Steen, R. (2019). On the Application of the Safety-II Concept in a Security Context. *European Journal for Security Research*. doi:10.1007/s41125-019-00041-0

On the application of the Safety-II concept in a security context

Riana Steen
BI Norwegian Business School

ABSTRACT

This paper presents an alternative and broader security risk perspective, incorporating uncertainty, as a two-dimensional combination of *i.* Threat (T_h) on value (V_l), *ii.* Vulnerability (Vu) given coping capabilities (Cc), and associated uncertainties U (will the threat scenario occur? and to what degree are we vulnerable?). Moreover, this work attempts to provide an integrated approach to the safety and security fields. We look closely into the issues related to Safety-I, Safety-II and security. Whereas conventional safety management approaches (Safety-I) are based on hindsight knowledge and risk assessments calculating historical data-based probabilities, the concept of Safety-II looks for ways to enhance the ability of organisations to be resilient in the sense that they recognise, adapt to and absorb disturbances. Three determinants that shape the Safety-II concept in the security perspective are the capacity of organisations to operate in changing circumstances; formulating strategies that promote a willingness to devote resources to security purposes, driven mainly by the organisation's leader; and an organisational culture that encourage people to speak up (respond), think creatively (anticipate), and act as mindful participants (monitor and learn). Based on clarifying some of the fundamental building blocks of security risk assessment, this work develops an extended security risk assessment, including an analysis of both vulnerability and resilience. The analysis explores how the system works following any type of threat scenarios and determines whether key functions and operations can be sustained.

Keywords: security, vulnerability, uncertainty, resilience, Safety-II

1 Introduction

Many researchers in the field of security management believe the safety concept does not capture all aspects of the security setting. Therefore, a distinction is made between these two contexts, based on the intentionality behind unwanted events, the way risk is understood and the methods used to assess and manage risk in these contexts. Various arguments reveal inadequacies in the application of conventional safety management in the security setting. One is related to the definition of risk. In the safety context, risk is commonly defined based on a combination of probability and the severity of the consequences (P,C), whereas probability is interpreted as an objective property of the activity being studied. According to objective interpretation, a probability expresses the relative fraction of times the event occurs if the situation analysed were hypothetically "repeated" an infinite number of times. The underlying probability is unknown, and the aim of the assessment is to estimate the proportion, p, of the population being studied, having a certain property. For example, a failure condition (risk

source), based on a sample from a population. From a statistical point of view, with a large sample, the estimation error would be negligible (Aven and Steen 2010).

Many scientific research studies and handbooks in the security field argue that objective probability (frequentist) does not capture all aspects of concern in a security context. For instance, consider a scenario in the wake of an insider threat. Two main aspects of an insider threat are the intention of the threat actor to conduct a malicious act, and the perpetrator's capability and competence for carrying out a malicious plan. Clearly, as we see in this example, using (P,C) has limitations in defining security risk. It is not meaningful to talk about fraction of times the event of interest (insider) occurs when we can repeat the situation over and over again (infinite number). Accordingly, a distinction is made between how risk is defined in the safety and security field. In the context of security, a risk perspective is proposed that excludes the probability aspect. This perspective is based on three factors: value (asset), threat and vulnerability (consequences) (Alberts et al. 1999; Landoll 2011; Masse et al. 2007; NS 5831, 2014).

Researches in the safety management field, on the other hand, have criticised the value, threat and vulnerability risk perspective for not considering uncertainty as a main component in the risk perspective. For instance, Amundrud et al. (2017) questioned this risk perspective by asking: "If risk is considered the triplet, value, threat and vulnerability, we may ask where is the uncertainty component?". This question leads us to the first research question in this paper:

RQ1: How do we incorporate uncertainty in the three factors risk perspective?

As an answer to this research question a new definition of security risk is provided, as the two-dimensional combination of threat on value and vulnerability given coping capabilities, and the associated uncertainties (will the threat scenario occur*?* and to what degree are we vulnerable?). Besides the definition of risk, there are other arguments related to insufficiencies in the application of safety management in a security setting. One concerns the risk management methodology. For instance, Jore et al. (2018) claim that "current risk management methodology for long-term security planning is insufficient to capture black swan events", in particular "during the risk-assessment phase" (Jore and Egeli, 2015). In the safety management field, Probabilistic Risk Assessment (PRA) is commonly used to analyse probabilities (objective) and risk, based on statistical estimation theory.
The other approaches that apply in risk assessment systems analysis methods are Fault Tree analysis and Event Tree analysis, where there is a lack of data for accurately predicting system performance (Aven 2011, p. 2). An underlying assumption in conventional risk assessment, referred to as 'Safety-I' by Hollnagel (2017), that are based on statistical methods is that data are available to predict the future performance of a system and systems are tractable. This means that how a system functions is known, and subsystem details and descriptions are uncomplicated, and systems do not change while they are being described (Hollnagel, 2006).

Along with a conventional PRA-based risk assessment (Safety-I), aiming to provide an accurate estimation of probability, there is a widely-accepted staple of the literature on risk assessments that are based on a broader risk perspective, where the uncertainty is considered as a main component of risk definition. These 'modified' risk assessment approaches emphasize the uncertainty assessment rather than probability estimation during the risk assessment process (Amundrud et al. 2017; Aven 2015; Bjerga et al. 2016; Zio and Aven 2011). Here, subjective probability provides likelihood dimensions, conditional on background knowledge and applies this to express uncertainty. The main idea is that the ambition of precise risk estimation has to be replaced by uncertainty assessments and characterisation. This type of probability interpretation usually applies in risk assessment processes in the security context, where the process itself consists of risk identification, risk analysis, and risk evaluation. While a conventional risk assessment (Safety-I) aims to provide accurate risk estimation, the objective of conducting a modified (security) risk assessment is to provide insights about threat scenarios, processes, activities and systems being analysed, and to reveal uncertainties and describe them. The insight provided by risk assessment is to support decision making by finding, prioritising and implementing risk-reduction measures.

Conducting a security risk assessment, such as for a modified risk assessment in the safety context, relies on the assumption that how the system is functioning is to "some extent" known. Nevertheless, due to the complexity and high level of uncertainty involved in the security context, for instance acts of terror or insider scenarios, which are caused by the interaction of multiple factors, this assumption may oversimplify the nature of the security threat. In contrast to safety context, a security threat is mostly rooted outside the organization and may also be untraceable (e.g., hackers). Moreover, details and descriptions of a security threat are often complicated, such as information about threat scenarios that are provided by intelligence services, which are more general than at the individual organizational level (Jore 2017). Consequently, in many cases organisations don't have sufficient knowledge and means to understand and reduce the threat. Thus, the identification of threat scenarios based on existing knowledge may not cover all types of scenarios. As an example, consider the insider threat, how does one assign a probability to such a threat when it comes from people within the organization who have inside intelligence about the organization's security practices? How does the risk analysis team take into account all relevant factors when the perpetrator could be intelligent, adaptable and strategic enough to adjust their performance? Depending on, for example the actor's motivation and capabilities, the actual threat evaluation could dramatically change, and as a result the risk analysis would give poor predictions.

Consequently, current approaches to security assessment are unable to deal with these types of threats. Surprises can occur relative to the background knowledge that the predictions and assessments are based on. In order to achieve a higher level of security, in many cases, it is not sufficient to rely only on the results of a risk assessment. We need to find a way to enhance organisation resilience to deal with any type of threats and do this under varying conditions. Hence the goal of conducting a risk assessment should not only be to provide insights about threat scenarios and systems vulnerability (what can go wrong and what are the

consequences), but also to enhance an organization's resilience to ensure that 'as many things as possible go right'.

Organization resilience is about having the capacity and ability to anticipate potential opportunities and threats, respond adequately to internal and external disturbances, and monitor critical success variables to manage organizational behavior, sustain operations, recover from disturbances, and exploit opportunities to build a desirable future (performing even better) (Steen and Tangenes 2017). This is what the alternative perspective on safety is about, namely 'Safety-II'. In the Safety-II approach, safety is defined as the ability to succeed under varying conditions. The focus in Safety-II is to find ways "to enhance the ability of organisations to be resilient in the sense that they recognise, adapt to and absorb disturbances." (Hollnagel et al. 2015).

The present paper addresses the application of the Safety-II concept in the security context by attempting to answer the second research question:

RQ2: What is the link between security risk and the Safety-II concept?

To find the link between the Safety-II concept and security risk, we examine the objectives of the Safety-II concept and the purpose of the security risk assessment simultaneously to identify areas of convergence. The results reveal that the three determinants shaping the Safety-II concept in the security perspective are the capacity of an organisation (human, institutional, physical and financial) to operate in changing circumstances; forming strategies that promote a willingness to devote resources for security purposes and an organisational culture that encourages people to speak up (respond), think creatively (anticipate), and act as mindful participants (monitor and learn). These capacities are the cornerstones of resilience engineering (Hollnagel 2006). Then, the question arises: How can resilience engineering be incorporated in security risk assessment (RQ3)?

RQ3: How can security risk assessment support basic ideas of resilience engineering?

In attempting to answer RQ3, we present a new security risk assessment framework that provides a structure for linking the concepts of security and resilience engineering. This framework is based on the above-mentioned new definition of a security risk perspective, where uncertainty is a main component of the risk definition. Moreover, the framework consists of both a vulnerability analysis, and also resilience analysis. The aim of conducting a vulnerability-resilience based risk assessment is twofold. Firstly, to provide insights about potential threat scenarios, causes and consequences, then to produce a risk picture based on the available background knowledge and identify key uncertainty factors. These elements are in line with the components of existing risk assessment frameworks. The novelty of this part is in how risk is defined, where here more focus is given on the resilience abilities (LARM) of a system (ability to Learn, Anticipate, Response and Monitor). The aim of this part is to secure that 'as few things as possible go wrong'. The second objective in utilizing the vulnerability-resilience based framework is to enhance the system's capacity to withstand any

type of threat and disruption, and to rapidly recover to the normal functionality of the system. The aim of this part is to ensure that 'as many things as possible go right'. This part corresponds with the Safety-II concept.

We argue that, due to the complexity of the security context, the security assessment must provide a broader risk picture, and consider uncertainties "hidden" in background knowledge about the threat scenarios and systems' vulnerability. We need to understand how the system functions in every detail and conduct risk assessment focusing on how the system works following any threat scenario. Can key functions and operations be sustained? Why? The extended processes in risk assessment ensure a broader perspective, link risk, vulnerability and resilience, and provide insights from different traditions and perspectives.

To answer the three research questions outlined above, this work applies an explorative qualitative research approach based on structured reviews of scientific literature. We conduct a search of relevant studies published from 2000 until 2018 using scientific databases (e.g., Academic Search Premier, BIBSYS, Oria and Google Scholar). The searches targeted the main concepts of this paper – security risk assessment, Safety-I and II, and resilience engineering. Other non-journal literature is also used to understand key issues and concepts in this work. Using a prison escape scenario as an example, we demonstrate how to understand the proposed security risk perspective and how different stages of security risk assessment could be carried out.

The remainder of this paper is organised as follows. Section 2 presents a case example, which will be referred to throughout the paper. Section 3 presents a theoretical basis of the main concepts used in this study, including security risk and security risk assessment, followed by a review of the main ideas in the Safety II concept. Section 4 presents an alternative security risk perspective, incorporating uncertainty as a main element of risk perspective. Section 5 suggests improved security risk assessment measures, highlighting issues from the Safety-II concept and presenting an extended risk assessment framework. Section 6 discusses the practical implications of the extended risk assessment framework with respect to our case example. Finally, Section 7 draws conclusions and provides recommendations for further research.

2. Case example

To illustrate the issues and discussion, we use the following case example, inspired by the true story of a South Carolina inmate serving a life sentence for kidnapping, who used a drone and a makeshift dummy to escape from prison in July 2017 (Levenson and Jones, 2017).

The prisoner escaped from a maximum-security prison using wire cutters flown in by an accomplice piloting a drone. The inmate used a cellphone to manage the drone's delivery. A more detailed illustration is provided in Figure 2.

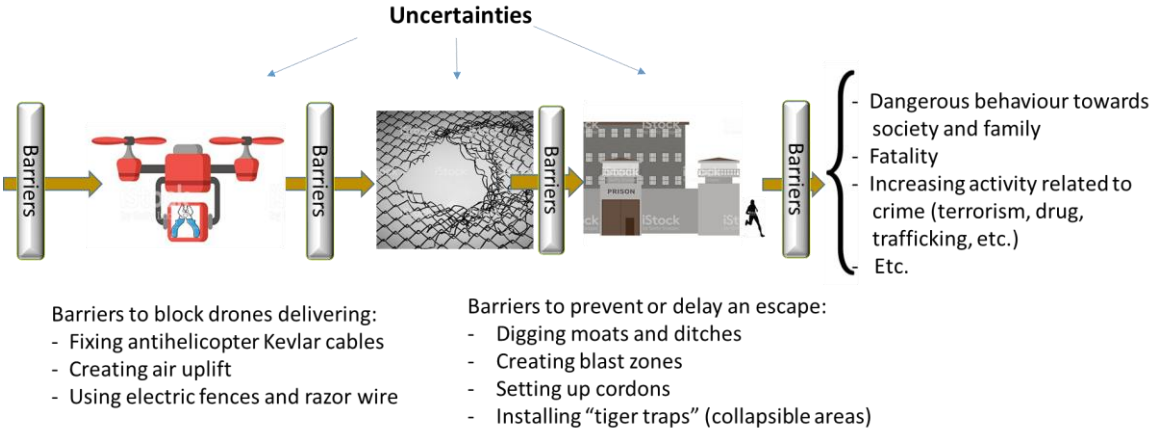


Figure 1 Illustration of key concepts in the case of a prison escape

In line with Norwegian standard NS 5831, threat (T_h), value (V_l) and vulnerability (V_u) are the components of security risk in this paper. Let us explain these three components in light of our case example. A value at stake (V_l) could be the secure custody of the prisoners. Related to this value, a threat scenario (T_h) is an "escape" by an inmate (the threat agent). An escape can lead to different consequences, for example, fatalities, increased organised crime, drugs, etc. The degree to which the prison is vulnerable (V_u) regarding an escape scenario depends on the functionality of different barriers in the prison; for example, the existence of air uplift, blast zones and a shield around the prison. It is beyond the scope of this paper to discuss the events surrounding the escape and all the barriers presented in Figure 2. However, some overall reflections are provided.

3. Theoretical background

As the main issues used in this paper are security risk, security risk assessment, Safety-II and resilience, the research is built on a theoretical background related to these issues. We start by reviewing the main components of a risk description in line with the three-factor model security risk perspective.

3.1 Introduction to the main components of the triplet security risk perspective

Value

Value, denoted as an asset, is tangible and non-tangible elements of a system that represents value to an organisation and is "subject to protection" (Beyerer and Geisler 2016). Value could be related to human life, human rights, health, environment, reputation, property, components, items or processes, functions and operations. According to Katsikas (2012) one way of expressing value is to use the example of the negative effects unwanted incidents can have on business interests. Such incidents could be disclosure, destruction and lack of availability.

Value could also be related to a system's functionality, as our case example (Figure 2) shows. Imprisonment serves primarily as a correctional service to protect society against crime, and to ensure execution of remand. To achieve this mission, there are four main goals in correction services: retribution, deterrence, incapacitation, and rehabilitation (Kifer et al. 2003). Incapacitation involves physically removing prisoners from the society they have offended or may jeopardise. In the process of incapacitation, a value at stake (V_1) could be the secure custody of prisoners.

Threat

Landoll (2011) defines threat as "an undesired event that may result in the loss, disclosure or damage to an organisational asset." The term is also referred to as a "risk source" in the Society for Risk Analysis Glossary (2018). According to Ojanen (2017) defining threat involves making decisions on what to work against. In the security context, a threat can lead to different consequences e.g., fatalities, environmental damage, reputational damage and economic loss. From a subjective school of thought, George (1986) considers "threat" to be an appraisal of a situation:

> " …an individual experiences in a situation that he perceives poses a severe threat to one or more of his values. Thus, perception of threat in the situation must occur in order for the individual to experience arousal of anxiety or fear. Threat, in other words, is not an attribute of the stimulus situation; it depends on the subject's appraisal of the implications of the situation for his values."

George's explanation is based on the thesis that "perception of threat" is a subjective judgement that one may make about the characteristics and severity of a risk source. In the security context, in a prison for instance, threat could be an "escape" posed by an inmate (the threat agent). Threat of escape is subject to uncertainty. An escape is an event that may happen in future. Three main components of an inmate's escape are the motivation, means and opportunity. As these components are subject to uncertainty, we do not know for sure the likelihood of the inmate escaping from prison (means), the reason for escaping (motive) and whether he has a chance to escape (opportunity). Accordingly, the judgment is conditional on a set of assumptions and underlying factors. We can use probabilities as a means to express the uncertainty associated with various threat scenarios. It is important to emphasise that probability is just a tool used to articulate uncertainties. In a security context, strong knowledge about the probabilities does not exist and we often cannot obtain accurate estimates. In this regard, degree of belief (DoB) can be used as an approach to express uncertainty (Beyerer and Geisler 2016; Häring et al. 2016).

Vulnerability

Vulnerability implies the level of risk an organisation or community faces regarding specific threats to their values. Aven and Renn (2010) refer to vulnerability as the "quality of the risk absorbing system" to withstand or tolerate different degrees of the threat agent to which it may be exposed. Landoll (2011) describes vulnerability as "a flaw or oversight in an existing

control that may possibly allow a threat agent to exploit it to gain unauthorized access to organizational assets." Scholars have classified an organisation's vulnerability in different groups. For instance, in the Security Risk Assessment Handbook, vulnerability is categorised into three main groups; administrative (gaps in policies, procedures or security activities), physical vulnerability (gaps in physical, geographical, personnel or utility provisioning controls), and technical vulnerability (gaps in logical controls in the organisation's system; for example, weak passwords) (Landoll, 2011).

In measuring the degree of vulnerability, scholars have developed different methods and metrics that incorporate variables such as sensitivity to stress or changes, probability of being exposed to stress, restoration time, access to entitlements, threat impacts and communication and physical connectivity, (Adger 2006; Dwyer et al. 2004; Rajesh etal. 2018). Adger (2006) defines vulnerability as "the state of susceptibility to harm from exposure to stresses associated with environmental and social change and from the absence of capacity to adapt." He points to the three key parameters of vulnerability: the threat to which a system is exposed, its sensitivity, and its adaptive capacity. The degree to which a system is vulnerable to threats depends on the system's ability to absorb the shocks, the autonomy of self-organisation, and the ability to adapt both in advance and in reaction to threats. Fraser and Greenhalgh (2001) describe capability as the "extent to which individuals can adapt to change, generate new knowledge and continue to improve their performance" (p. 799). Adaptation is about having human, technical, institutional and structural capabilities to intervene in vulnerability (Birkmann et al. 2013). For instance, in our example, the likelihood of an inmate's escape serves to make the prison vulnerable, but the degree of likelihood depends on the prison authorities' responses to this type of threat, namely to escape, assisted by a drone. The authorities' responses, on the other hand, are shaped by a subjective evaluation of the threat and having the capabilities to anticipate, monitor and cope with it. There is also uncertainty about the prison authorities' ability to cope with the threat.

3.2 Security risk assessment

The aim of conducting security risk assessments is to provide insights about the threat phenomena, processes, activities and vulnerability of the system being analysed. By identifying threats, studying their causes and consequences and describing risk, decision-makers are informed about the risk level and main contributors to risk. In this way, security risk assessment's main function is to support decision-making on how to respond to threats. There are a number of security risk assessment frameworks with the same basic elements in conducting the assessment, namely asset valuation, threat analysis, vulnerability analyses and security risk evaluation (Landoll 2011). Examples include:

- OCTAVE framework: Operationally Critical Threat, Asset and Vulnerability Evaluation risk assessment framework. (Alberts 2002)
- ISO (27005) Standard on security management (International Organization for Standardization 2018)

The OCTAVE framework is based on the triplet factors risk perspective (threat, asset and vulnerability). The risk assessment process is built on three phases. The first focuses on identifying assets and their value, threats to those assets, and identifying security requirements based on knowledge from staff at multiple levels within an organisation, along with standard catalogues of information. The second phase concentrates on the identified threat scenarios and evaluating vulnerability. The results of the second phase provide insights about "the high-priority information infrastructure components, missing policies and practices". The third phase is founded on the previous phases and estimates the "impact and probability of the risks". The third phase results help develop a protection strategy and establish a plan to manage security risk. (p. 4-5)

Risk assessment in ISO (27005) is one of the main parts of this risk management framework, founded on the triplet factors risk perspective. It consists of the following processes; risk identification (including identification of assets, threats, existing controls, vulnerabilities and consequences), risk analysis (including assessment of consequences, likelihood of incidence and level of risk determination) and risk evaluation. Risk assessment helps find proper risk treatment options and produce a risk treatment plan. (p. 8-12)

3.3 Safety II concept

In the traditional safety management approaches (Safety-I), safety is defined as a condition where "as few things as possible go wrong". In the Safety-II concept, safety is understood as a situation in which "as many things as possible go right (under varying conditions)" (Hollnagel et al. 2015). The underlying idea of the Safety-II concept is that "we cannot make things go right simply by preventing them from going wrong" (Hollnagel 2016a). In the Safety-II perspective, safety "is about how to support, augment and facilitate the everyday activities that are necessary for acceptable outcomes on all levels of an organisation" (Hollnagel 2017). Whereas the risk assessment process in the Safety-I perspective focuses on accidents caused by failures and malfunctions, and aim to identify causes and contributory factors, risk assessment in Safety-II aims to understand the conditions in which performance variability can become difficult or impossible to monitor and control (Hollnagel et al. 2015).

Aven (2016) suggests that "the way we understand and describe risk strongly influences the way risk is analysed, and hence it may have serious implications for risk management." Accordingly, changes in the definition of safety affect safety management approaches and the application of methods and techniques. Safety-I approaches aim to avoid things going wrong and focus on hindsight knowledge, failure in reporting, and risk assessments to calculate historical data-based probabilities. In contrast, Safety-II tools and concepts aim to enhance organisations' resilience. Safety-II focuses on eliminating hazards, preventing failures and malfunctions, and developing an organisation's potential for resilient performance in the sense that the organisation recognises, adapts to and absorbs variations, changes, disturbances, disruptions and surprises (Hollnagel 2017). A simple illustration of the Safety-II school of thought is presented in Figure 3.

Figure 2 Safety-II as a combination of Safety-I and resilience

The Safety-II concept is based on a "mechanisms that allow for the detection of the variability, for the understanding of its potentially surprising nature or scope, and for the timely reconfiguration of the system to manage it successfully." (Hollnagel 2014) Rather than anticipating and responding to some specific unwanted events, Safety-II's agenda is the improvement of *resilience* abilities in an organisation (or system) by managing "a system's adaptive capacity based on empirical evidence" (ibid). There are many definitions of resilience, which vary according to the discipline and scientific field. In the security context, it is defined as "the ability to withstand and recover from deliberate attacks, accidents or naturally occurring threats or incidents" (Masse and Rollins 2007). Examples of resilience measures include developing a business continuity plan, having a generator for back-up power and using building materials that are more durable.

An approach for developing resilience in an organisation or system is Resilience Engineering (RE). Hollnagel (2006) provides a summary of the RE approach, and the basic premises and features of the field. For the purpose of the present paper it is sufficient to draw attention to a few key points (p. 10-14):

- Many adverse events cannot be attributed to a breakdown or malfunctioning of components and normal system functions (intractable events). They are best understood as the result of unexpected combinations of normal performance variability.
- Effective safety management cannot be based on hindsight, nor rely on error tabulation and the calculation of failure probabilities. Safety management must not only be reactive, but also proactive.
- The conventional view on safety management considers performance variability of any kind as a threat to be avoided, and to counter such threat, constraining measures such as barriers, interlocks, rules and procedures are employed.
- In resilience engineering, performance variability is considered normal and necessary. Safety cannot be obtained by constraining performance variability, since that would also affect the ability to achieve desired outcomes. Instead, the solution is to dampen variability that may lead to negative outcomes and simultaneously reinforce variability that may lead to positive outcomes.

To be resilient, a system or organisation must meet the four potentials (Hollnagel, 2011):

I) Knowing what to expect; that is, how to *anticipate* future developments, threats and opportunities, such as potential changes, disruptions, pressures and their consequences. This is the ability to address the potential.
II) Knowing what to look for; that is, how to *monitor* that which is or can become a threat in the near term. This is the ability to address the critical.
III) Knowing what to do; that is, how to *respond* to regular and irregular disruptions and disturbances. This is the ability to address the actual.
IV) Knowing what has happened; that is, how to *learn* from experience. This is the ability to address the factual.

These four potentials are interdependent and are known as the four cornerstones of resilience engineering. (Hollnagel 2011). A systemic view could be used to describe and analyse a system's resilience potentials. The significant characteristic of the systemic view is that it considers an organisation as "a multi-minded, sociocultural system, a voluntary association of purposeful members who have come together to serve themselves by serving a need in the environment" (Gharajedaghi 2011). The Functional Resonance Analysis Method (FRAM) is an example of the application of the systemic view approach in the Safety-II context. The FRAM provides a way to describe outcomes of different functions in an organisation/system by focusing on the relationship between different functions in a system, where a function describes an activity (work-as-done) in an actual work situation. In the FRAM approach a function describes "the means that are necessary to achieve a goal" (Hollnagel 2012, p. 39). Different functions in an organisation represent operational, technical and organisational activities. For instance, in the context of security in a prison, different functions could be incident reporting, interstate and international transfer of prisoners and control activities. According to Hollnagel (2012, p. 39), implementing the FRAM approach involves two stages. First, providing a model of the different functions, which is the focus of the analysis. Then, in the second stage, the model from the first stage is used to analyse how these functions are interrelated with regard to the six aspects of: Input (I: that which the function processes or transforms or what it starts with), Output (O: the result of the function), Preconditions (P: that must exist before a function can be executed), Resources (R: that which the function needs to produce the output), Time (T: related to starting time, finishing time or duration) and Control (C: how the function is monitored or controlled. The results generated in output can represent both a change of state in the system or in one or more aspects of other functions (I, P, R, T or C), and a decision or a signal that starts a downstream function (Bellini et al. 2017).

When all of the functions in the system under study are identified and linked, in order to assess the resilience of a system related to identified couplings, resilience-enhancing measures (4R), introduced by Bruneau et al. (2003), could be applied. These four properties are as follow:

- R1: Robustness refers to the strength of elements in systems to withstand a given level of stress or demand without suffering degradation or loss of function.

- R2: Rapidity is the capacity to meet priorities and achieve goals in a timely manner to contain losses and avoid future disruption; for example, quick access to sources of financing to support recovery.
- R3: Redundancy refers to the availability of substitutable elements or systems that can be activated when earthquake-related disruptions occur; for example, having many evacuation routes.
- R4: Resourcefulness is the capacity to mobilise and apply material and human resources to achieve goals in the event of disruptions

According to Bruneau et al. (2003), while robustness and rapidity could be considered the desired results of resilience-enhancing measures, redundancy and resourcefulness could be taken into account as means to these ends.

4. Incorporating uncertainty in the triplet security risk perspective

This section addresses the first research question: How can uncertainty be incorporated in the three-factors risk perspective? In Section 3.1 we referred to our case example and explained how the main components in the triplet security risk perspective, threat (T_h), value (V_l) and vulnerability (V_u), could be understood. We also made some comments on how uncertainty is involved in making a judgment about the threat and vulnerability level.

Threat is always attached to some sort of value at stake, which means that without considering what the subject of protection (value) is, threat makes no sense. This is why these two factors should be treated together as "threat given value" ($T_h|V_l$). Moreover, the vulnerability of a system related to a threat to a given value depends on how that system responds to the threat (see Section 3.1). The form and quality of response relies on the significance and effectiveness of the organisation's coping capabilities (C_c). Coping capabilities could comprise three aspects: individual (staff's explicit and tacit knowledge, judgment skills, experience, and abilities); theological (an information-gathering and reporting system, a monitoring system) and the organisational aspect (culture, communication, compliance, etc.). From these arguments, we identify two main dimensions of security risk:

$$[(T_h|V_l), (V_u|C_c)] \qquad (1)$$

Beside the threat towards values and vulnerability given coping capabilities, the uncertainty involved with a threat (is the threat real? how serious is it?), and the ability to cope with the threat (could we deal with the threat, given our current resources?) will affect the consequences of being exposed to some threats. In this way, uncertainty affects vulnerability. Consequently, as vulnerability is one of the main components of security risk, the level of uncertainty has a direct effect on security risk. Incorporating uncertainty as an influencing factor on security risk in our model leads us to the following setting:

$$\text{Security risk: } [(T_h|V_l), (V_u|C_c), u] \qquad (2)$$

Based on Setting *2*, by security risk we understand the two-dimensional combination of

Threat *(T_h)* on value *(V_l)* and Vulnerability *(V_u)* given coping capabilities *(C_c)*, and the associated uncertainties (*u)* (will the threat scenario occur*?* and to what degree are we vulnerable?).

This risk perspective could be understood as follows: When a threat actor (an event or some sort of risk source that may cause a threat to arise) intentionally exploits vulnerability in an organisation (person, process, system) and causes harm to the value at stake, security risk is realised. The extent to which the system is vulnerable depends on its coping capabilities, see Figure 2.

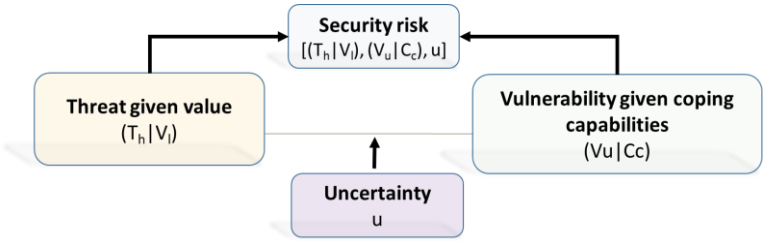


Figure 3 Formative conceptual security risk perspective

As Figure 4 depicts, both threat and vulnerability are subject to uncertainty through its moderating effects, which are generally considered as an interaction between factors or variables, where the effects of one variable depend on levels of the other variable in analysis (Fairchild and MacKinnon 2009). Uncertainty as a moderating factor in this work means it affects the strength of the relation between a predictor (threat and vulnerability) and the level of outcome, security risk. The higher the level of uncertainty, the higher the level of risk. Aven (2014) defines uncertainty as a lack of knowledge, i.e., "not knowing about something, where 'something' refers to the true value of a quantity or the true future consequences of an activity (p. 235)." There are many sources of uncertainty, including the subjectivity of the analyst's judgments and linguistic ambiguity (Zio and Aven 2011). De Berker et al. (2016) classified uncertainty into three categories: the randomness inherent in any complex environment, imperfect knowledge of the relationships between predictors and outcomes and volatility uncertainty, which is about the stability of the context.

Different assessment tools can be applied to reflect uncertainties about threat scenarios and a system's vulnerability given these identified scenarios. Examples include historical data-based probability assessment and expected values, as well as judgement (knowledge) based probability (subjective) such as Degree of Belief, ranking threats and vulnerability in terms of likelihood and potential impact. Bayesian analysis may also be used to update subjective probabilities to formally incorporate new information. Bayesian analysis is done in two steps (Aven 2014); assigning a prior distribution on the parameters of interest, and applying Bayes' theorem to establish the posterior distribution of the parameters (p. 110). The reference in a subjective probability assessment is a certain standard such as drawing a ball from an urn. For instance, if we assign a probability of 0.2 for a threat scenario 's', we compare our uncertainty of 's' occurring with the draw of two favourable balls from an urn having eight unfavourable balls. Following this perspective, true probabilities do not exist as there is no reference to a true fraction that is unknown (Bjerga et al. 2016).

Various types of a risk matrix can be applied to represent a risk picture, to rate threat scenarios, vulnerability, and to determine the need for risk reduction. An example of a risk matrix, suitable for a [($T_h|V_l$), ($V_u|C_c$), u] risk perspective is presented in Figure 5.

		Threat on value						Vulnerability given coping capabilities				
		Insignificant	Negligible	Moderate	Extensive	Significant		Insignificant	Negligible	Moderate	Extensive	Significant
Likelihood	Rare	Low	Low	Low	Medium	Medium		Low	Low	Low	Medium	Medium
	Unlikely	Low	Low	Medium	Medium	Medium		Low	Low	Medium	Medium	Medium
	Possible	Low	Medium	Medium	Medium	High		Low	Medium	Medium	Medium	High
	likely	Medium	Medium	Medium	High	High		Medium	Medium	Medium	High	High
	Almost certain	Medium	Medium	High	High	High		Medium	Medium	High	High	High

Figure 4 An example of security risk matrix

The likelihood dimension indicates the state of being likely or probable and is divided into the five levels: almost certain (the first level), when the threat scenario is expected to occur in most circumstances, and the fifth level, rare, which reflects that the threat scenario may occur only in exceptional circumstances. The combination of threat on value and associated uncertainties provides a threat score (low to high level). Vulnerability given coping capabilities also has a score from a low to high level. The level of threat and vulnerability can be assigned by obtaining a numerical value from zero to ten. The higher score implies a higher threat. For each identified threat scenario, a security risk description covers threat, vulnerability (the impact of the event to occur, given coping capabilities), and associated uncertainties.

Subjective probabilities, as for objective probabilities are tools. A tool has constraints. Its value is comparable with the extent of the risk analyst's ability to determine, characterise and analyse uncertainty factors involved with threat scenarios. However, black swan events may happen with different types, such as those that are identified in possible threat scenarios, "but whose probability of occurrence are judged negligible, and thus are not believed to occur"; and those that are completely unknown (Flage and Aven 2015). Nonetheless, risk assessment is not able to capture black swan and unknown-unknown (events that we don't know that we don't know) types of events. As pointed out by Taleb (2007), a black swan event is "is an outlier" and "nothing in the past can convincingly point to its possibility" (p. xvii). Thus, it is impossible to predict these types of threats that an organization may face. In order to deal with unforeseen events and (potential) surprises it is necessary to see beyond risk analyses. We need to find a way to enhance an organisation's resilience to deal with any type of threats and under varying conditions.

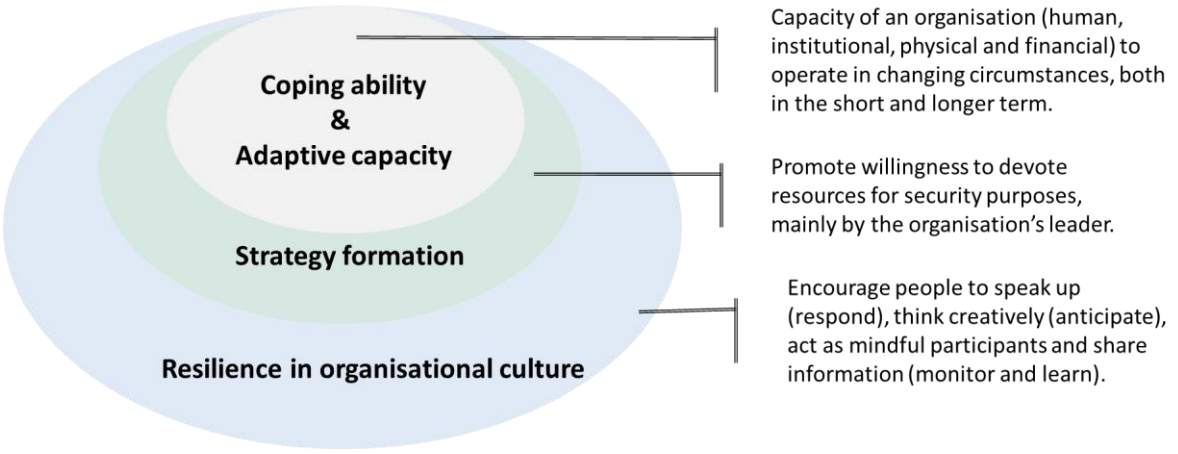
5. What is the link between the definition of security risk and the Safety-II concept?

As discussed in Section 3.2, a crucial step in assessing security risk is identifying potential threats, which is influenced largely by an organisation's ability to make sense of what is going on and to "understand issues or events that are novel, ambiguous, or in some other way violate expectations" (Maitlis and Christianson 2014). In the security context, sense making is a crucial process because "it is the primary site where meanings materialise that inform and constrain identity and action" (Weick et al. 2005). In the Safety-II concept this is known as the ability to monitor (see Section 3.3). According to Hollnagel (2011) "monitoring must cover both what happens in the environment, outside the organisation (system) and that which happens in the organisation itself, i.e., its own performance and functions." (p.279)

The extent to which an organisation is able to monitor and address critical issues and respond appropriately to minimise the impact of undesirable events is highly dependent on having *organisational capacity* in the form of human, institutional, physical and financial resources to operate in changing circumstances, both in the short term (coping ability) and longer term (adaptive capacity) (Turner et al. 2003). Moreover, an organisation's attention structure affects senior managers, operational managers and the staff's attention to changes through so-called attention regulators. One important regulator is the "rules of the game" (Ocasio 1997), which relates to organisational culture, *strategy formation*, and formal and informal principles of action, interaction, and interpretation that guide organisational behaviour in the security context.

Mintzberg (1978) defines strategy formation in a descriptive context as "a pattern in a stream of decisions". As an abstraction of decisions and actions in an organisation, strategy formation is interwoven with *organisational culture* because it affects the ability of an organisation to learn from experience and respond to disruptions and change (Sutcliffe and Vogus 2003; Tangenes and Steen 2017; Weick 2016). Both organisational culture and strategy content can be condensed to hypothesize the cause and effect that provide behavioural guidelines for organisational actions in the quest for security.

From the discussion above we can conclude that the three determinants that shape the Safety-II concept in the security perspective are: the capacity of organisation (human, institutional, physical, and financial) to operate in changing circumstances; strategy formation that promotes a willingness to devote resources for security purposes, driven mainly by the organisation's leader; and an organisational culture that encourages people to speak up (respond), think creatively (anticipate) and act as mindful participants (monitor and learn). Figure 6 illustrates these three determinants and their coupling to each other.



Figur 5 The three determinants that shape the Safety-II concept in the security perspective

We use Figure 6 to prepare an answer to our second research question: What is the link between security risk and the Safety-II concept? Section 3.3 presented the objective and main steps of the security risk assessment process and how the concept of Safety-I, resilience and Safety-II are connected (Figure 3). On the one hand, the security risk assessment objective is to support decision-making on how to respond to undesirable events in each identified threat scenario and what risk reduction/prevention measures to choose. However, to move toward the Safety-II paradigm, we need to enhance the system's capacity for resilience (robustness and rapidity) (R1 and R2, Section 3.3) to withstand and recover from disturbance and incidents. Enhancing organisation capacity to detect and respond appropriately to security incidents is the common objective in both the Safety-II concept and security risk assessment. Increasing resilience potentials in an organisation affects redundancy (R3) and resourcefulness (R4) in a system (Section 3.3), and consequently will increase coping capability and adaptive capacity. Improving coping capabilities will reduce the vulnerability of the system to a threat scenario, hence the security risk.

6. How can security risk assessment support the basic ideas of the Safety-II concept?

While existing security risk assessment frameworks include vulnerability, they lack a focus on resilience in line with the Safety-II concept. To ensure that the risk assessment includes the Safety-II dimension, we added resilience analysis in the risk-assessment process. We refer to such risk assessments as "vulnerability-resilience based risk assessment". The aim of the suggested risk assessment is twofold. Firstly, to provide insights about the phenomena, processes, activities and systems being analysed, and to reveal uncertainties and describe them. Secondly, to enhance the system's capacity to withstand any type of threat and disruption and to rapidly recover to the normal functionality of the system. The aim of this part is to ensure that 'as many things as possible go right'. This part corresponds with the Safety-II concept. Figure 7 summarises the main elements of the vulnerability-resilience based risk assessment. Note that the term "root cause" should be interpreted as the threat conditions or risk source (see Section 3.2) that lead to a specific outcome. This recognition

requires enlargements in the basic design of root analysis, including the capacity to treat (means, motive and opportunity) coupled with human–environment systems and those linkages within and outside the systems that affect systems vulnerability.

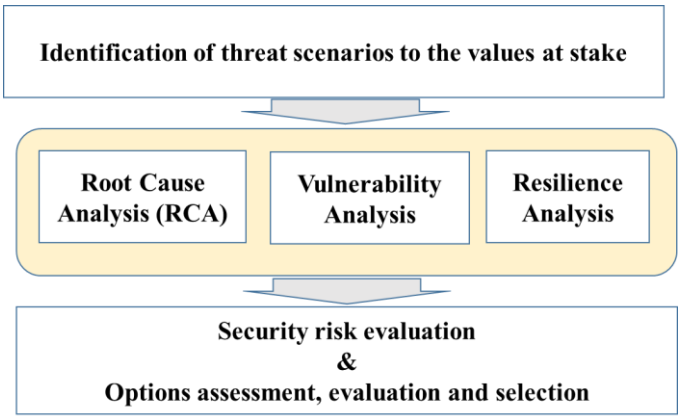


Figure 6 Main elements of vulnerability-resilience based risk assessment framework

As Figure 7 illustrates, the framework comprises three main phases: "scenario identification", "analysis" and "the evaluation and option section". The detailed sequence of phases and steps are as follows:

Identification of threat scenarios to the values at stake:

- Identify and evaluate asset (values)
- Continuously anticipate and monitor the security-related behaviour of the different components of the systems
- Characterise the threats and associated uncertainties on values

Root Cause Analysis:

- Identify possible causes of each threat scenario (means, motive and opportunity)
- Collect and analyse data

Vulnerability analysis:

- Identify the vulnerable impact area, given each threat scenario
- Security risk and consequence analysis: Analyse the cause-and-effect links through the entire system under study
- Analyse the cause-and-effect links beyond the system under study

Resilience analysis:

- Identify different functions and interdependencies among them
- Identify uncertainties that arise in interdependencies
- Investigate how the system works following any type of variation in security-related behaviour
- Perform functionality analysis: Can key functions and operations be sustained?

Security risk evaluation, options assessment and selection:

- Using objective criteria, analyse the security risk and consequences, given the resilience of the system
- Identify available security risk treatment measures to improve resilience
- Identify constraints (i.e., budget, time and organisational resources)
- Assess the options and select the best strategies

What follows is a closer look at the key components of the [(T_h|V_l), (V_u|C_c), u] security risk perspective and the different steps in the extended security risk assessment presented above, considering our case example (Section 2).

7. DISCUSSION

The issue we raise in the present section is the extent to which a [(T_h|V_l), (V_u|C_c), u] type of security risk perspective and the use of extended security risk assessments as described in the previous section can support the process of meeting resilience engineering potentials (i.e., the ability to anticipate, monitor, respond and learn), which are the main pillars of the Safety-II concept. We start by explaining the various elements of an extended risk analysis regarding our case example.

7.1 Identification of threat scenarios to the values at stake:

As mentioned earlier, threat is attached to the value at stake, so a priority in assessing risk is to provide an estimation of organisational values (assets) and to identify possible threats to each asset. Our threat scenario (scenario s) is an inmate using wire cutters dropped from a drone as part of an elaborate escape plan. Uncertainty prevails in this plan because the inmate needs many available resources, such as access to a cell phone smuggled into prison. How likely is that? Different techniques – including historical data – are used to identify threats. However, a historical data-based approach would not be sufficient (particularly in our case) as it is unlikely to cover all the relevant events. Hollnagel (2006) agrees, saying that anticipating what may happen (potential I, Section 3.3) "must go beyond the classical risk assessment, and consider not only individual events but also how they may combine and affect each other". This is associated with some uncertainty, but failing to think ahead will inevitably leave a system, in our case prison authorities, unprepared, and hence more vulnerable. To compensate for a lack of data, we need to adopt alternative approaches, such as Degree of Belief (DoB), brainstorming and Delphi-type exercises. An important task is to be creative and come up with scenarios that have not happened before but which are plausible.

This risk assessment element (6.1) obviously relates to four resilience potentials (criterion I-IV, Section 3.3). Having the ability to anticipate risk events in the long run (I) may enhance prison authorities' capacity to be prepared, and hence be less vulnerable. The ability to monitor (II) what is going on and address the critical issues has a direct effect on surveillance in prison, which requires having available resources (e.g., closed circuit television (CCTV)). It is also strongly related to criterion III, the ability to respond to actual threats such as an escape attempt. Regarding criterion IV, the ability to learn from experience and address the factual issues, Hollnagel and Speziali (2008) emphasise that "Learning requires more than collecting data from accidents, incidents, and near-misses or building up a company-wide database. Some organisations unfortunately seem to confuse data with experience. But whereas data are relatively easy to amass and can be collected more or less as a routine or procedure, experience requires the investment of considerable effort and time in a more or

less continuous fashion" (p.11). Conducting a broad approach to this analysis step (6.1), as indicated above, can improve the basis for learning from experience, revealing relevant threats and escape scenarios, which can be followed up in the vulnerability and resilience analyses.

7.2 Root cause analysis

Root cause analysis (RCA) is closely related to Step 1: identifying threats and escape scenarios. In this paper we are concerned mainly with the resilience dimension – an escape attempt may occur and we need to be able to sustain the functioning of the system (prison). Many different techniques are used in RCA, such as barrier analysis, which focuses on what control functions exist in the system to detect or prevent an unwanted event (escape attempt), and which might have failed. Change analysis is another approach that could be applied in RCA, when a system's performance has been subject to major change. This analysis explores variations made in organisation/systems (e.g., processes, technologies, information and people) and technology. In this regard, drone technology and its availability could be a related subject in our example. The other technique is fault trees and influence diagrams, which provide system insights enabling uncertainty/probability/frequency indices to be computed (see Aven 2011). These indices can be conducted quantitatively, based on data for the system being studied. However, in the security context, data are often lacking and expert judgments and lay knowledge are required. Several methods have been developed to account for organisational factors, see e.g., I-Risk (Papazoglou et al. 2003), the SAM approach (Murphy and Paté-Cornell, 1996) and the Hybrid Causal Logic Method (Mohaghegh, Kazemi, and Mosleh, 2009).

The root cause analysis element relates to criterion II (the ability to monitor what is going on), and to criterion I (the ability to anticipate risk events and opportunities in the long run). As the RCA may reveal new threat scenarios, it enhances the ability to respond (criterion III) in a more timely and effective manner.

7.3 Vulnerability (consequence) analysis given the occurrence of scenario "s"

An escape event (scenario "s") can lead to different consequences, C, with respect to violence, fatalities and environmental damage, for example, depending on the existence of barriers and their effectiveness. Examples of barriers in our case are air uplift, parallel control operating systems, and detection and warning systems. The barriers and system performance in general are influenced by a number of performance influencing factors (PIFs), such as technical qualities, the competence of the operators, procedures and time pressure. Many different techniques are used to analyse vulnerabilities (see also the methods mentioned in Section 3.1). These techniques provide system insights and informative indices. For a system in operation as in prison, indices (indicators) can be defined reflecting the system's operational features, such as alarms of different types.

The vulnerability element relates to criterion III (ability to respond) through barrier analysis and criterion II (ability to monitor what is going on) through the barrier and consequences indices (indicators) used. To some extent vulnerability also relates to criterion I (ability to

anticipate risk events and opportunities), as the consequence analysis may reveal such vulnerability. The vulnerability analysis also depends on the ability to learn by experience (criterion IV). For example, even though the prison in our example didn't experience a drone-assisted escape attempt, it is possible the prison authorities will use the experience from other prison escape situations, learn from them, and increase the ability to deal with such an escape scenario.

7.4 Resilience analysis

Whereas vulnerability analysis studies the performance of a system given a specific event "A" based on scenario "s", resilience analysis investigates the system without specifying events. The analysis focuses on how the system works following any type of variation, for example different types of escape attempts, and raises the question of whether key functions and operations can be sustained. The FRAM approach (see Section 3.3) can be used to conduct resilience analysis. In the first step, we need to identify the essential functions of the system, then study the interdependencies (couplings) among the functions through so-called functional resonance. In the second step we should characterize the variability of each function and assess how the variability of multiple functions can be coupled.

In our example a function could refer to something the prison as an organisation does: for example, it is the function of an emergency department to respond to an emergency situation. It could also be what a technical system (for example a warning system) in prison does either by itself (an anti-drone fence) or together with staff (monitoring the CCTV). Figure 8 illustrates how the FRAM approach could be utilised in our case example.

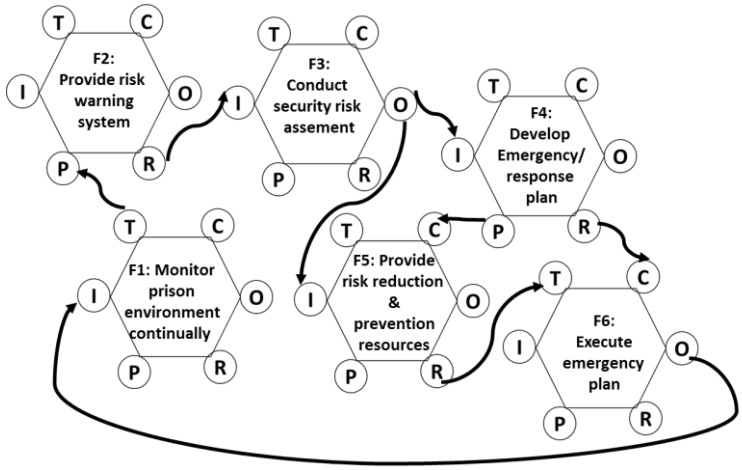


Figure 7 Application of the FRAM approach in an escape from prison – case example

It is important to mention that the illustration (Figure 8) is just an example and includes just a few functions. In an actual case study, a FRAM model should be developed further by considering each function and its interdependencies in detail. The FRAM model illustrated in Figure 8 started with function (3), conducting security risk assessment, and improved by several iterations. Through its development, different versions of the FRAM for our case

example were developed. Table 1 presents a short overview of how to interpret the couplings described in our FRAM model, considering coupling between F3, F5 and F6, and critical questions to ask regarding each coupling.

Table 1 Coupling between functions, comments and critical questions

Coupling	***Comments***	***Critical Questions****
F5 (R) – F6 (T) ***F3 (O) – F5 (I)***	*In order to execute an emergency plan (F6) in a timely manner and respond effectively when incidents occur (T), it is necessary to have resources available (F5-R). The scope of the response depends on the type of incident/event. The answer should be provided as output from function 3(F3-O).*	*-What are the events for which the system has a prepared response?* *- How fast can the system respond?* *- How many resources are allocated to ensure response readiness (people, equipment, materials)? How many people are available for the response potential? Who is responsible for maintaining the ability to respond?* *-What is the criterion for ending the response and returning to a "normal" state?*

**Questions derived from RAG – Resilience Analysis Grid (Hollnagel 2011)*

The answers to the critical question in Table 1 could be used to evaluate the functionality of each aspect in the system. For instance, if the resources allocated to ensure response readiness (people, equipment, materials) are inadequate to respond to the threat scenario, further implementation of risk reduction measures should be considered. In this regard, cost-benefit analyses and other types of analyses in a larger context (referred to as a managerial review and judgment) could be used where the limitations and constraints of the analyses are considered.

After identification of essential system functions (step 1), the second step is to characterize the potential and expected variability of each function, including sources and impacts. For each function in our example there can be variability. For instance, a defined function, F1, is continuous monitoring of a prison environment. The variability of F1 depends on different factors, including pressure and working environment. The variability could be well intended or caused by a malicious act (e.g., a prison officer hacks the CCTV and assists an inmate to escape). The characterisation of variability and understanding how the variability can affect other functions, and thereby the system as a whole, is essential to figuring out both how things go right and how they go wrong. This understanding provides useful insights to improve a system's resilience.

It is important to mention that the FRAM is just a tool to map functions and their coupling in a system. We have to acknowledge that this tool with its limitations. Functional resonance mapping is conditional on a number of assumptions and knowledge about interrelationship

between different functions in an organisation. This knowledge is subject to uncertainty, which is not reflected in the FRAM model.

7.5 Security risk evaluation, options assessment and selection:

The results from the previous section (6.4) provide insight to evaluate security risk, and consequently determine the need for risk prevention/reduction measures and the selection of alternatives (Renn 2008). This evaluation and selection are based on value judgements, weighing criteria such as (p. 175):

- *Effectiveness:* Does the alternative achieve the desired effect?
- *Efficiency:* Does the alternative achieve the desired effect with the least consumption of resources?
- *Sustainability:* Does the option contribute to the overall goal of sustainability?
- *Fairness:* Does the alternative burden the subjects of regulation in a fair and equitable manner?
- *Political and legal implementation ability:* Is the alternative compatible with legal requirements and political programmes?
- *Ethical and public acceptability:* Is the option morally acceptable? Will the alternative be accepted by the individuals affected by it?

Measuring risk reduction/prevention options against these criteria may create conflicting results. For instance, related to our case example, implementation of technologies and options for drone detection and deterrence, such as DroneShield could be considered an effective alternative and a good preventive measure against drone-assisted crime in prison. However, there are some issues concerning the expected cost and utility of this measure. The question that needs to be addressed is whether the security risk is high compared to relevant reference values of alternative measures? The selection among different alternatives is a challenging task as it requires balancing different concerns, such as risks and costs. It is the decision maker's responsibility to undertake such considerations to decide on an appropriate balance of the various concerns. However, it is beyond the scope of this paper to discuss this issue any further. We refer to Abrahamsen et al. (2017) for guidance on the selection of a suitable type of security risk-management strategy for the implementation of various types of security measures in different decision-making contexts.

8. Conclusions and final remarks

The purpose of this paper has been to explore how to apply the Safety-II concept in a security context. We addressed three research questions, concerning the definition of risk, the link between Safety-II and security, and security risk assessment. We presented an alternative, security risk perspective, where uncertainty is one of the main components of risk definition. We believe this broader security risk perspective is more suitable for assessing and managing

risk and resilience in the security context, as it allows for questions to be asked concerning existing knowledge in expressing uncertainty. We discussed how organisational culture, the formulation of strategy that promotes a willingness to devote resources for security purposes and the capacity of an organisation to operate in changing circumstances, affect an organisation's potential resilience. We also presented an extended framework that links security risk assessment and the Safety-II concept. The extended processes ensure a broader perspective, link risk, vulnerability and resilience, and provide insights from different traditions and perspectives.

A fundamental problem in analysing risk and resilience in the security context is that it is difficult to express uncertainty and determine how likely it is that an incident/event happened (e.g., an attempted escape assisted by a drone); we are unable to give strong arguments for specific likelihood assignments of threat occurrence. Yet, a likelihood can always be assigned based on available knowledge. An extended risk assessment acknowledges this and considers a set of methods, both qualitative and quantitative, to reflect this (lack of) knowledge. Addressing uncertainties and knowledge, we obtain a stronger focus on the factors that are important for obtaining resilience (I-IV), and hence, the application of Safety-II in a security context.

This article does not present how to apply resilience analysis, using the Functional Resonance Analysis Method (FRAM) in a real case study in a security context. Future research in this promising area could include further conceptualising of Safety-II in the security context, strategy formation to increase resilience in the security context and the application of the FRAM to assess the risk of organisational change in a security context.

References

Abrahamsen EB, Pettersen K, Aven, T, Kaufmann M, Rosqvist, T (2017) A framework for selection of strategy for management of security measures. Journal of Risk Research 20(3):404-417, doi:101080/1366987720151057205

Adger WN (2006) Vulnerability. Global Environmental Change 16(3):268-281

Alberts CJ (2002) Managing information security risks: the OCTAVE approach. Boston, Addison-Wesley. Available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.461.7807&rep=rep1&type=pdf. Accessed 18 Nov 2018

Alberts CJ, Behrens GS, Pethia DR, Wilson RW (1999) Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE) Framework, Version 10. US Department of Defense The Software Engineering Institute. Available at https://resources.sei.cmu.edu/asset_files/TechnicalReport/1999_005_001_16769.pdf. Accessed 21 Nov 2018

Amundrud Ø, Aven T, Flage R (2017) How the definition of security risk can be made compatible with safety definitions. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 231(3):286-294 doi:101177/1748006X17699145

Aven T (2011) Quantitative risk assessment: The scientific platform Cambridge. Cambridge, University Press

Aven T (2014) Risk, surprises and black swans: fundamental ideas and concepts in risk assessment and risk management. London: Routledge
Aven T (2015) Implications of black swans to the foundations and practice of risk assessment and management. Reliability Engineering and System Safety 134:83-91
Aven T (2016) Risk assessment and risk management: Review of recent advances on their foundation. European Journal of Operational Research 253(1):1-13 doi:101016/jejor201512023
Aven T, Renn O (2010) Risk Management and Governance Concepts, Guidelines and Applications. Springer-Verlag Berlin Heidelberg
Bellini E, Ceravolo P, Nesi P (2017) Quantify resilience enhancement of UTS through exploiting connected community and Internet of everything emerging technologies ACM Transactions on Internet Technology (TOIT) 18(1):114-147. doi: 10.1145/3137572
Beyerer J, Geisler J (2016) A Framework for a Uniform Quantitative Description of Risk with Respect to Safety and Security. European Journal for Security Research 1(2):135-150. doi:101007/s41125-016-0008-y
Birkmann J, Cardona OD, Carreño ML, Barbat AH, Pelling M, Schneiderbauer S, Welle T (2013) Framing vulnerability, risk and societal responses: the MOVE framework. Natural Hazards 67: 93-211
Bjerga T, Aven T, Zio E (2016) Uncertainty treatment in risk analysis of complex systems: The cases of STAMP and FRAM. Reliability Engineering and System Safety 156:203-209. doi:10.1016/j.ress.2016.08.004
Bruneau M, Chang SE, Eguchi RT, Lee GC, O'Rourke TD, Reinhorn AM, von Winterfeldt D (2003) A framework to quantitatively assess and enhance the seismic resilience of communities Earthquake Spectra,19: 733–752
De Berker AO, Rutledge, RB, Mathys, C, Marshall, L, Cross, GF, Dolan, RJ, and Bestmann, S (2016) Computations of uncertainty mediate acute stress responses in humans. Nature Communications 7. doi: 101038/ncomms10996
Dwyer A, Zoppou C, Nielsen O, Day S, Roberts S (2004) Quantifying Social Vulnerability:A methodology for identifying those at risk to natural hazards Canberra. Australia: Geoscience Australia
Fairchild A, MacKinnon D (2009) A General Model for Testing Mediation and Moderation. Effects Prev Sci 10(2): 87-99. doi:101007/s11121-008-0109-6
Flage R, Aven T (2015) Emerging risk – Conceptual definition and a relation to black swan type of events. Reliability Engineering and System Safety 144:61-67. doi:10.1016/j.ress.2015.07.008
Fraser SW, Greenhalgh T (2001) Coping with complexity: educating for capability. British Medical Journal 323(7316):799-803
George LA (1986) The Impact of Crisis-Induced Stress on Decision Making. In: Fredric Solomon and RQ Marston (Eds), The Medical Implications of Nuclear War (pp 528-552) Washington DC: National Academies Press
Gharajedaghi J (2011) Systems Thinking: Managing Chaos and Complexity: A Platform for Designing Business Architecture (3 ed): Elsevier Science
Hollnagel E (2006) Resilience: The challenge of the unstable. In: E H David D Woods, Nancy Leveson (Ed), Resilience Engineering: Concepts and Precepts. Aldershot, UK: Ashgate, pp 275-296
Hollnagel E (2011) Epilogue: RAG- The Resilience Analysis Grid. In: E Hollnagel, J Pariès, J Wreathall, and DD Woods (Eds), Resilience engineering in practice: A guidebook. Farnham, UK: Ashgate, pp 275-296

Hollnagel E (2012) FRAM: The Functional Resonance Analysis Method: Modelling Complex Socio- technical Systems. Farnham, Ashgate Publishing Ltd
Hollnagel E (2014) Becoming Resilient. In: PC Nemeth and E Hollnagel (Eds), Resilience engineering in practice: Volume 2: Becoming resilient. Farnham, UK: Ashgate publishing, pp 179-192
Hollnagel E (2016a) Resilience engineering: a new understanding of safety. Journal of the Ergonomics Society of Korea 35:185-191
Hollnagel E (2017) Safety-II in Practice: Developing the Resilience Potentials. Routledge Ltd
Hollnagel E, Speziali J (2008) Study on Developments in Accident Investigation Methods: A Survey of the "State-of-the-Art". (1104-1374) Available at https://hal-mines-paristech.archives-ouvertes.fr/hal-00569424/document Accessed 12 Oct 2017
Hollnagel E, Wears RL, Braithwaite J (2015) From Safety-I to Safety-II: A White Paper. Published simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie University, Australia: The Resilient Health Care Net Available at https://www.england.nhs.uk/signuptosafety/wp-content/uploads/sites/16/2015/10/safety-1-safety-2-whte-papr.pdf Accessed 10 Aug 2017
Häring I, Ebenhöch S, Stolz A (2016) Quantifying Resilience for Resilience Engineering of Socio Technical Systems. European Journal for Security Research 1(1):21-58 doi:101007/s41125-015-0001-x
International Organization for Standardization (2018) Information technology - security techniques - information security risk management (Third ed, International standard ISO/IEC). Geneva: ISO
Jore, SH (2017) Safety and Security- Is there a need for an Integrated approach? In: Walls L, Revie M, Bedford T (eds) Risk, Reliability and Safety: Innovating Theory and Practice. Taylor and Francis Group, CRC Press, London, pp 852-859
Jore SH, Egeli, A (2015) Risk management methodology for protecting against malicious acts: are probabilities adequate means for describing terrorism and other security risks? In: Podofillini L, Sudret B Stojadinovic B, Zio E, Kräger W (eds) Safety and reliability of complex engineered systems. CRC Press, London, pp 807–815
Jore SH, Utland I-LF, Vatnamo VH (2018) The contribution of foresight to improve long-term security planning Foresight. Journal of Futures Studies, Strategic Thinking and Policy 20(1):68-83. doi:101108/FS-08-2017-0045
Katsikas SK (2012) Risk Management. In: Vacca JR (ed) Computer and Information Security Handbook. Elsevier Science and Technology, pp 905-927
Kifer M, Hemmens C, Stohr MK (2003) The Goals of Corrections: Perspectives from the Line. Criminal Justice Review 28(1):47-69. doi:101177/073401680302800104
Landoll D (2011) Security Risk Assessment Handbook. CRC Press
Levenson E, Jones S (2017) South Carolina inmate used drone, makeshift dummy to escape prison. Available at https://edition.cnn.com/2017/07/07/us/sc-prison-escape-drone/index.html Accessed 10 Sep 2017
Maitlis S, Christianson M (2014) Sensemaking in Organizations: Taking Stock and Moving Forward. The Academy of Management Annals 8(1):57-125. doi:101080/194165202014873177
Masse T, O'Neil S, Rollins J (2007) The Department of homeland security's risk assessment methodology: evolution, issues, and options for congress. Congressional Research Service Washington DC
Mintzberg H (1978) Patterns in Strategy Formation. Management Science 24(9):934-948. doi:10.1287/mnsc.24.9.934

Mohaghegh Z, Kazemi R, Mosle A (2009) Incorporating organizational factors into Probabilistic Risk Assessment (PRA) of complex socio-technical systems: A hybrid technique formalization. Reliability Engineering and System Safety 94(5):1000-1018. doi:10.1016/j.ress.2008.11.006

Murphy DM, Paté-Cornell ME (1996) The SAM Framework: Modeling the Effects of Management Factors on Human Behavior in Risk Analysis. Risk Analysis 16(4):501-515. doi:10.1111/j.1539-6924.1996.tb01096.x

NS 5831 (2014) In Samfunnssikkerhet – Beskyttelse mot tilsiktede uønskede handlinger – Krav til sikringsrisikostyring: Societal Safety – protection against intentional unwanted actions – requirements to security risk management. Available at https://www.standard.no/no/Nettbutikk/produktkatalogen/Produktpresentasjon/?ProductID=718201 Accessed 10 Sep 2018

Ocasio W (1997) Towards An Attention-Based View Of The Firm. Strategic Management Journal 18(1):187-206. Available at https://onlinelibrary.wiley.com/doi/epdf/10.1002/%28SICI%291097-0266%28199707%2918%3A1%2B%3C187%3A%3AAID-SMJ936%3E3.0.CO%3B2-K Accessed 05 Oct 2018

Ojanen H (2017) The EU's Power in Inter-Organisational Relations. Springer. p.122 doi:10.1057/978-1-137-40908-9

Papazoglou IA, Bellamy LJ, Hale AR, Aneziris ON, Ale BJM, Post JG, Oh, JIH (2003) I-Risk: development of an integrated technical and management risk methodology for chemical installations. Journal of Loss Prevention in the Process Industries 16(6): 575-591. doi:10.1016/j.jlp.2003.08.008

Rajesh S, Jain S, Sharma P (2018) Inherent vulnerability assessment of rural households based on socio- economic indicators using categorical principal component analysis: A case study of Kimsar region, Uttarakhand. Ecological Indicators 85:93-104. doi:10.1016/j.ecolind.2017.10.014

Renn O (2008) Risk governance: Coping with uncertainty in a complex world (Earthscan risk in society series). London: Earthscan

Society for Risk Analysis (2018) Society for Risk Analysis Glossary Available at http://sra.org/sites/default/files/pdf/SRA%20Glossary%20-%20FINAL.pdf Accessed 05 Oct 2018

Steen R, Aven T (2011) A risk perspective suitable for resilience engineering Safety Science 49:292-297 doi:10.1016/j.ssci.2010.09.003

Sutcliffe KM, Vogus TJ (2003) Organizing for resilience. In: Cameron KS, Dutton JE, Quinn RE (eds) Positive organizational scholarship: foundations of a new discipline San Francisco, Calif: Berrett-Koehler, pp 94-110

Taleb NN (2007) The black swan : the impact of the highly improbable. Allen Lane, London

Tangenes T, Steen R (2017) The trinity of resilient organisation: aligning performance management with organisational culture and strategy formation. International Journal of Business Continuity and Risk Management 7(2):127-150

Turner BL, Kasperson RE, Matson PA, McCarthy JJ, Corell RW, Christensen L, Schiller A (2003) A Framework for Vulnerability Analysis in Sustainability Science. Proceedings of the National Academy of Sciences of the United States of America, 100(14):8074-8079. doi:10.1073/pnas.1231335100

Weick KE (2016) D. Christopher Kayes: Organizational Resilience: How Learning Sustains Organizations in Crisis, Disaster, and Breakdowns. Administrative Science Quarterly 61(1). doi:10.1177/0001839215615333

Weick KE, Sutcliffe KM, Obstfeld D (2005) Organizing and the process of sensemaking Organization Science 16(4):409-421. doi:10.1177/0001839215615333

Zio E, Aven T (2011) Uncertainties in smart grids behavior and modeling: What are the risks and vulnerabilities? How to analyze them? Energy Policy 39(10):6308–6320. doi:10.1016/j.enpol.2011.07.030