Chapter 2

Literature Review

This chapter gives an overview of the various definitions of resilience for different systems and a brief description of the interpretations of resilience in the various disciplines. The chapter also covers the types of disruptions that create the need for implementing resilience in systems and the methodologies and metrics that have been used to assess resilience to date. Additionally, the elements of resilience upon which this book is based are introduced. Later on in the chapter, the main factors that contribute towards achieving resilience in organizations are summarized. The chapter concludes with an overview of the relationship between resilience and risk management as detailed in the literature.

2.1 Resilience Definitions

The term resilience comes from the Latin word *resilire*, meaning “to leap back.” Resilience in systems extends this definition to mean the ability of systems to “bounce back” after a disturbance has occurred. Many efforts have gone into defining and characterizing resilience, some of these definitions are summarized in the following paragraphs.

Folke *et al.* (2002) defines resilience as the capacity of the system to absorb disturbance and reorganize while undergoing change so as to retain the same function, structure, identity and feedbacks. Dalizell and McManus (2004) state that resilience describes the overarching goal of a system to continue to function to the fullest possible extent in the face of stress to achieve its purpose. Rose and Liao (2005) present a similar definition for systems resilience; they define resilience to be the ability of the system to apply adaptive responses in the face of disruptions in order to avoid potential losses.
Andersson (2006) views resilience as the ability of an actor, in the face of natural hazards, to cope with or adapt to hazard stress. He also mentions that the resilience is influenced by the proactive measures taken as a preparation for the hazard and the reactive measures in response to a hazard.

Fiksel (2003) defines a resilient system as a system that has the ability to return to a stable equilibrium state after a perturbation.

Bruneau et al. (2003) define a resilient system as one that has reduced failure probability, reduced consequences of failures, and reduced time to recovery. They also defined robustness, redundancy, resourcefulness and rapidity to be properties of resilient systems. Recent efforts by Reed et al. (2009) follow in these footsteps and identify quality as the capacity of the infrastructure. They define robustness to be the ratio of the lost capacity of the system as a result of a disruptive event to the capacity of a fully functioning structural system, and rapidity to be the measure of rapidity of recovery.

Pavard et al. (2006) define resilient systems to be those capable of maintaining a constant output value level when the system suffers from a perturbation.

The counterpart of resilience is brittleness. A resilient system is able to adapt to the shock and contain it; a brittle system, on the other hand, lacks the ability to adapt and transmitting exogenous shocks (Wears and Perry 2008).

2.2 Resilience in Different Disciplines

The concept of resilience existed long before it was adopted for engineering systems. It was first introduced by Holling in 1973 (Holling 1973) for ecological systems. He defined the resilience of ecological systems to be “the measure of the persistence of systems and their ability to absorb change and disturbance and still maintain the same relationship between populations or state variables”; that is, how far the system can be perturbed without shifting to a different regime. More resilient systems can absorb a larger amount of disturbance before shifting into a new regime (Walker et al. 2006).
Holling (1996) describes two views of resilience. The first is engineering resilience, which concentrates on the stability near an equilibrium steady state. It is measured through the resistance to disturbance and the speed of return to the steady state; more resilient systems have more than one equilibrium state. The second is ecological resilience, where the conditions are not at an equilibrium steady state. Here, the resilience is measured by the magnitude of disturbance that can be absorbed before the system changes its structure. Sugden (2001) relates the resilience of an ecosystem to recovery. He states that the resilience of an ecosystem is the extent to which the system can recover if an alien species is introduced to the system and then removed.

In psychology, resilience is the ability of an individual to withstand stresses and to bounce back or recover from traumatic situations; an example of a particularly resilient individual is Nelson Mandela, who went on to lead a country after years of solitary captivity (Neill 2006).

In materials science, resilience is the ability of the material to absorb energy when deformed elastically and to return to its original form when unloaded. It is represented by the area under the curve in the elastic region of the stress–strain diagram.

In business organizations, resilience is an issue of great importance since business disruptions can be fatal to an organization. Understanding and improving the resilience of the system entails looking into the interdependencies within and across organizations (Goble et al. 2002).

Figure 2.1 shows the resilience in different disciplines found in the literature.

![Resilience Diagram](image-url)
2.3 Resilience and Disruptions (Shocks)

The primary factor for implementing resilience in systems is to improve the reaction of the system after the occurrence of disruptions or system shocks and improve its ability to resume functionality. The following sections will introduce the main factors that cause disruptions in infrastructure systems and the typical disruption profile that describes the state of the infrastructure system following a disruption.

2.3.1 Categories of potential disruptions to systems

Disruptions can be caused by various sources of threats. The level of resilience required by a system depends on the potential level of damage that may result from disruptions. Sources of disruptions have been grouped by Mansouri et al. (2009b) into the following (see Fig. 2.2):

- **Human Factors** — Disruptions caused by humans operating or using the system, such as malicious attacks or accidental disruptions as a result of human error.

- **Natural Factors** — Disruptions due to damage caused by nature, such as hurricanes and floods.

- **Organizational Factors** — Events that occur at the organizational level, such as a workers strike.

- **Technical Factors** — Failures in system components, such as when equipment is faulty.

![Fig. 2.2. Sources of disruptions. Adapted from Mansouri et al. (2009b).](image)
Disruptions have been categorized by Westrum (2006) to be caused by regular and irregular threats. Regular threats occur frequently and these can be managed by preparing the systems to execute a standard response to overcome disruptions. Regular threats are predictable on a probabilistic basis, such as via statistical models, which are often used for estimating the likelihood of occurrence as well as the magnitude of natural disasters.

Most systems are designed according to system safety analysis guidelines that prepare the system to recover from predictable disruptions. Irregular threats are one-off events; predicting these types of disruptions is difficult, and in some cases, impossible. In general, whereas unpredictable disruptions may belong to any of the four categories, human disruptions are a little harder to predict than other sources of disruptions (Jackson 2009).

2.3.2 Disruption profile

When disaster strikes, the profile of the disruption on an infrastructure will typically have eight phases, as shown in Fig. 2.3 (Sheffi 2005).

The first phase, “Preparation,” indicates how well the system is prepared against threats. This stage will include proactive resilience measures that reduce the damage of predictable events. Unforeseen threats could lead to a higher degree of damage.

The second phase, the “Disruptive Event,” is the moment when the disruption occurs, followed by the “First Response,” during which the initial damage is attended to. An example of such a response is providing first aid if there were human casualties. At the same time, the recovery preparation takes place during this stage, aiming to bring the system back to its normal state. The full impact of the disruption is often not felt immediately; it typically takes a certain amount of time for the full extent of the damage to be felt. The severity of the disruptive event dictates how long it will take the system to recover. In some cases, it is not possible to recover completely from the disruption.

The purpose of implementing resilience is to change the appearance of the disruptive profile by reducing the area between the dotted line representing the normal performance of the system and the dip,
illustrated by the solid line, reflecting the impact of the disruptive event on the performance.

![Disruptive event profile](source: Sheffi (2005)).

### 2.4 Methodologies for Characterizing Resilience

Bruneau and Reinhorn (2007) defined the following four attributes to be dimensions of a resilient system:

- **Robustness** — the strength of a system and its elements to withstand a disruption
- **Redundancy** — the extent the system and its elements have substitutes to continue functioning after a disruptive event
- **Resourcefulness** — the capacity of the system to identify, prioritize and apply resources in the face of a disruption
- **Rapidity** — the capacity to meet priorities and achieve goals to limit loss and thwart future disruptions

Fiksel (2003) characterizes a resilient system as having multiple equilibrium points, and more resilient systems as having more equilibrium points. Figure 2.4 illustrates the different system states. The first curve shows a resistant system, which is able to recover rapidly
from a small perturbation but may not survive a large perturbation. The second curve shows a resilient system that can operate across a broad spectrum and is able to survive large perturbations. The third system is the most resilient as it has multiple equilibrium states. This type of system is able to continue operation after a large perturbation by having multiple equilibrium states achievable through a structural change.

![Adjacent system states](image)

Fig. 2.4. Adjacent system states. Source: Fiksel (2003).

Similar concepts that define the state of resilience in systems were proposed by Walker *et al.* (2004):

- **Latitude** — the amount by which the system can change before it loses its ability to recover
- **Resistance** — the ease of changing the system
- **Precariousness** — how close the system is to its threshold

The benefits of implementing resilience in systems are seen after the occurrence of disruptions; these benefits are intensified based on the measures taken before the occurrence of the disruptive events. Richards *et al.* (2007) identify two phases of resilience — namely, the anticipation/avoidance phase, and the survival/recovery phase. In the anticipation/avoidance phase, stakeholders need to identify foreseen and emergent vulnerabilities in order to take action that avoids or limits
The survival/recovery phase entails an adaptation to disruptions while maintaining operations.

Madni and Jackson (2009) have defined the pillars of resilience as: disruptions, systems attributes, methods, and metrics. The conceptual framework for the pillars of resilience is shown in Fig. 2.5.

The disruptions that can assault infrastructure systems result in a deterioration of the attributes that define the functionality of the infrastructure system. Traditional methods, such as risk and safety assessments, aid decision makers in making effective production/safety tradeoffs. However, metrics provide insight into various time and cost of restoration efforts, and how well the system responds to them.
2.5 Resilience Measurement Approaches

Holling introduced the concept of resilience in systems in 1973 (Holling 1973), but there was little motivation in quantitatively assessing resilience in the years prior to the September 11, 2001. Since the attacks, infrastructure resilience has become a topic of increasing interest. The following sections will review some of the resilience metrics that have been proposed for infrastructure systems.

2.5.1 Infrastructure resilience metrics

Bruneau and Reinhorn (2007) proposed a metric for measuring the seismic resilience of infrastructure systems. Their metric assumes a degradation of the quality of the infrastructure from 100 percent following an earthquake. The graph in Fig. 2.6 shows that if the time of occurrence of the disruptive event is \( t_0 \) and recovery is achieved at time \( t_1 \), the resilience can be measured as the time to recovery. This relationship is expressed in Eq. (2.1):

\[
R = \int_{t_0}^{t_1} [100 - Q(t)] \, dt ,
\]

where \( R \) is the infrastructure resilience, \( Q(t) \) is the quality of the infrastructure system, \( t_0 \) is the time of occurrence of the disruptive event, and \( t_1 \) marks the end of the recovery and restoration efforts.

Fig. 2.6. Schematic representation of the seismic resilience concept. Source: Bruneau et al. (2003).
The “resilience triangle” in Fig. 2.6 represents the loss of functionality from damage and disruption, as well as the pattern of restoration and recovery over time. Although Bruneau and Reinhorn’s efforts focus primarily on seismic resilience, their proposed metric can be tailored for any type of threat.

Reed et al. (2009) use this definition of resilience and identify the quality of the infrastructure as the infrastructure capacity. Infrastructure systems are interdependent by nature. Their research captures the interdependencies between the infrastructures through a linear function, although they state that a second-order trend might be more appropriate.

Attoh-Okine et al. (2009) propose the use of belief functions for measuring the resilience index of interdependent urban infrastructure systems.

Other efforts in defining resilience metrics for networked infrastructure systems include the methodology outlined by Garbin and Shortle (2007) where resilience metrics are the percentage of nodes or links damaged in the network versus the network performance.

2.5.2 Service infrastructures resilience metrics

Resilience metrics have also been defined for service infrastructures. Considering power infrastructures, Shinozuka et al. (2004) measure the resilience of power systems in terms of speed of restoration and repair efficiency. Chang and Chamberlin (2005) measure resilience in terms of economic loss.

In the field of transportation infrastructure, Werner et al. (2005) measure resilience as a function of the increase in travel time following a disruption. Murray-Tuite (2006) proposes metrics for evaluating 10 components of transport infrastructure resilience that she identifies as redundancy, diversity, efficiency, autonomous components, strength, adaptability, collaboration, mobility, safety, and the ability to recover quickly, and compares the system optimum and user equilibrium traffic assignments.

Heaslip et al. (2009) categorize the metrics with regards to the transportation infrastructure under individual resilience, community
resilience, economic resilience, and recovery metrics; each metric is evaluated as a combination of several indices that describe each metric.

Considering telecommunications infrastructure, specifically the Internet, Cohen et al. (2001) study the tolerance of the Net to intentional attack; the proposed resilience metric being a measurement of the number of sites needed for the disintegration of the network. Dolev et al. (2006) explore the resilience of the Internet at the autonomous system level. They propose resilience metrics derived from network properties such as the number of connected nodes in the network. Omer et al. (2009) propose a metric for the global transoceanic telecommunications network as the ratio of the value delivery of a network after a disruption to the value delivery before a disruption — the value delivery being identified as the amount of data transferred across the cables.

### 2.6 Elements of Resilience

The elements of resilience are the elements that need to be addressed in order to achieve resilience. Achieving resilience in systems is addressed in terms of vulnerability and adaptive capacity. Making a system less vulnerable and increasing its adaptive capacity results in a more resilient system. This relationship is illustrated in Fig. 2.7.

The following sections establish the linkages of vulnerability and adaptive capacity to resilience.

![Fig. 2.7. Relationship between resilience, vulnerability and adaptive capacity.](image-url)
2.6.1 Resilience and vulnerability

In the context of critical infrastructure systems, Haimes (2004) defined vulnerability to be the manifestations of the inherent states of the system that renders it susceptible to damage or loss. Sheffi (2005) characterizes the system’s vulnerability as a combination of the likelihood of a disruption and the potential severity of the disruption. A system with low vulnerability has a low probability of disruption with light consequences, while a system has high vulnerability when the probability of a disruption is high and the consequences are severe. Figure 2.8 illustrates Sheffi’s dimensions of vulnerability.

Dalziell and MacManus (2004) define a system of high vulnerability as a system that may be easily pushed from one state of stability or equilibrium to another, whereas it is not easy to push a system with low vulnerability from a state of stability or equilibrium.

Several studies establish the linkages between vulnerability and resilience. Gallopin (2006) relates resilience to the capacity of response, which he categorizes as a component of vulnerability. He also argues that even though a resilient system is less vulnerable, the relationship between the two is not symmetrical. Dalziell and MacManus (2004) refer to resilience as being a function of both the vulnerability of the system and its adaptive capacity. The relationship between resilience and
vulnerability is bidirectional: loss of resilience will result in the increasing vulnerability of the system and would increase the risk of shifting the system into an undesirable state (Resilience-Alliance 2007).

2.6.2 Resilience and adaptive capacity

The adaptive capacity of a system can be viewed as the capacity to adapt and to reconfigure in the face of disruptions without losing functionality (Resilience-Alliance 2007). It can also be viewed as the ability of a system to adjust to changing internal demands and external circumstances (Carpenter and Brock 2008).

Folke et al. (2002) define the adaptive capacity of the system to be the ability to cope with novel situations. Systems with a high adaptive capacity are able to reconfigure themselves while preserving functionality in the face of disruptions. Walker et al. (2004) define the adaptive capacity as a component that can be used to manage resilience. Gallopín (2006) suggests that the adaptive capacity of the system is an attribute that exists before the occurrence of a disruption.

Dalziell and McManus (2004) refer to resilience as a function of adaptive capacity and vulnerability. The adaptive capacity quality of the system before a disruptive event is an indication of the likelihood of successful adaptation following a disruptive event. More resilient systems have a larger adaptive capacity envelope, which enables the system to cope with changes imposed by a disruptive event.

2.7 Resilience in Organizations

Infrastructure systems have several layers amongst which is the organizational layer. Therefore, no infrastructure resilience discussion is complete without addressing the organizational aspect of infrastructure systems.

Organizations and enterprises are human-intensive systems that do not have a physical network for operation. In this type of system, people are themselves the physical infrastructure. Organizational systems have both the merits and drawbacks of being human-intensive systems. During
disruptive events, people can make adjustments to improve the situation; compensating, recovering and compromising. However, humans also make mistakes and even violations (Reason 2001).

Figure 2.9 identifies several key factors that lead to a resilient organizational structure. These are elaborated upon in the following:

- **Leadership**
  This is the most important element in organizational resilience. Enterprise leadership is about setting priorities, making commitments and the ability to make the right decisions about the courses of action to take when faced with adverse situations (O’Rourke 2007).

- **Awareness**
  Resilient organizations monitor change that occurs within the organization and are hence able to identify disruptions in advance. Data gathering provides management with the current state of affairs and reveals the extent of the problem as well as how prepared the organization is to deal with it (Wreathall 2006). Good communication is key to raising awareness. An organization with a strong communications infrastructure can more easily detect disruptions and alert the responsible persons.
• **Preparedness/Emergency Planning**

Organizations can actively anticipate problems and prepare for them by building a team that is able to imagine different possibilities and is able to apply inventive solutions (Johnson-Lenz 2009). Frequently deployed and even necessary schemes in emergency planning are emergency drill-response exercises (O’Rourke 2007). These exercises prepare the organization to deal with problems by training individuals in the courses of action to take in the event of emergencies.

• **Flexibility**

Flexibility allows organizations to adapt to new problems. Resilience through flexibility is achieved by allowing individuals in the organization to make decisions (Wreathall 2006). Flexibility is also achieved by creating redundancy or backup systems. Cross-training within an organization allows individuals to substitute for one another to a certain extent in cases of emergency.

• **Culture**

Resilience is achieved through a culture that is built on trust and accountability (Bell 2002), engaging individuals at all levels by developing a sense of shared purpose (Johnson–Lenz 2009), encouraging a culture that is more aware of its environment, and supporting communication throughout the organization (Wreathall 2006).

In organizational theory, loose coupling is a conceptual tool that coordinates the interactions between actors and specifies a certain course of action. Grote (2006) suggests the application of loose coupling in organizations to enhance resilience. The following are four examples of loose coupling that are in line with the aforementioned factors for achieving resilience in organizations:

(i) motivation through task orientation,
(ii) higher-order autonomy,
(iii) flexible changes between organizational modes, and
(iv) culture as basis for coordination/integration.
2.8 Resilience and Risk Management

Risk management deals with the identification of threats and vulnerabilities to determine risk and to apply mitigation strategies that reduce or avoid that risk. On the other hand, resilience is the ability of the system to cope with disruptive events so that the system can resume its regular functionality. Moody (2007), the president of the Critical Infrastructure Protection (CIP) program, related risk management and resilience in the following statement:

> Resilience is the empowerment of being aware of your situation, your risks, vulnerabilities and current capabilities to deal with them, and being able to make informed tactical and strategic decisions.

Risk management and resilience go hand in hand. Risk management procedures are used for the identification of schemes that enhance resilience based on the vulnerabilities of the system and their potential impact. Implementation of the schemes requires an evaluation of the available options in terms of cost-benefit tradeoffs and the implications of current decisions for the future.

Risk management applies measures such as detection and prevention that affect vulnerability while preserving the original inherent state of the system (Haimes 2006). Achieving resilience in systems extends beyond making the system less vulnerable to disruptions; it also entails the implementation of adaptive measures that will enable the system to resume functionality with minimum losses if a threat does impact the system.

In terms of critical infrastructures, Haimes (2006) states that the solution to infrastructure security is a result of protection and resilience. He refers to protection as the set of risk management activities that reduce the vulnerability of the system, whereas resilience includes risk management activities that enable the system to withstand major disruptions within acceptable cost and time limits.
2.9 Summary

This chapter began by defining resilience and asserting why resilience also applies to engineering systems. Resilience is a topic often discussed by stakeholders; the September 11 attacks in the U.S. are in large part responsible for the growing interest in building resilience into systems of all kinds. Although the primary focus of discussion is resilience in engineering and infrastructure systems, resilience finds its roots in ecological systems, in psychology and in materials science.

A review of the literature reveals several definitions of resilient systems. The qualities of a resilient system are illustrated in Fig. 2.10.

![Fig. 2.10. Qualities of a resilient system.](image)

A resilient system can:

- absorb and contain shock,
- reorganize after a disturbance to resume functionality,
- apply adaptive responses in the face of threats,
- cope with hazards, and
- recover quickly from a disruption.
Resilience is a mechanism by which the infrastructure system can prepare itself against disruptive events. The factors that can potentially cause disruptive events have been categorized into human, natural, organizational and technical factors.

Researchers have put forward several proposals for measuring infrastructure resilience, especially after seismic events. Also, several metrics have been suggested to measure the resilience of service infrastructure systems. However, no common framework for measuring resilience has emerged to date.

Resilience is achieved by addressing the two elements of resilience: vulnerability and adaptive capacity. A resilient response reduces the vulnerability of the system and increases its adaptive capacity. Reducing the vulnerability allows the system to absorb shock; increasing its adaptive capacity allows the system to reconfigure and continue functionality in adverse situations.

Infrastructure systems are socio-technical components where the organizational layer of the infrastructure plays a role that is as important as that of the physical layer. The key factors for achieving organizational resilience are leadership, awareness, flexibility and culture.

This chapter also presented an overview of the relationship between resilience and risk management.