Principles for Resilient Design - A Guide for Understanding and Implementation[i]

Scott Jackson[1, 2]

[1]Burnham Systems Consulting – Greater Los Angeles Area

[2]University of South Australia

Contact: jackson@burnhamsystems.net

Keywords: Resilience, Principles, Design, Processes, Heuristics, Measurement, Adversity, Capability

Introduction

In the context of engineered design, according to the International Council on Systems Engineering (2015) resilience is "the ability to provide required capability in the face of adversity". Resilience is somewhat different from risk. While risk pertains to the loss of value due to uncertain future events, resilience has to do with designing a system to maintain a pre-designated level of capability following a disturbance. A key consideration of resilience is the concept of satisficing as described by Adams et al. (2014, p. 118). Satisficing means that the desirable end state of a system is an acceptable level of functionality and that full recovery is not necessarily required.

Resilience Perspectives

Within the study of resilience there are two perspectives: reactive and proactive. Traditionally resilience has been considered to be a reactive concept, that is, the study of the effect on a system following an encounter with a disturbance. Psychology, materials science, and ecology have adopted this perspective. Even some work in engineering has also adopted this perspective, for example, Haimes (2009, pp. 498-501). On the other hand, the study of resilience in an engineering context has adopted the proactive perspective, that is, it considers events prior to the encounter with the threat. Foremost among these sources, the book by Hollnagel et al. (2006, p.36).

Resilience Principles for an Engineered Design

If a current design is not resilient, these are the features that need to be added to the system to make it resilient. There are two types of principles, physical and process principles. In addition, all principles are abstract. As described in Table 1, *physical redundancy*, for example, is an abstract physical principle. All it says is that the system should consist of two identical branches with equal functionality. A communications system, for example, is a concrete system; a communications

[i] This paper is part of the IRGC Resource Guide on Resilience, available at: https://www.irgc.org/risk-governance/resilience/. Please cite like a book chapter including the following information: IRGC (2016). Resource Guide on Resilience. Lausanne: EPFL International Risk Governance Center. *v29-07-2016*

system with two identical and independent branches is a concrete example of a system with physical redundancy. Also described in Table 1, *loose coupling* is an example of an abstract process principle. Electrical power systems are typical concrete systems that incorporate this principle.

Principles
• *Support Principles*
1. Absorption – The system should be capable of withstanding the design level disruption. Hollnagel et al. (2006)
• *Margin* – The design level should be increased to allow for an increase in the disruption. Hollnagel et al. (2006)
• *Hardening* – The system should be resistant to deformation. Richards (2009)
• *Context spanning* – The system should be designed for both the maximum disruption level and the most likely disruption. Madni (2008)
• *Limit degradation* – The absorption capability should not be allowed to degrade due to aging or poor maintenance. Derived; Jackson and Ferris (2013)
2. Restructuring – The system should be capable of restructuring itself. Hollnagel et al. (2006)
• *Authority escalation* – Authority to manage crises should escalate in accordance with the severity of the crisis. Maxwell et al. (2009)
• *Regroup* - The system should restructure itself after an encounter with a threat. Raveh (2008)
3. Reparability – The system should be capable of repairing itself. Richards (2009)
4. Drift correction – When approaching the boundary of resilience, the system should be able to avoid or perform corrective action; action can be taken against either real-time or latent threats.. Hollnagel et al. (2006)
• *Detection* – The system should be capable of detecting an approaching threat. Derived: Jackson and Ferris (2013)
• *Corrective action* – The system should be capable of performing a corrective action following a detection. Derived: Jackson and Ferris (2013)
• *Independent review* – The system should be capable of detecting faults that may result in a disruption at a later time. Derived, Haddon-Cave (2009)
5. Cross-scale interaction – Every node of a system should be capable of communicating, cooperating, and collaborating with every other node. Hollnagel et al.
• *Knowledge between nodes* – All nodes of the system should be capable of knowing what all the other nodes are doing. Billings (1997)
• *Human monitoring* – Automated systems should understand the intent of the human operator. Billings (1997)
• *Automated system monitoring* - The human should understand the intent of the automated system. Billings (1997)
• *Intent awareness* – All the nodes of a system should understand the intent of the other nodes. • Source: Billings (1997)
• *Informed operator* - The human should be informed as to all aspects of an automated system. Billings (1997)

• *Internode impediment* – There should be no administrative or technical obstacle to the interactions among elements of a system. Derived from case studies
6. Complexity Avoidance – The system should not be more complex than necessary. Madni (2009), derived from Perrow (1999)
• *Reduce Variability* – The relationship between the elements of the system should be as stable as possible. Marczyk (2012)
7. Functional redundancy – There should be two or more independent and physically different ways to perform a critical task. Leveson (1995), Madni (2009); Leveson uses the term "design diversity"
8. Physical redundancy – The system should possess two or more independent and identical legs to perform critical tasks. Leveson (1995); Leveson uses the term "design redundancy"
9. Defence in depth – The system should be capable of having two or more ways to address a single vulnerability. Derived from Reason (1997)
10. Human in the loop - There should always be human in the system when there is a need for human cognition. Madni (2009)
• *Automated function* – It is preferable for humans to perform a function rather than automated systems when conditions are acceptable. Billings (1997)
• *Reduce Human Error* – Standard strategies should be used to reduce human error. Derived from Billings (1997) and Reason (1990)
• *Human in Control* – Humans should have final decision-making authority unless conditions preclude it. Billings (1997)
11. Loose Coupling – The system should have the capability of limiting cascading failures by intentional delays at the nodes. Perrow (1999)
• *Containment* – The system will assure that failures cannot propagate from node to node. Derived; Jackson and Ferris (2013)
12. Modularity. Madni (2009), Perrow (2011) the functionality of a system should be distributed through various nodes of that system so that if a single node is damaged or destroyed, the remaining nodes will continue to function.
13. Neutral State – Human agents should delay in taking action to make a more reasoned judgement as to what the best action might be. Madni (2009)
14. Reduce Hidden Interactions – Potentially harmful interactions between elements of the system should be reduced. Derived from Leveson (1995) and Perrow (1999)

Table 1: Resilience principles and sources

Most of the principles are in reality heuristics rather than scientifically accepted principles meaning that they are practices adopted by experts in the various domains based on their experience. As heuristics the designer can be confident that they will be effective most of the time but not all of the time. Hence each principle can be expected to exhibit a vulnerability meaning that its incorporation may actually lead to occasional failures. This fact leads to the incorporation of one of the most important principles, the principle of *defence in depth*, meaning that second and third principles may be required to compensate for the vulnerabilities of the primary principle.

Jackson and Ferris (2013, pp. 152-164) have identified the principles that include both architectural physical and process principles. This paper also identifies those principles that would most likely be used as backup principles. They are called dependent principle. Among these there are 14 primary

principles and 20 support principles. The primary principles are applicable across the broadest range of domains and scenarios. The support principles are a subset of the primary principles, that is, they apply to a defined limited set of conditions. The following paragraphs discuss two of the most important principles: *absorption* and *restructuring*. The reader is referred to the Jackson and Ferris paper for a more exhaustive discussion of the principles and the support principles. Table 1 presents the complete set of primary and support principles.

Regarding the *absorption* principle, in very few cases can the system absorb all possible threat levels. That is when the *defence in depth* principle come into play. A good example is the US Airways Flight 1549, also known as the Miracle on the Hudson case described by Pariès (2011, pp. 9-27). In this case the aircraft was unable to absorb the flock of geese that it had struck (the *absorption* principle), so it was forced to employ the *functional redundancy* principle (alternative sources of power and control) and the *human in the loop* principle (the pilot). The result was an example of satisficing in which the aircraft was lost while the humans were saved.

The best example of the use of the *restructuring* principle is when the authorities in New York were able to deploy a spontaneous power system after the World Trade Centre attacks to restore power within five hours as described by Mendoça and Wallace (2006, pp. 209-219).

Finally, after having identified all the possible principles, how does the analyst decide which ones constitute a solution for a resilient system for a specific case? The answer to that question requires a little more analysis. The reader is referred to the paper by Jackson, Cook and Ferris (2015). This paper describes the path of a system from a nominal operational state to a final acceptable and satisficing state. This path is called a state-transition analysis. It defines seven possible states in which the system can exist and 28 possible transitions from state to state. Each transition will require the employment of one or more principles. In reality the number of practical principles will be very small, at least it is hoped. In short these principles will constitute the candidate principles for a final solution and will be the inputs to a simulation to determine the most appropriate one for a given scenario.

Measurement of Resilience

The measurement of resilience is dependent on what information is available and when is it available. This section defines four stages and the level of measure that may be possible in each stage.

Stage 1 – A system exists and no improvements have been made. Its vulnerability is well known from events in the past. For example, prior to the San Francisco Earthquake and Fires in 1906 the city lacked a redundant water system. The events of 1906 left the city without water with which to extinguish the fires. In agreement with Haimes (2009), measurement of resilience in this phase would be very difficult since the characteristics of the system are unknown.

Stage 2 – In this phase resilience principles have been invoked to improve the resilience of the system. The most useful metrics are the principles that result from the recommendations of experts in each domain. In short a compilation of these recommendations and their frequency would constitute a valid and useful metric.

Stage 3 - At this point specific designs will have been defined with the appropriate principles incorporated in them. These characteristics will have been incorporated into a computer model to simulate the encounter with the threat and the resulting condition of the system being defined. In

addition, threat and the encounter of the threat will also be in the model. This is the stage at which the highest quality metrics would be possible. Ganin et al. (2016) have proposed metrics which can be quantified and evaluated during this phase.

Stage 4 - At this point the system will have been built and it will have encountered the predicted threat. It will be possible to determine what actually happened and how much of the system including humans survived. Unfortunately there are very few cases like this. In short, it cannot be expected that many good metrics will come out of this phase.

Annotated Bibliography

Adams, K. M., Hester, P. T., Bradley, J. M., Meyers, T. J., & Keating, C. B. (2014). Systems Theory as the Foundation for Understanding Systems. *Systems Engineering, 17*(1), 112-123.
This paper is a survey of the various principles of systems theory and they apply to systems. Many of these principles apply directly to the study of resilience.

Billings, C. (1997). *Aviation Automation: The Search for Human-Centered Approach*. Mahwah, NJ: Lawrence Erlbaum Associates.
This book documents several principles related to the interface between humans and automated systems. The focus of this book was aviation systems, but the principles can apply to any system in which humans interface with automated systems.

Ganin, A. A., Massaro, E. M., Gutfraind, A., Steen, N., Keisler, J. M., Kott, A., et al. (2016). *Operational resilience: concepts, design, and analysis*.
This report suggests analytic measures for the resilience of complex infrastructure systems.

Haddon-Cave, C. (2009). *An Independent Review of the Broader Issues Surrounding the Loss of the RAF Nimrod MR2 Aircraft XV230 in Afganistan in 2006*. London: The House of Commons.
This report examines the 2006 Nimrod accident and the establishment of additional rigour in the safety and airworthiness processes.

Haimes, Y. Y. (2009). On the Definition of Resilience in Systems. *Risk Analysis, 29*(43), 498-501.
This paper was written in the context of infrastructure resilience for the US Homeland Security for which the reactive perspective is more common.This paper also stresses the difficulty in measuring resilience.

Hollnagel, E., Woods, D. D., & Leveson, N. (Eds.). (2006). *Resilience Engineering: Concepts and Precepts*. Aldershot, UK: Ashgate Publishing Limited.
This book was a pioneering effort in the adaptation of the concept of resilience to the world of engineering. It contains chapters written by several experts in this field. For the most part this book focuses on organisations rather than physical systems. This book is the source of several of the principles discussed in this paper.

Jackson, S., & Ferris, T. (2013). Resilience Principles for Engineered Systems. *Systems Engineering, 16*(2), 152-164.
This paper is a compilation of resilience principles from various sources including the Hollnagel

et al. (2006) book. It also discusses the concept of the vulnerability of principles and the dependency among them.

Jackson, S., Cook, S. C., & Ferris, T. (2015). *Towards a Method to Describe Resilience to Assist in System Specification*. Paper presented at the IS 2015.
This paper outlines the state-transition analysis for determining the most appropriate design for a resilient system.

Leveson, N. (1995). *Safeware: System Safety and Computers*. Reading, Massachusetts: Addison-Wesley.
This book was the sources of several principles including physical redundancy, functional redundancy, and reduce hidden interactions.

Madni, A. (2008). Suggested Heuristics and Validation. In S. Jackson (Ed.) (Various suggested resilience heuristics and their validation. This item covers several conversations. ed.). Los Angeles, CA.
Professor Madni suggested several principles including the neutral state principle and the context spanning support principle.

Madni, Azad,, & Jackson, S. (2009). Towards a conceptual framework for resilience engineering. *Institute of Electrical and Electronics Engineers (IEEE) Systems Journal, 3*(2), 181-191.
This paper was an early source of many of the principles found in this paper.

Marczyk, J. (2012). Complexity Reduction. In S. Jackson (Ed.) (Email communication ed., pp. 1). Como, Italy.
Dr. Marczyk shared his work on complexity including the importance of both the structural aspects of complexity and the uncertainty between elements of the system known as Shannon entropy.

Maxwell, J., & Emerson, R. (2009, July 2009). Observations of the resilience architecture of the firefighting and emergency response infrastructure. *Insight,* 45-46.
This paper was a result of a joint meeting between the INCOSE Resilient Systems Working Group and several senior members of the San Francisco Fire Department. One principle that emerged from this meeting was the authority escalation support principle.

Mendoça, D., & Wallace, W. (2006). *Adaptive capacity: Electric power restoration in New York City following the 11 September 2001 attacks*. Paper presented at the Second Resilience Engineering Symposium.
This paper is a description of how a resilient system was created during a crisis environment following the 9/11 event.

Pariès, J. (2011). Lessons from the Hudson. In E. Hollnagel, J. Pariès, D. D. Woods & J. Wreathhall (Eds.), *Resilience Engineering in Practice: A Guidebook*. Farnham, Surrey: Ashgate Publishing Limited.
This is a chapter in Hollnagel's second book and is a discussion of the US Airways Flight 1549 and its significance to resilience especially with respect to the defence in depth principle.

Perrow, C. (1999). *Normal Accidents: Living With High-Risk Technologies*. Princeton, NJ: Princeton University Press.
This book is the source of several principles including complexity avoidance and loose coupling.

Perrow, C. (2011). Modular Power Grids. In S. Jackson (Ed.). New Haven.
Via personal communications Dr. Perrow stressed the importance of modularity in electrical grid systems.

Raveh, A. (2008). Regroup Heuristic. In S. Jackson (Ed.) (Comment during tutorial on resilience ed.). Utrecht, the Netherlands.
At a tutorial on resilience, Mr. Raveh suggested the regroup support principle.

Reason, J. (1990). *Human Error*. Cambridge: Cambridge University Press.
This book provides a detailed list of how human error can be avoided. This information was the source of the reduce human error support principle.

Reason, J. (1997). *Managing the Risks of Organisational Accidents*. Aldershot, UK: Ashgate Publishing Limited.
This book was the primary source of the defence in depth principle which was inferred from the Swiss cheese model of the author.

Richards, M. G. (2009). *Multi-Attribute Tradespace Exploration for Survivability.* Unpublished Dissertation, Massachusetts Institute of Technology, Cambridge, MA.
This thesis provides a view of resilience from a defence point of view. The reparability principle was the primary contribution.