 2016. "The Frame Problem," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta (ed.), Stanford, CA: The Metaphysics Research Lab, Stanford University.

Simha, A., and Kishore, R. 2011. "Social Capital and IT as Predicates of Collective Mindfulness and Business Risk Mitigation: A Grounded Theory Development," in *Proceedings of the 32nd International Conference on Information Systems*, paper 32.

Sonenshein, S. 2016. "Routines and Creativity: From Dualism to Duality," *Organization Science* (27:3), pp. 739-758.

Swanson, E. B., and Ramiller, N. 2004. "Innovating Mindfully with Information Technology," *MIS Quarterly* (28:4), pp. 553-583.

Taleb, N. N. 2007. *The Black Swan: The Impact of the Highly Improbable*, New York: Random House.

Thatcher, J. B., Wright, R. T., Sun, H., Zagenczyk, T. J., and Klein, R. 2018. "Mindfulness in Information Technology Use: Definitions, Distinctions, and a New Measure," *MIS Quarterly* (42:3), pp. 831-847.

Tilson, D., Lyytinen, K., and Sørensen, C. 2010. "Digital Infrastructures: The Missing IS Research Agenda," *Information Systems Research* (21:4), pp. 748-759.

Valorinta, M. 2009. "Information Technology and Mindfulness in Organizations," *Industrial and Corporate Change* (18:5), pp. 963-997.

Van de Walle, B., and Turoff, M. 2008. "Decision Support for Emergency Situations," *Information Systems and e-Business Management* (6:3), pp. 295-316.

Van Den Eede, G., Van de Walle, B., and Rutkowski, A.-F. 2006. "Dealing with Risk in Incident Management: An Application of High Reliability Theory," in *Proceedings of the 39th Hawaii International Conference on System Sciences*, Los Alamitos, CA: IEEE Computer Society Press.

Vicente, K. J., and Rasmussen, J. 1990. "The Ecology of Human–Machine Systems: Mediating Direct Perception in Complex Work Domains," *Ecological Psychology* (2:3), pp. 207-249.

Vogus, T. J., and Welbourne, T. M. 2003. "Structuring for High Reliability: HR Practices and Mindful Processes in Reliability-Seeking Organizations," *Journal of Organizational Behavior* (24:7), pp. 877-903.

Wahlström, M., Salovaara, A., Salo, L., and Oulasvirta, A. 2011. "Resolving Safety-Critical Incidents in a Rally Control Center," *Human–Computer Interaction* (26:1 & 2), pp. 9-37.

Waller, M. J., and Roberts, K. H. 2003. "High Reliability and Organizational Behavior: Finally the Twain Must Meet," *Journal of Organizational Behavior* (24:7), pp. 813-814.

Weick, K. E., and Roberts, K. H. 1993. "Collective Mind in Organizations: Heedful Interrelating on Flight Decks," *Administrative Science Quarterly* (38:3), pp. 357-381.

Weick, K. E., and Sutcliffe, K. M. 2001. *Managing the Unexpected: Assuring High Performance in an Age of Complexity*, San Francisco: Jossey-Bass.

Weick, K. E., and Sutcliffe, K. M. 2006. "Mindfulness and the Quality of Organizational Attention," *Organization Science* (17:4), pp. 514-524.

Weick, K. E., Sutcliffe, K. M., and Obstfeld, D. 1999. "Organizing for High Reliability: Processes of Collective Mindfulness," in *Research in Organizational Behavior* (Volume 1), R. S. Sutton and B. M. Staw (eds.), Stanford, CT: JAI Press, pp. 81-123.

Whetten, D., and Cameron, K. 1994. "Organizational Effectiveness: Old Models and New Constructs," in *Organizational Behavior: The State of the Science*, J. Greenberg (ed.), Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 135-154.

Wildavsky, A. 1991. *Searching for Safety*, New Brunswick, NJ: Transaction Books.

Winograd, T., and Flores, F. 1986. *Understanding Computers and Cognition: A New Foundation for Design*. Reading, MA: Addison-Wesley.

Wittgenstein, L. 1922. *Tractatus Logico-Philosophicus*, London: Kegan Paul.

Yin, R. K. 2009. *Case Study Research: Design and Methods* (Vol. 5), Thousand Oaks, CA: SAGE Publications.

Yoo, Y., Henfridsson, O., and Lyytinen, K. 2010. "The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research," *Information Systems Research* (21:5), pp. 724-735.

About the Authors

Antti Salovaara (Ph.D., Cognitive Science, University of Helsinki) is an adjunct professor of Computer Science at University of Helsinki and a Lecturer of Computer Science at Aalto University. His research interests include appropriation and creative use of information systems, field research methods, and knowledge work. His publications have appeared both in human–computer interaction and information systems journals and conferences.

Kalle Lyytinen (Ph.D., Computer Science, University of Jyväskylä; Dr. h.c. Umeå University, Copenhagen Business School, Lappeenranta University of Technology) is Distinguished University Professor and Iris S. Wolstein professor of Management Design at Case Western Reserve University, and a distinguished visiting professor at Aalto University, Finland. Between 1998 and 2018 he was the third most productive scholar in the IS field when measured by the AIS basket of eight journals; he is among the top five IS scholars in terms of his h-index (81); he is a LEO Award recipient (2013), a Fellow of the AIS (2004), and the former chair of IFIP WG 8.2, "Information Systems and Organizations." His Erdos number is 3. He has published over 350 refereed articles and edited or written over 30 books or special issues. He conducts research that explores digital innovation especially in relation to the nature and organization of digital innovation, design work, requirements in large scale systems, diffusion and assimilation of digital innovations, and emergence of digital infrastructures.

Esko Penttinen (Ph.D., Information Systems Science, Helsinki School of Economics) is Professor of Practice in Information Systems Science at Aalto University School of Business, Helsinki, Finland. Esko leads the Real-Time Economy Competence Center and is chairman of XBRL Finland. His main research interests include adoption and economic implications of interorganizational information systems, meaningful work allocation between humans and computers, and governance issues related to cloud-based information systems. His work has appeared in outlets such as *Information Systems Journal, Journal of Information Technology, International Journal of Electronic Commerce,* and *Electronic Markets*.

MIS Quarterly

RESEARCH ARTICLE

HIGH RELIABILITY IN DIGITAL ORGANIZING: MINDLESSNESS, THE FRAME PROBLEM, AND DIGITAL OPERATIONS[1]

Antti Salovaara
Department of Computer Science, University of Helsinki, FI-00014 Helsinki, FINLAND, and
Department of Information and Service Economy, School of Business, Aalto University, FI-00076 Aalto, FINLAND
{antti.salovaara@alumni.helsinki.fi}

Kalle Lyytinen
The Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106-7235, U.S.A., and
Department of Information and Service Economy, School of Business, Aalto University, FI-00076 Aalto, FINLAND
{kalle@case.edu}

Esko Penttinen
Department of Information and Service Economy, School of Business, Aalto University, FI-00076 Aalto, FINLAND
{esko.penttinen@aalto.fi}

Appendix A

Software Operation Stages

F-Secure protects end-user computers in three stages, each providing a distinct type and level of protection. When a user's computer decides to retrieve content, reputation inspection checks whether F-Secure has listed the address as having been compromised. If the site is not listed, the retrieval command is allowed to pass through. The content retrieved gets examined in malware-sample-based detection, sometimes called the computer's firewall. If no fingerprints from malicious content match the data, the detection stage is passed: the data may enter the computer. Finally, behavior-monitoring continuously checks for suspicious actions within the computer and, if necessary, activates a removal mechanism. All these operations are automated and involve communications between the end-user client software and F-Secure's servers.

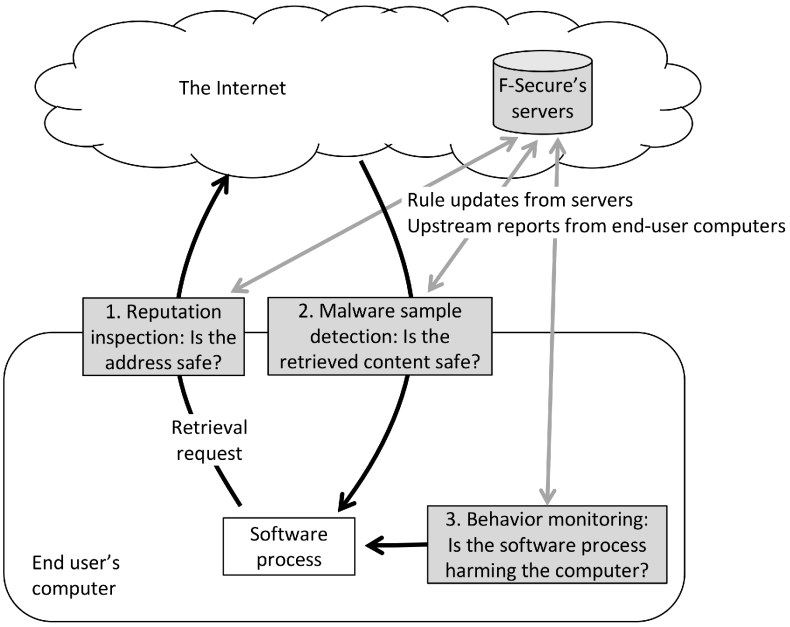

Figure A1. F-Secure's Digital Architecture of Malware Protection

Appendix B

Question Protocols

Background items for interviewee (used for each respondent):

- What is your education?
- Can you describe your prior work experience?
- What positions have you held at F-Secure?
- What are your current responsibilities at F-Secure?

Overall operations (items used mainly in Phase 1):

- Could you describe, in general terms and with adjectives, the environment where F-Secure operates (industry, type of operations, complexity of operations, and real-time operations)?
- Can you explain the sharing of fingerprints between security-software firms?
- Can you describe the process for sharing fingerprints?
- What kinds of systems do you employ?
- What kinds of contractual agreements do you have with other security-software firms?
- What are F-Secure's objectives? What outcomes does F-Secure try to bring about? What outcomes are you trying to avoid occurring?
- What are F-Secure's main hazards, and how important are they? Give examples.

Workflow (items used mainly in Phase 3):

Organization

- Can you describe the work processes, management, and allocation of tasks?
- How are the response unit's operations geographically managed 24/7?
- What kinds of incentives are in place?
- How independent are the employees and sites?

Response unit

- Can you describe timeline of changes within the response unit over the last 10 years?
- How do you manage day-to-day hand-over of tasks from shift to shift?

Division of work between the research and response units
- Can you elaborate on the breakdown by types of employee? IT vs. business?
- Can you describe the education background and heterogeneity of employees?

Systems
- Can you describe the methods and tools related to malware protection at F-Secure?
- Can you describe the analysis tools: which ones are unique to every employee and which ones are company-wide tools? Which tools are shared and which are personal?

Knowledge management
- Can you describe the knowledge-management practices at F-Secure: how knowledge is codified?
- What systems are used to store knowledge?
- How dependent knowledge management is on individual workers?
- How difficult it is to find expertise?
- Can you describe F-Secure's recruitment policies?

The rule engine
- How the rule engine works (e.g., on pseudo-code level)?
- Where the rule engine is located?
- How the workflow is organized around the rule engine?
- Can you clarify the concepts of Deep Guard and behavioral inspection systems?

F-Secure as a digital high-reliability organization (items used mainly in Phase 2):

"HOMEWORK": Could you give an example of a particularly complex/tricky case, one from your own experience within the past 12 months? (Interview note: anchoring a difficult case to be considered for all five processes)

Preoccupation with failure
- What kinds of failures/errors can occur in your operations? Could you give examples?
- How do you deal with these kinds of failures?
- How do you analyze them? (Tools, *post-mortem* analysis, etc.)
- What mechanisms do you use to prevent failures?
- Think of a situation in which identifying/recognizing a virus takes a long time. If it takes a long time, what does this mean? (Interview note: Could you cite an example in which you would have tried to generalize from a failure instead of localizing it?)
- Are you familiar with the concept "near miss"? What does that mean in the case of F-Secure?
- How do you analyze near misses?
- If it is not clear whether a sample is malicious, how do you make the determination?
- Do you encourage and reward reporting of errors? (Are there any punishments/incentives?)
- What mechanisms do you have for self-reporting of errors?
- F-Secure has received several awards and victories in head-to-head testing. How do you avoid becoming over-confident/too proud of your work? (Interview note: How do you avoid drifting into complacency, inattention, and habituated routines?)

Interview note: RETURN to particularly complex case:

**Can you tell me how you recognized the problem in that complex case, why it became a complex problem, and how the case changed the way F-Secure evaluates its operations?*

Reluctance to simplify interpretations
- When a new sample comes in, can you describe the process of analyzing it? (DRAW!)
- How easy is it to identify and classify what the sample involves?
- How easy is it to misclassify a sample? How do you avoid the temptation to simplify actions and make compromises?
- Do you have redundancy, such as cross-checks, or do you attempt to always question your levels of competence?
- Could you describe an example in which an employee was skeptical? Do you conduct adversarial (conflicting) reviews of your software frequently?
- Can you give an example of an actual false positive's analysis? Or a false negative's?
- Do you actively diagnose the limitations of preplanned procedures?
- Do you recruit for employees with non-typical prior experience?
- Describe teams' composition? What about the variety of expertise within a response team?
- Are you concerned about knowledge silos? Dependence on a single expert?
- How often do you rotate positions?
- What are the typical ways of retraining employees?

Interview note: RETURN to particularly complex case:

** Thinking about that complex case, could you tell how self-criticism, double-checking, or other means of avoiding simplification helped resolve the case?*

Under-specification of structures

- We are interested in the relative benefits of flexible and structured/pre-defined processes. Can you cite examples wherein flexibility has been critical? Can you cite examples of when clear structure has been critical?
- How do you see the evolution of flexibility and structure at F-Secure? What are the main sources of change in organizational structures?
- If you had to characterize F-Secure with the term "organized anarchy" or "structured hierarchy," how would you position F-Secure?
- What is its degree of flexibility in the rule engine? How do you plan for changes in the rule engine?
- How automated is the escalation process?
- If an employee takes a break (say, a year's maternity leave), how easy is it to return to the job? What is the pace of change?

Interview note: RETURN to particularly complex case:

** Could you elaborate on the complex case in terms of flexibility vs. structure (under-specification of structures)?*

Sensitivity to operations

- How would you describe the situational alertness of your team members at F-Secure?
- Can you give an example of a situation wherein an employee would have been in a "flow"/"bubble" (extreme situational alertness and responsiveness to unexpected inputs)?
- How often, if ever, do your team members achieve a state of "flow"? Are there different types of "flow" situations in which your team members reach a highly focused state?
- Can you describe some of the routine tasks within response teams?
- In your team's daily work, do you find similarities to organizations with high sensitivity to operations (for example, air-traffic control)? In what sense?
- How well are your team members aware of other team members' doings? Do you consider the employees to have shared mental representations (common understanding)?

Interview note: RETURN to particularly complex case:

** In the complex case, were there moments of "flow"? In what stages, and how was it manifested?*

Commitment to resilience

- Can you cite examples in which F-Secure would have shown resilience, coping with unanticipated dangers after they manifest themselves, learning to bounce back?
- Do you consider F-Secure capable of responding to threats independently without industry-community assistance?
- What if within-shift escalation is needed but not available? What happens?
- Have there been self-organized *ad-hoc* networks for expert problem-solving (dissolved later if these networks lacked formal status)?

Interview note: RETURN to particularly complex case:
** Can you elaborate on the complex case in terms of F-Secure's commitment to resilience and independent problem-solving? Did you establish* ad-hoc *networks to deal with the case?*

Concluding items

- What other aspects do you find important?
- What is specific to the operations of F-Secure?
- Can you elaborate on the community aspect? What characterizes the security software community?

Appendix C

Examples of Coding for HRO Characteristics and Operating Environment

Element/Code	Indicators	Example	No.
Operating environment	Type of business Malware volume Pace of change Time-criticality	On the pace of change: "In a game of cat-and-mouse, so we develop certain technologies to prevent this kind of infection from certain known malware within the week. Then the bad guys come up with new innovations to [counter] that."	63
Organizing operations	Maintaining tension between flexibility and structure Escalating cases Being reactive vs. proactive Sharing malware samples Avoiding false positives and false negatives	On reactive operations: "My impression of the on-call operations is that we monitor the samples we get from customers and add detections to them. It is purely reactive. Very often, detections are such that they only detect one sample. This, in turn, is a result of time-criticality: we must be able to process a large number of samples. [In on-call work], we don't have time to do thorough analysis and write sophisticated rules."	25
IT's role in F-Secure operations	Emergence of cloud operations The rule engine New protection technologies Dualism vs. duality between IT and human work	On the rule engine: "The rule engine categorizes the incoming malware samples and assigns them various actions. [This is based on] heuristics, clustering. automatically detecting samples [and placing them] into clusters. Algorithmic processes assign samples to [known malware] families. [Lowering] the rate of false positives is critical here. We take the algorithms from the academic world, the criteria from practice."	63
Preoccupation with failure	Being concerned about failure Analyzing near misses Encouraging reporting of errors	On reporting of errors: "We have a sync meeting. It's really not self-reporting of errors, but they ask a lot of questions. I encourage them to ask a lot of questions. So whenever they don't know, they ask. They have a tendency to keep on asking. If they keep on asking, then they don't have many errors because then they don't push forward [acritically]."	44
Reluctance to simplify interpretations	Considering multiple perspectives when making decisions Encouraging skepticism	On reluctance to simplify operations: "So we did that, and then there was no obvious clue something is still … there. But we had high suspicions that there was still something there. It did happen. Then we had to look more closely at what the malware is doing, at the level of really looking at how it actually connects with all the different files. And how it integrates with the system, how it is affecting the Internet connection to the system and all these things. That took us several days. Because whenever we look at a sample that deeply, it can take a while."	36
Under-specification of structure	Balancing between flexibility and structure	On flexibility of operations: "They write the code directly, and we have the rule-based mechanism, which means that we are very flexible. We can change our workflow in the same work day."	78

Element/Code	Indicators	Example	No.
Commitment to resilience	Establishing informal problem-solving teams Recovering from errors	On problem-solving: "There's a certain team that handles our front end, but this case was a special case since it was handled directly by our second-tier analysts and then it went even to our third level. We had to improvise solutions to help this reviewer install our product on their system. What happened in the end was that we had the solutions for the product, the fix, and then plan B for how we can disable the malware blocking our Web site as well as completely remove the traces of the malware so that we can install our product. [Plan B was needed], but in the end when our developer was able to fix that quickly enough and we gave it to our reviewer, luckily it solved the problem. Q: What number of people was involved? Three of our best guys [were] working on that. Q: Similar expertise? No, different expertise. On different levels: product experts, malware experts, then we had someone collaborating with the reviewer."	12
Sensitivity to operations	Having a heightened awareness of operations Generating a holistic picture of operations	On heightened awareness of operations: "if you do something [difficult] like reverse-engineering complex malware, which requires a high level of concentration. There might be heavy obfuscation, complex algorithms that you need to understand. Then, you need to get into a 'flow' state so that you can proceed." (2a)	11
Problem-solving cases	Using in-depth descriptions of difficult cases		13