# RAG – Resilience Analysis Grid

Erik Hollnagel

# Introduction

A system<sup>1</sup> cannot **be** resilient, but a system can have a potential for resilient performance. A system is said to perform in a manner that is resilient when it sustains required operations under both expected and unexpected conditions by adjusting its functioning prior to, during, or following events (changes, disturbances, and opportunities). Whereas current safety management (Safety-I) focuses on reducing the number of adverse outcomes by preventing adverse events, Resilience Engineering (RE) looks for ways to enhance the ability of systems to succeed under varying conditions (Safety-II). It is therefore necessary to understand what this ability really means, since it clearly is not satisfactory just to call it 'resilience'.

The purpose of the rather roundabout definition given above is to avoid statements such as 'a system **is** resilient if ...', since this narrows resilience to a specific quality. (Or even worse, that 'a system **has** resilience if ...'.) RE has from the very beginning maintained that resilience is a characteristic of how a system performs, not a quality that the system as such has or possesses. Resilience is functional and not structural. If we want to use a short description, we should therefore refer to a system's **resilient performance** rather than a system's resilience.

# Safety as a Quality

A system is traditionally considered to be safe if the number of adverse outcomes is acceptably low. Such outcomes are typically accidents and incidents, but may also include work time injury, work related illnesses, etc. The level of safety corresponds to the number of such outcomes, and the common interpretation is that a higher level of safety

<sup>1</sup> In this Technical Note, a 'system' is used in a broad sense and includes, for instance, the organisation.

corresponds to a lower number of adverse outcomes. One example of that is the International Civil Aviation Organisation's definition of safety as:

"... the state in which the risk of harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and risk management."

There is, however, more to safety than reducing the number of adverse events. RE defines safety as the ability to succeed under varying conditions, cf., above. This definition includes the traditional meaning of safety, since the ability to succeed under varying conditions will lead to fewer adverse outcomes – something that goes right cannot at the same time go wrong. To distinguish the two definitions, they have been called Safety-I and Safety-II, respectively (Hollnagel, 2014). Where the focus of the Safety-I definition is on protection and prevention against harmful events (*protective safety*), the focus of the Safety-II definition is more broadly on the system's ability to function in a way that produces acceptable outcomes (*productive safety*). RE is about what a system needs for its continued existence and growth, hence addresses both safety and core business processes (productivity, quality, and effectiveness). This has consequences for how safety is understood or defined, for how it is measured, and for how it is managed.

# **Reactive and Proactive Adjustments**

The key feature of a resilient system is its ability to adjust how it functions. Adjustments can in principle take place either after something has happened (be reactive, responding to feedback), or take place before something happens (be anticipatory or proactive, controlled by feedforward).<sup>2</sup>

• Reactive adjustments are by far the most common. For instance, if there is a major accident in a community, such as a large fire or an explosion, local responders will change their state of functioning and prepare for the many different types of consequences that may follow. These are the short-term or single-loop responses. Responding when something has happened is, however, not enough to guarantee a system's safety and survivability. One reason is that a system can only be prepared to

<sup>2</sup> The meaning of feedforward is that actions are based on calculations or assumptions about what will happen in the future – either in the short run or the long run.

respond to a limited set of events or conditions, and usually only for a limited duration. Another reason is that the damage then will have had time to grow and spread.

• Proactive adjustment means that the system can change from a state of normal operation to a state of heightened readiness before something happens. In a state of readiness, resources are allocated to match the needs of the expected event and special functions may be activated. A trivial example from the world of aviation is to secure the seat belts before start and landing or during turbulence. In this case the future events are consequences of regular, scheduled activities, hence highly predictable. In other cases the criteria for changing from a normal state to a state of readiness may be less obvious either because of a lack of experience, because the future is uncertain, because the validity of indicators is questionable, or because the signals are 'weak'.

# The Four Abilities - The Basis of Resilient Performance

The broad working definition of resilient performance can be made more precise and operational by considering what makes resilient performance possible. Since resilient performance is possible for most, if not all, systems, the explanation must refer to something that is independent of any specific domain. RE has proposed the following four abilities are necessary for resilient performance (Hollnagel, 2011) :

- The ability to **respond**. Knowing what to do, or being able to respond to regular and irregular changes, disturbances, and opportunities by activating prepared actions or by adjusting current mode of functioning.
- The ability to monitor. Knowing what to look for, or being able to monitor that which
  is or could seriously affect the system's performance in the near term<sup>3</sup> positively or
  negatively. The monitoring must cover the system's own performance as well as what
  happens in the environment.

<sup>3</sup> In practice this means within the time-frame of ongoing operations, such as the duration of a flight or the current segment of a procedure.

- The ability to **learn**. Knowing what has happened, or being able to learn from experience, in particular to learn the right lessons from the right experience. (This corresponds to the double-loop learning described by Argyris & Schön, 1974.)
- The ability to **anticipate**. Knowing what to expect, or being able to anticipate developments further into the future, such as potential disruptions, novel demands or constraints, new opportunities, or changing operating conditions.

The reason why there are four abilities rather than three or five (or some other number) is simply pragmatic. The four abilities proposed here can be easily be recognised in historical as well as present event analyses, and seem together to be sufficient without any being redundant.

The reason why the set of four is constituted by <response, monitoring, learning and anticipation><sup>4</sup> and not by a different set is likewise pragmatic. In other words, there is no strong theory that leads to the inevitable conclusion that it must be these four abilities and not another set. Having said that, it is nevertheless easy to argue that all four are necessary.

A system that is unable to respond is doomed, possibly in the short run and definitely in the long. Responding can, however, not be effective if the set of responses is fixed, no matter how large the initial set is. Unless the system's environment is completely stable, the responses must change and develop over time, which means that the system must be able to learn.

The ability to respond also depends on the ability to monitor. Without monitoring the system must constantly be in a high state of alert for every possible condition for which a response has been prepared. That is neither possible nor reasonable (from an economic or productivity point of view). Without monitoring, without some kind of forewarning, every situation will be a surprise. That is clearly not a sustainable condition.

Both responding and monitoring must furthermore be revised or adjusted based on experiences, i.e., based on learning. Learning must serve to strengthen or reinforce that which worked well, and change or adjust that which did not work well.

<sup>4</sup> As a brief reference to the four abilities, the term <RMLA> will be used for the rest of this Technical Note.

It is finally an advantage to be prepared for something that is potentially possible, although it may not have happened yet, cf., Westrum's discussion of regular, irregular and unexampled threats (Westrum, 2006). If the working environment is dynamic but stable, then anticipation may not be necessary. But if the environment changes even a little during the lifetime of the system, then anticipation clearly becomes necessary.

It can rather easily be argued that the four abilities are necessary, since the absence of any of them makes it impossible for a system to have a resilient performance. Another question is whether the four are sufficient, or whether additional abilities should be included. While there are good reasons to for considering the four as both necessary and sufficient, the argument for the latter is too long to go into here. But consider two candidates for additional abilities that have been proposed at one time or another.

One is the ability to adapt. While there is no denying that it is important to be able to adapt, adaptation is a composite rather than a primary ability. A system that is adaptive can adjust or modify itself, or rather the way it functions, to different conditions. This requires a combination of the ability to respond and the ability to learn, and possibly also the ability to monitor. Adaptation is therefore not a primary ability. Another is the ability to communicate. Communication may rightly be considered a primary ability, but primary for a system to exist rather than for resilient performance, hence on the same level as energy uptake and waste removal. For a system such as an organisation, explicit communication is necessary to coordinate how the various parts function. But communication itself does not provide a response.

#### Measurements of the Potential for Resilient Performance

'Resilience' refers to to something that the system does rather than to something that the system has; but it refers to something that is multifaceted rather than something that can be described by a single quality or dimension. There is no 'resilience', hence no quantity, amount, or level of resilience. The literature on RE shows that there are many different opinions about the 'phenomenology' of resilience – of what it is that characterises resilient performance (e.g., Hollnagel, Woods & Leveson, 2006; Hollnagel et al. 2011). So instead of considering what resilient performance **is**, we should consider what **enables** resilient

performance, what makes it possible – and conversely what would make it impossible, if it was missing.

From this point of view it makes sense to consider the four abilities that provide the basis for resilient performance. In principle we might simply try to determine the extent to which each is present in, or supported by, a system. Indeed, on an overall level we might ask about how well a system is able to <RMLA>. While it in some cases could be meaningful to address each ability as a simple, uniform quality, it will be far more practical to look at the details of each ability. This can be done, for instance, by using a goals-means analysis or a functional decomposition to reveal which specific functions or sub-functions are needed to enable a system to <RMLA>. The answers to such detailed questions can be used to develop a profile of the potential for each ability, hence the potential for resilient performance overall, and in that way serve a (composite) proxy measure for 'resilience'. This proxy measure has been called the Resilience Analysis Grid or RAG.<sup>5</sup>

# Generic and Specific Questions

The basic idea of the RAG is to develop a set of questions to determine how well a system does on each of the four basic abilities. But rather than asking the single question "How well is system X able to <RMLA>", a set of more precise questions is developed which address important aspects of each ability. RE provides a set of generic questions for each ability, as described in the following. It is, however, important to point out that these sets cannot be used without first being tailored to the target particular domain or application. Their main purpose is to serve as the starting point for developing sets of (diagnostic) questions that are specific for the chosen system.

With this in mind, the following sections will describe how each of the four abilities can be analysed in more detail.

## The ability to respond

No system, organisation, or organism can survive unless it is able to respond to what happens. Responses must furthermore be both timely and effective so that they can bring

<sup>5</sup> The name is a vague and possibly misleading allusion to the psychological technique known as Kelly's Repertory Grid. In hindsight, it might have been wiser to use the homophone RAQ, meaning Resilience Assessment Questionnaire.

about the desired outcome before it is too late. In order to respond, the system must therefore first detect that something has happened, then recognise what it is and determine whether a response is necessary, and finally know how to respond, when to begin, and when to stop.

In order to be able to respond it is necessary either to have prepared responses and resources at the ready, or to be flexible enough to reconfigure the existing configuration so that the necessary resources become available. A response, e.g., to an alert, may also be that the system changes into a state of readiness without interrupting what it otherwise is doing. In responding to events, it is essential to be able to distinguish between what is urgent and what is important.

Event list	What are the events for which the system has a prepared response?
Background	How were these events selected (tradition, regulator requirements, design basis, experience,
	expertise, risk assessment, industry standard, etc.)?
Relevance	When was the list created? How often is it revised? On which basis is it revised? Who is
	responsible for maintaining and evaluating the list?
Threshold	When is a response activated? What is the triggering criterion or threshold? Is the criterion
	absolute or does it depend on internal / external factors? Is there a trade-off between, e.g.,
	safety and productivity?
Response list	How was the specific type of response list decided? How is it ascertained that it is adequate?
	(Empirically, or based on analyses or models?)
Speed	How fast is full response ability available? How fast can an effective response be
	implemented?
Duration	For how long can a 100% effective response be sustained? What is the minimum acceptable
	response level and how long can it be sustained?
Stop rule	What is the criterion for ending the response and returning to a "normal" state?
Response	How many resources are allocated to ensure response readiness (people, equipment,
capability	materials)? How many are exclusive for the response potential? Who is responsible for
	maintaining the response ability?
Verification	How is the readiness to respond maintained? How and when is the readiness to respond
	verified?

Table 1: Examples of detailed issues relating to the ability to respond

# The ability to monitor

Resilient performance is not possible unless a system is able flexibly to monitor both its own performance (what happens inside the system's boundary) and what happens in the environment (outside the system's boundary). Monitoring improves the system's ability to cope with possible near-term events – threats and opportunities alike. In order for the monitoring to be flexible, its basis must be revised from time to time.

Monitoring normally relies on indicators. One type of indicators are called 'leading' indicators, because they can be used as valid precursors for changes and events that are about to happen. 'Leading' indicators are generally seen as very attractive (Hopkins, 2009). The main difficulty with 'leading' indicators is that the interpretation requires an articulated description, or model, of how the system functions. In the absence of that, 'leading' indicators are defined by association or spurious correlations. Because of this, most systems rely on 'lagging' indicators, such as accident statistics.

The dilemma of 'lagging' indicators is that while the likelihood of success increases the smaller the lag is (because early interventions are more effective than late ones), the validity or certainty of the indicator increases the longer the lag (or sampling period) is.

Indicator list	How have the indicators been defined? (By analysis, by tradition, by industry consensus, by
	the regulator, by international standards, etc.)
Relevance	When was the list created? How often is it revised? On which basis is it revised? Who is
	responsible for maintaining the list?
Indicator type	How many of the indicators are of the 'leading,' type and how many are of the 'lagging'? Do
	indicators refer to single or aggregated measurements?
Validity	How is the validity of an indicator established (regardless of whether it is 'leading' or
	'lagging')? Do indicators refer to an articulated process model, or just to 'common sense'?
Delay	For 'lagging' indicators, how long is the typical lag? Is it acceptable?
Measurement	What is the nature of the 'measurements'? Qualitative or quantitative? (If quantitative, what
type	kind of scaling is used?)
Measurement	How often are the measurements made? (Continuously, regularly, every now and then?)
frequency	

Table 2: Examples of detailed issues relating to the ability to monitor

Analysis /	What is the delay between measurement and analysis/interpretation? How many of the
interpretation	measurements are directly meaningful and how many require analysis of some kind? How
	are the results communicated and used?
Stability	Are the measured effects transient or permanent?
Organisational	Is there a regular inspection scheme or -schedule? Is it properly resourced?
support	

# The ability to learn

The ability to respond and the ability to monitor both depend on the ability to learn, unless the environment is perfectly stable and perfectly predictable. Efficient and systematic learning from experience requires careful planning and ample resources. The effectiveness of learning depends on the basis for learning, i.e., which events or experiences are taken into account as well as on how the events are analysed and understood.

In learning from experience it is important to separate what is *easy* to learn from what is *meaningful* to learn. The level of safety is often couched in terms of the number or frequency of occurrence of adverse events. But compiling extensive accident statistics does not mean that anyone actually learns anything. Counting how often something happens is not learning. Knowing how many accidents have occurred, for instance, says nothing about why they have occurred, nor anything about the many situations when accidents did not occur. And without knowing why something happens, as well as knowing why it does not happen, it is impossible to propose effective ways to improve safety.<sup>6</sup>

In safety management, learning has traditionally focused on things that go wrong (accidents and incidents) both because they are easy to perceive and because they are a cause of concern. But since the number of things that go right, including near misses, is many order of magnitudes larger than the number of things that go wrong, it makes obvious sense to try to learn from representative events (frequency) rather then only from failures (severity).

<sup>6</sup> This goes for Safety-I as well as for Safety-II.

Selection	Which events are investigated and which are not (frequency, severity, value, etc.)? How is the
criteria	selection made, which criteria are used? Who makes the selection?
Learning basis	Does the system try to learn from successes (things that go right) as well as from failures
	(things that go wrong)?
Classification	How are events described? How are data collected and categorised?
Formalisation	Are there any formal procedures for data collection, analysis and learning?
Training	Is there any formal training or organisational support for data collection, analysis and
	learning?
Learning style	Is learning a continuous or discrete (event-driven) activity?
Resources	How many resources are allocated to investigation and learning? Are they adequate? Which
	criteria do they depend upon?
Delay	What is the delay in reporting and learning? How are the outcomes communicated
	internally and externally?
Learning target	On which level does the learning take effect? (For instance, individual, collective,
	organisational.)
Implementation	How are 'lessons learned' implemented? Regulations, procedures, norms, training,
	instructions, redesign, reorganisation, etc.?

Table 3: Examples of detailed issues relating to the ability to learn

# The ability to anticipate

While monitoring makes immediate sense, it may be less obvious that it is useful to look at the more distant future as well. The purpose of looking at the potential is to anticipate possible future events, conditions, threats, and opportunities that may either be beneficial or detrimental to the system's continued functioning.

Risk assessment focus on future threats and is suitable for systems where the principles of functioning are known, where descriptions do not contain too many details, where descriptions can be made relatively quickly, and where systems are so stable so that descriptions remain valid for a long time. Many present day systems where industrial safety is a concern are unfortunately not like that, but are instead underspecified. For such systems the principles of functioning are only partly known, descriptions contain (too) many details and take a long time to make, and the systems keep changing so that

descriptions must be frequently updated.<sup>7</sup> Traditional risk assessment methods are therefore inadequate, if not downright inappropriate.

The anticipation for future opportunities has little support in current methods, although it rightly ought to be considered as important as the search for threats.

Expertise	What kind of expertise is relied upon to look into the future? (In-house, outsourced?)
Frequency	How often are future threat and opportunities assessed?
Communication	How are the expectations about future events communicated or shared within the system?
Strategy	Does the system have a clearly formulated 'model of the future'?
Model	Is the model or assumptions about the future explicit or implicit? Qualitative or
	quantitative?
Time horizon	How far ahead does the system look ahead? Is the time horizon different for, e.g., business
	and safety?
Acceptability of	Which risks are considered acceptable and which unacceptable? On which basis?
risks	
Aetiology	What is the assumed nature of the future (threats, opportunities)?
Culture	Is risk awareness part of the organisational culture?

Table 4: Examples of detailed issues relating to the ability to anticipate

# Rating the Potential for Resilient Performance

The four sets of questions described above constitute the Resilience Analysis Grid (RAG). The purpose of using the RAG is **not** to provide an absolute rating of how well a system does on the four basic abilities. There are several reasons for this. The most important is probably that there is no meaningful standard or norm that can be used as either a reference or a criterion. A second reason is that answers to the RAG questions represent a more or less arbitrary point in time. A third that the ratings refer to an ordinal scale at best; and a fourth that the questions may have different meanings for different organisations and contexts, etc.

The purpose of the RAG is rather to provide a well-defined characterisation (or profile) of a system that can be used to manage the system and specifically to develop its potential for resilient performance. The intention is that the RAG is applied regularly so

<sup>7</sup> In extreme cases, the system may change faster than a description can be produced. Descriptions will therefore always be incomplete and the system therefore underspecified.

that it becomes possible to see if there have been any changes. In that way the RAG can be used to monitor system changes, hence to manage the changes.<sup>8</sup>

In order for the RAG to be useful as a tool, it is necessary that the answer to each item can be rated. The rating can, for example, use the following Likert-type scale:

- Excellent the system meets and exceeds the criteria for the required ability.
- Satisfactory the system fully meets all reasonable criteria for the required ability
- Acceptable the system meets the nominal criteria for the required ability.
- Unacceptable the system does not meet the nominal criteria for the required ability.
- Deficient there is insufficient ability to provide the required ability.
- Missing there is no ability to provide the required ability.

(The Likert-type scale is proposed because it is widely used. Other forms of rating may, of course, be used instead.)

The ratings of individual items can be presented in a variety of ways. It might seem attractive to produce a single measure, for instance by aggregating the answers to a set of questions. The Likert-type scale makes this tempting since each answer easily can be assigned a numerical value. There are, however, two serious objections. First, that the ratings are given on an ordinal rather than an interval scale. This means that the difference between, e.g., 'acceptable' and 'unacceptable' is not the same as the difference between, e.g., 'unacceptable' and 'deficient'. Second, that the relative importance or weight of the questions is undefined. Is the 'event list', for instance, more or less important than the 'background' for the ability to respond? Unless a precise answer can be given for each pairwise combination of the questions in each set, an aggregated measure is best avoided.

Another way of presenting the ratings is by means of a radar chart or star-plot. The radar chart uses a number of equi-angular spokes; each spoke represents one of the questions and the length of a spoke is proportional to the how the question was rated on

<sup>8</sup> The RAG is therefore one way in which the system can monitor itself.

the Likert-type scale. The result is a star-like polygon, which provides a clear signature of how well the system does with regard to the particular ability.<sup>9</sup>

Figure 1 below uses radar charts to show what the ratings for the ability to respond could look like for a system. The ratings are here assumed to have been made with a fourmonth interval. The differences between the two ratings are easy to see, and can be used to determine both whether the system develops in the right direction and where specific interventions should focus.

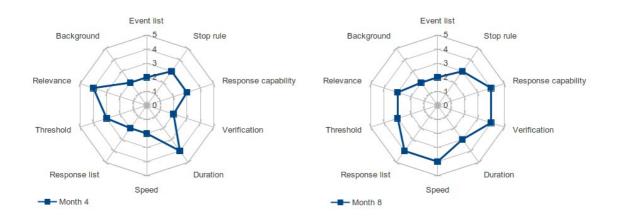


Figure 1: Radar charts illustration the use of the RAG.

# Managing the Potential for Resilient Performance

The sets of RAG questions that are developed for a specific use of the RAG should be formulated so that they can easily be assessed. This means that they should refer to concrete relations or characteristics of the system's performance, to something that the respondents have experience with or something that is described in the system's documentation. This has the added benefit that the questions themselves can be the basis for interventions to improve resilient performance.

Consider, for instance, the first question in Table 1: "What are the events for which the system has a prepared response?" If the answer is found to be *unacceptable*, meaning that the

<sup>9</sup> A further advantage is that this type of chart is a standard function in most spreadsheets.

list of events is incomplete or inappropriate, then this can be the starting point for proposing remedial activities. The consequences of such remedial activities can then be gauged by a later application of the RAG. (When that should happen obviously depends on how fast a change can be expected to take place.)

While a "one problem – one solution" approach is appealing – and indeed seems to be the preferred way to respond in Safety-I management, it disregards the fact that the issues addressed by the individual questions cannot be seen in isolation. A system cannot just be understood as a linear combination of its parts, but must be recognised as a whole where the dependencies or couplings among the parts is critical for overall performance.<sup>10</sup> As an example of that, consider the ability 'to respond' as a function. If we use the Functional Resonance Analysis Method (Hollnagel, 2012), we may find the following dependencies:

Name of function	Respond
Description	A system's ability to respond to what happens or may happen.
Aspect	Description of Aspect
Input	Alerts
	Interruptions
Output	Responses
Precondition	State of readiness
Resource	Tools, staff, materials
Control	Plans and procedures
Time	Work schedules

Table 5: The ability to respond described as a FRAM function.

The dependencies described in Table 5 can also be shown graphically as shown in Figure 2. (An explanation of the graphical elements used in Figure 2 can be found at <u>www.functionalresonance.com</u>.)

It will go beyond the scope of this Technical Note to provide a more detailed model of how the four basic abilities depend on each other.<sup>11</sup> Suffice it to say that it is important to resist from using a "one problem – one solution" approach in any kind of system management, whether the focus is safety, quality, productivity, or resilience.

<sup>10</sup> The determination of what constitutes the 'parts' is relative rather than absolute, and should refer to how the system functions rather than to how it is structure.

<sup>11</sup> That will be the subject of a forthcoming book.

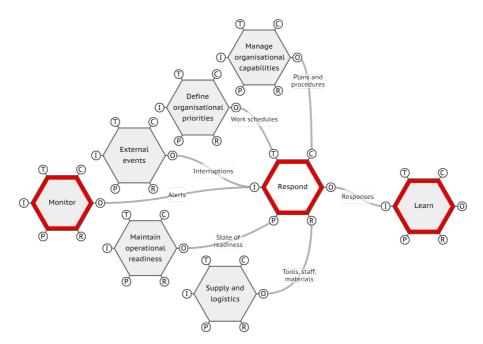


Figure 2: The ability to respond rendered as a FRAM function.

## Summary

The Resilience Analysis Grid is not an off-the-shelf tool that can be used directly. It is rather intended as a basis from which more specific sets of questions can be developed. The questions must clearly be relevant for the system where they are intended to be used, and may therefore require clarification and reformulation.

This Technical Note has outlined the principles for how the evaluations can be rated, and how the results can be presented. The radar chart is not in itself a measure of resilience, but a compact representation of how the various items were rated. It is also a process measure rather than a product measure, i.e., it shows the current potential for resilient performance in terms of how well the system does on each of the four main abilities.

RE does not prescribe a certain balance or proportion among the four abilities. This balance clearly is domain dependent, it is therefore impossible to propose a 'standard' value. For a fire brigade, for instance, it is more important to be able to respond than to

anticipate. Whereas for a sales organisation, the ability to anticipate may be just as important as the ability to respond. But RE does make clear that it is necessary for a system to posses each of these abilities to some extent, in order to have the potential for resilient performance. All systems traditionally put some effort into the ability to respond. Many also put some effort into the ability to learn, although it often is in a very stereotyped manner. Fewer systems make a sustained effort to monitor, particularly if there has been a long period of stability. And very few systems put any serious effort into the ability to anticipate.

## References

- Argyris, C. & Schön, D. (1974). Theory in practice: Increasing professional effectiveness. San Francisco: Jossey-Bass.
- Hollnagel, E. (2011). RAG The resilience analysis grid. In: E. Hollnagel, J. Pariès, D. D.
  Woods & J. Wreathall (Eds). *Resilience Engineering in Practice. A Guidebook*. Farnham, UK: Ashgate.
- Hollnagel, E. (2012). FRAM. The Functional resonance Analysis Method for modelling complex sociotechnicals systems. Farnham, UK: Ashgate.
- Hollnagel, E. (2014). Safety-I and Safety-II: The past and future of safety management. Farnham, UK: Ashgate.
- Hopkins, A. (2009). Thinking about process safety indicators. Safety Science, 47, 460-465.
- Westrum, R. (2006). A typology of resilience situations. In: E. Hollnagel, D. D. Woods & N. G. Leveson (Eds.), Resilience Engineering: Concepts and Precepts. Aldershot, UK: Ashgate.