Four concepts for resilience and the implications for the future of resilience engineering

Article in Reliability Engineering & System Safety - April 2015
DOI: 10.1016/j.ress.2015.03.018

1 author:

David D Woods
The Ohio State University
420 PUBLICATIONS 22,512 CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project
Patient Safety with HIT View project

Project
Innovation of emergency medical service system to enhance resilience in disaster risk reduction View project
Four concepts for resilience and the implications for the future of resilience engineering

David D. Woods
Initiative on Complexity in Natural, Social & Engineered Systems, The Ohio State University, United States

Keywords:
Resilience engineering
Resilience
Sustainability
Resilient control
Robust control
Complexity
Complex adaptive systems
Socio-technical systems

1. Introduction

Today’s systems exist in an extensive network of interdependencies as a result of opportunities afforded by new technology and by increasing pressures to become faster, better and cheaper for various stakeholders. But the effects of operating in interdependent networks has also created unanticipated side effects and sudden dramatic failures [42,1]. These unintended consequences have led many different people from different areas of inquiry to note that some systems appear to be more resilient than others. This idea that systems have a property called ‘resilience’ has emerged and grown extremely popular in the last decade (for example, articles in scientific journals on the topic of resilience increased by an order of magnitude between 2000 and 2013 based on search of Web of Science, e.g., Longstaff et al. [26]). The idea arose from multiple sources and has been examined from multiple disciplinary perspectives including: systems safety (see Hollnagel et al. (2006)), complexity (see [1]), human organizations (see [42,40,22,32,31]), ecology (see [41]), and others. However, with popularity has come confusion as the label continues to be used in multiple and diverse ways.

As multiple observers from different disciplines began to study the characteristics that affect the ability to create, manage, and sustain resilience, four core concepts appear and recur. This paper organizes the diverse uses of the label ‘resilience’ into groups based on these four conceptual perspectives. The paper refers to these four concepts as resilience [1] through [4]. First, people use the label resilience to refer to how a system rebounds from disrupting or traumatic events and returns to previous or normal activities (rebound = resilience [1]).

Second, people use the label resilience as the equivalent to the concept of system robustness. These two concepts have recurred repeatedly in work on resilience, especially in the early stages of exploring how systems manage complexity as they appear to provide a path to generate explanations of how some systems are able to manage increasing complexity, stressors, and challenges (robustness = resilience [2]).

As researchers have continued to study the problem of complexity and how systems adapt to manage complexity, two additional concepts have emerged. Upon further inquiry, the empirical results begin to reveal how some systems overcome the risk of brittleness, i.e., the risk of a sudden failure when events push the system up to and beyond its boundaries for handling changing disturbances and variations [7,43,44]. From the perspective of overcoming the risk of brittleness, a third use of the label resilience becomes the idea of graceful extensibility [47,45] — how a system extends performance, or brings extra adaptive capacity to bear, when surprise events challenge its boundaries (graceful extensibility = resilience [3]).

Another line of inquiry has pursued formal models of systems that have proved to be evolvable in biology and technology (e.g., the internet). A fourth use of the label resilience emerged from this work that focuses on the question: what are the architectural properties of layered networks that produce sustained adaptability—the ability to adapt to future surprises as conditions continue to evolve? [14,32,31]. This line of work centers on how networks can manage fundamental trade-offs that constrain all systems
It seeks to identify governance policies that operate across layered networks in biological systems, social systems, and technological systems—what governance policies sustain the ability of the network to continue to function well and avoid falling into traps in the trade spaces as conditions change over long time scales (sustained adaptability = resilience [4]).

This paper briefly considers each of the four, in turn, to explore how each has stimulated lines of inquiry and led to new and sometimes unexpected results. The intent of the paper is to set a new baseline for future work. Whatever the historical contributions of each of these four concepts, the question is how to advance productive lines of inquiry. Organizing the numerous and continuing attempts to define resilience around these four concepts blocks out a great deal of noise (see the overview in [27]). The review of the four concepts sets the stage to debate which concepts have the potential to continue to advance our understanding of complex adaptive systems.

2. Four concepts for resilience

2.1. Resilience as rebound (or resilience [1])

The rebound concept begins with the question: why do some communities, groups, or individuals recover from traumatic disrupting events or repeated stressors better than others to resume previous normal functioning? A representative example of this approach is a recent compilation of papers assembled when an organization asked the Institute of Medicine to help it answer the above question [6]. We also find this question asked by business continuity centers as organizations confront extreme weather events that can produce surprising cascades of effects [11].

This use of the label resilience as [1] = rebound – is common, but pursuing what produces better rebound merely serves to restate the question. Where progress has been made, the focus is not on the period of rebound but on what capabilities and resources were present before the rebound period. Finkel’s analysis of contrasting cases of recovery from or inability to recover from surprise provides compelling evidence [16]. First, it is not what happens after a surprise that affects ability to recover; it is what capacities are present before the surprise that can be deployed or mobilized to deal with the surprise. This issue was noted early on by Lagadec with respect to major external trigger events [20, p. 54]: “the ability to deal with a crisis situation is largely dependent on the structures that have been developed before chaos arrives. The event can in some ways be considered a brutal and abrupt audit: at a moment’s notice, everything that was left unprepared becomes a complex problem, and every weakness comes rushing to the forefront”.

Second, rebound considers responses to specific disruptions, but much more importantly the disrupting events represent surprises, that is, the event is a surprise when it falls outside the scope of variations and disturbances that the system in question is capable of handling [43,46]. In other words, the key is not simply the attributes of the event in itself as a disruption or its frequency of occurrence, but how the event challenges a model instantiated in the base capabilities of that system. The surprise event challenges the model and triggers learning and model revision—a kind of model surprise [48]. There are patterns to surprise, or, as Nemeth puts it, there are regularities to what on the surface appears to be irregular variations in terms of how disturbances challenge normal functioning [30].

These two points highlight a paradox about resilience, that shifts the focus from resilience [1] to resilience [3] (graceful extensibility) as research begins to consider resilience as multiple forms of adaptive capacity. To overcome the risk of brittleness in the face of surprising disruptions requires a system with the potential for adaptive action in the future when information varies, conditions change, or when new kinds of events occur, any of which challenge the viability of previous adaptations, models, plans, or assumptions. However, the data to measure resilience as this potential comes from observing/analyzing how the system has adapted to disrupting events and changes in the past [44].

There are other limits to the line of inquiry based on resilience [1], for example, the concept of recovery to normal or previous function (return to equilibrium) has not held up to inquiry (see for example, [41]). The process of adapting to disruptions, challenges and surprises over time changes the system in question in multiple ways. In adapting to new challenges, systems draw on their past but become something new. Even when adapting to preserve, the process of adapting transforms both the system and its environment. Continuity occurs over a lineage of challenge and adaptive response, a series of adaptive cycles that compose an adaptive history.

It is historically interesting that questions about resilience are often formulated around finding a way to explain variations in how systems rebound from challenge. But research progress has left this framing behind to focus on the fundamental properties of networks, systems and organizations that are able to build, modify and sustain the right kinds of adaptive capacities [14]. Studies of biological systems [17] and evolutionary computational modeling of biological systems [23,24] have shown that properties that will sustain adaptive capacity in the future can be selected for [4]. These are examples of results that shift in focus the focus from resilience [1] to resilience [4] —architectures for sustained adaptability.

2.2. Resilience as robustness (or resilience [2])

Resilience [2] – increased ability to absorb perturbations – confounds the labels robustness and resilience. Some of the earliest explorations of resilience confounded these two labels, and this confound continues to add noise to work on resilience (as noted in [43,29]).

An increase in robustness expands the set of disturbances the system can respond to effectively. This simple definition is the basis for the success in robust control as a subset of control engineering [15], “Robust control is risk-sensitive, optimizing worst case (rather than average or risk-neutral) performance to a variety of disturbances and perturbations” ([14, p. 15624]). Alderson and Doyle [1] point out that robustness is always of the form: system X has property Y that is robust in sense Z to perturbation W. In other words, robust control works, and only works, for cases where the disturbances are well-modeled.

If an increase in robustness expands the set of disturbances the system can respond to effectively, the question remains what happens if the system is challenged by an event outside of the current set? If the system cannot continue to respond to demands and meet some of its goals to some degree, then the system will experience a sudden failure or collapse – that is, the system is brittle at its boundaries—resilience [3]. In other words, resilience comes to the fore when the set disturbances is not well modeled and when this set is changing. And ironically, the set of poorly modeled variations and disturbances changes based on a record of past success which triggers adaptive responses by other nearby units in the layered network of interdependent systems. As a result of this fundamental result, and in a direct analogy to robust control, a new line of inquiry has emerged to develop resilient control systems for applications such as cybersecurity and cyber-physical systems (e.g., [36]).

Confounding resilience and robustness turns out to be erroneous in another way. If an increase in robustness expands the set

Please cite this article as: Woods DD. Four concepts for resilience and the implications for the future of resilience engineering. Reliability Engineering and System Safety (2015), http://dx.doi.org/10.1016/j.ress.2015.03.018
of disturbances the system can respond to effectively, the usual assumption is that this performance envelope only grows larger or more encompassing. But Doyle and colleagues have shown formally and theoretically (e.g., [9]) and safety research has shown empirically [43,19] that this simple expansion is not what happens. Instead, expanding a system's ability to handle some additional perturbations, increases the systems vulnerability in other ways to other kinds of events.

This is a fundamental trade-off for complex adaptive systems where becoming more optimal with respect to some variations, constraints, and disturbances increases brittleness in the face of variations, constraints, and disturbances that fall outside this set [1,18]. The search for good system architectures studies how some systems are able to continue to solve the trade-off as load increases [14,25]. A converging line of evidence comes from studies of human systems that escape from the tragedy of the commons [12,31,22]. The emerging understanding of heuristic and formal architectural principles points us to the fourth concept for resilience as some architectures are able to sustain the ability to adapt to future surprises over multiple cycles of change, or resilience [4].

2.3. Resilience as graceful extensibility (or resilience [3])

The third concept sees resilience as the opposite of brittleness, or, how to extend adaptive capacity in the face of surprise [46,47,7]. Resilience [3] juxtaposes brittleness versus graceful extensibility. Rather than asking the question how or why do people, systems, organizations bounce back, this line of approach asks: how do systems stretch to handle surprises? Systems with finite resources in changing environments are always experiencing and stretching to accommodate events that challenge boundaries. And what systems escape the constraints of finite resources and changing conditions?

Without some capability to continue to stretch in the face of events that challenge boundaries, systems are more brittle than stakeholders realize [45]. And all systems, however successful, have boundaries and experience events that fall outside these boundaries—surprises. Brittleness describes how a system performs near and beyond its boundary, separate from how well it performs when operating well within its boundaries. Descriptively and specifically, brittleness is how rapidly a system's performance declines when it nears and reaches its boundary. Brittle systems experience rapid performance collapses, or failures, when events challenge boundaries. Of course, one difficulty is that the location of the boundary is normally uncertain and moves as capabilities and conditions change.

There is always some rate and kind of events that occur to challenge the boundaries of more or less optimal or robust performance, and thus graceful extensibility, being prepared to adapt to handle surprise, is a necessary form of adaptive capacity for all systems [43,45]. Systems with low graceful extensibility risk collapse at the boundaries. But surprise has regular characteristics as many classes of challenge re-curr (e.g., [30]) which can be tracked and used as signals for adaptation. Caporale and Doyle express the point in the context of biological systems [4, p. 20]:

“However, many classes of environmental challenge re-curr. Hosts combat pathogens (and pathogens avoid host defenses); predators and prey do battle through biochemical adaptations; bird’s beaks must pick up and crack available seeds (or insects)—a menu that may change rapidly due, for example, to a drought.”

Challenges such as cascades of disturbances and friction in putting plans into time are generic classes of demands that require the ability to extend performance to avoid collapse due to brittleness [47].

Attempts to expand the base envelope (the competence envelope or base adaptive capacity) shift the dynamics and kinds of events that challenge the new boundaries (and how they challenge the boundaries). This process of change means that graceful extensibility is a dynamic capability. Graceful extensibility is a play on the traditional term—graceful degradation. However, graceful degradation only refers to breakdowns. Woods [45] uses graceful extensibility because adaptation at the boundaries can be very positive and lead to success, not simply less negative capability. Systems with high graceful extensibility have capabilities to anticipate bottlenecks ahead, to learn about the changing shape of disturbances and possess the readiness-to-respond to adjust responses to fit the challenges [16,46,48].

From the point of view of resilience [3], attempts to understand rebound, first, should change direction: search for previous disrupting events and analyze what the system drew on to stretch to accommodate those kinds of past events. Observing/analyzing how the system has adapted to disrupting events and changes in the past provides the data to assess that system's potential for adaptive action in the future when new variations and types of challenges occur [44]. Many studies of these kinds of adaptive cycles have identified basic patterns and empirical generalizations (recent examples are [8,28,3,33–35,37,39]).

Second, the desire to understand rebound should lead to studies and models of the consequences when a system has to stretch repeatedly to multiple challenges over time. Calling on resources to stretch repeatedly can overwork a system's readiness-to-respond capability, resulting in consequences associated with stress (e.g., in material science over-stressing a material changes that material and its ability to respond to challenges in the future).

Studies of how systems extend adaptive capacity to handle surprise have led to characterization of basic patterns in how adaptive systems succeed and fail [47]. The starting point is exhausting the capacity to deploy and mobilize responses as disturbances grow and cascade—this pattern is called decompensation. The positive pattern observed in systems with high graceful extensibility is anticipation of bottlenecks and crunches ahead.

Decompensation as a form of adaptive system breakdown subsumes a related finding called critical slowing down, where an increasing delay in recovery following disruption or stressor is an indicator of an impending collapse or a tipping point [38,10]. When the time to recovery increases and/or the level recovered to decreases, this pattern indicates that a system is exhausting its ability to handle growing or repeated challenges, in other words, the system is nearing saturation of its range of adaptive behavior. Risk of saturation signals the risk of the basic decompensation failure pattern. Risk of saturation turns out to play a key role in graceful extensibility as a basic form of adaptive capacity ([47,25,38,44]).

There are many other indicators of the risk of decompensation, and studies of systems that reduce the risk of decompensation provide valuable insight about where to invest to reduce brittleness/increase resilience [3]. For example, Finkel [16] identified characteristics of human systems that produce the ability to recover from surprise. Interestingly, these characteristics or sources of resilience represent the potential for adaptive action in the future. Sources of resilience [3] provide a system with the capability, in advance, to handle classes of surprises or challenges such as cascading events. Providing and sustaining these sources of resilience [3] has its own dynamics and difficulties that arise from fundamental trade-offs—resilience [4] [43,19,1]. For example, work has found that organizations can undermine, inadvertently, their own sources of resilience as they miss how people step into the breach to make up for adaptive shortfalls [43].
2.4. Resilience as sustained adaptability (or resilience [4])

Resilience [4] refers to the ability manage/regulate adaptive capacities of systems that are layered networks, and are also a part of larger layered networks, so as to produce sustained adaptability over longer scales [1]. Some layered networks, or complex adaptive systems demonstrate sustained adaptability, but most layered networks do not, i.e., they get stuck in adaptive shortfalls, unravel and collapse when confronting new periods of change, regardless of their past record of successes. Resilience [4] asks three questions: (1) what governance or architectural characteristics explain the difference between networks that produce sustained adaptability and those that fail to sustain adaptability? (2) What design principles and techniques would allow one to engineer a network that can produce sustained adaptability? (3) How would one know if one succeeded in their engineering (how can one confidently assess whether a system has the ability to sustain adaptability over time, like evolvability from a biological perspective and like a new kind of stability from a control engineering perspective)?

In socio-technical systems, sustained adaptability addresses a system’s dynamics over a life cycle or multiple cycles. The architecture of the system needs to be equipped at earlier stages with the wherewithal to adapt or be adaptable when it will face predictable changes and challenges across its life cycle. Predictable dynamics of challenge include:

- Over the life cycle, assumptions and boundary conditions will be challenged—surprises will continue to re-cur.
- Over the life cycle, conditions and contexts of use will change—therefore boundaries will change, especially if the system provides valuable capability to stakeholders.
- Over the life cycle, adaptive shortfalls will occur and some responsible people will have to step in to fill the breach.
- Over the life cycle, the need for graceful extensibility and the factors that produce or erode graceful extensibility will change, more than once.
- Over life cycles, classes of changes will occur, and the system in question will have to adapt to seize opportunities and respond to challenges by readjusting itself and its relationships in the layered network.

Central to resilience [4] is identifying what basic architectural principles are preserved over these changes and provide the needed flexibility to continue to adapt over long scales [14]. Advances on resilience [4] center on the finding that all adaptive systems are subject to fundamental constraints or trade-offs, that there are multiple trade-offs, and that there are basic architectural principles that allow some systems to adjust their position in the multi-dimensional trade space in ways that tend to move toward or find new positions along hard limit lines [14,25]. Prominent in this line of inquiry are questions about which trade-offs are fundamental and whether these are different for human systems as compared to biological or physical systems at various scales [13,18].

Resilience [4] also leads to the agenda to define resilient control mechanisms, i.e., control or management of adaptive capacities relative to the fundamental trade-offs. Thus, resilience [4] is a higher level concept in which multiple dimensions are balanced and traded off, given the laws that constrain how (human) adaptive systems work. In resilience [4] it makes sense to say a system is resilient, or not, based on how well it balances all the tradeoffs, or not. For example, success stories can be found in biology if we look at glycolysis as modeled by Chandra et al. [5], or selection for future adaptive capacity (as in [24]), and in human systems success stories can be found in the work of Finkel [16] on how successful military systems prepare to adapt to surprise, Ostrom on how human networks avoid the tragedy of the commons through polycentric governance principles as in examples such as managing limited water resources in Bali [32,12,21]. Progress is being made on mechanisms for resilient control in infrastructures (e.g., [2]) and in regulating the risk of brittleness (e. g. by regulating a system’s capacity for maneuver to handle potential upcoming surprises in [47,45]).

3. Implications for resilience engineering

As different people and disciplines pursue their journey of inquiry about complex systems and reducing risks of sudden failure in complex systems, a progression of concepts recur that capture different senses of the label resilience. This paper has organized the various senses and definitions into four groups: rebound, robustness, graceful extensibility, and architectures for sustained adaptability. This partition represents four core concepts that have recurred since the introduction of resilience as a critical systems property. This partition allows an assessment of progress and a projection of what is promising to create the ability to engineer resilience into diverse systems and networks in the future.

The first implication of the partition is that, through overuse, the label resilience only functions as a general pointer to one or another of the four concepts. For science and engineering purposes, one needs to be explicit about which of the four senses of resilience is meant when studying or modeling adaptive capacities (or to expand on the four anchor concepts as new results emerge).

Second, the value of the differing concepts depends on how they are productive in steering lines of inquiry toward what will prove to be fundamental findings, foundational theories, and engineering techniques. The yield from first two concepts about resilience, rebound and robustness, has been low. Resilience as rebound misdirects inquiry to reactive phases and restoration or return to previous states. It begs the question on what is needed in advance of a challenge event or shift in variations and disturbance, and how systems continue to change as they adapt, as well as how systems provoke changes through adaptation.

Confounding resilience and robustness begs the question of how systems and networks adapt when faced with poorly modeled events, disruptions, and variations. Control engineering already knows a great deal about how to engineer systems to handle well-modeled disturbances. The lines of inquiry relevant to resilience are about how systems and networks can be prepared to handle the model surprises that occur as change is ongoing. The empirical progress has come from finding, studying, and modeling the biological and human systems that are prepared to handle surprises.

The value of these two concepts is historical as they were the first approaches used to tackle issues related to resilience and stimulated multiple lines of inquiry. The disappointment is that both of these concepts continue to be recycled, both in reference to past work and in current efforts, as if they provide an adequate conceptual basis to move forward. Nevertheless, the lines of inquiry have progressed to tackle questions such as:

- how adaptive systems fail in general and across scales;
- how systems can be prepared for inevitable surprise while still meeting pressures to improve on efficiency of resource utilization;
- what mechanisms allow a system to manage the risk of brittleness at the boundaries of normal function;
- what architectures allow systems to sustain adaptability over long times and multiple cycles of change.

Please cite this article as: Woods DD. Four concepts for resilience and the implications for the future of resilience engineering. Reliability Engineering and System Safety (2015), http://dx.doi.org/10.1016/j.ress.2015.03.018
Studies of resilience in action have revealed a rich set of patterns and regularities about how some systems provide and adjust graceful extensibility to overcome brittleness. Models on what makes the difference between resilience and brittleness have been successful in specific areas to highlight fundamental processes that sustain adaptability over long scales. As a result, we can characterize different kinds of adaptive capacities, dynamic patterns about how these capacities develop or degrade, and the kinds of architectures that support or sustain the ability to adapt to future challenges.

However, the multiple lines of inquiry that intersect around the label resilience are young. The end story remains to be written of how to engineer in graceful extensibility and how to design architectures that will sustain adaptive capacities over time.

References