Environment Systems and Decisions (2018) 38:433–457
https://doi.org/10.1007/s10669-018-9708-3

The theory of graceful extensibility: basic rules that govern adaptive systems

David D. Woods[1]

Published online: 10 September 2018

Abstract
The paper introduces the theory of graceful extensibility which expresses fundamental characteristics of the adaptive universe that constrain the search for sustained adaptability. The theory explains the contrast between successful and unsuccessful cases of sustained adaptability for systems that serve human purposes. Sustained adaptability refers to the ability to continue to adapt to changing environments, stakeholders, demands, contexts, and constraints (in effect, to adapt how the system in question adapts). The key new concept at the heart of the theory is graceful extensibility. Graceful extensibility is the opposite of brittleness, where brittleness is a sudden collapse or failure when events push the system up to and beyond its boundaries for handling changing disturbances and variations. As the opposite of brittleness, graceful extensibility is the ability of a system to extend its capacity to adapt when surprise events challenge its boundaries. The theory is presented in the form of a set of 10 proto-theorems derived from just two assumptions—in the adaptive universe, resources are always finite and change continues. The theory contains three subsets of fundamentals: managing the risk of saturation, networks of adaptive units, and outmaneuvering constraints. The theory attempts to provide a formal base and common language that characterizes how complex systems sustain and fail to sustain adaptability as demands change.

Keywords Resilience · Resilience Engineering · Complex adaptive systems · Human systems integration · Adaptability · Complexity · Socio-technical systems · Agility · Resilient control · Sustainability · Robust yet fragile · Resilient infrastructures

1 Introduction

1.1 The mystery of sustained adaptability

Control systems are everywhere regulating processes to meet targets in the service of human goals. Science and engineering advances over decades have developed the theory and practice, and linked the two together to field many forms of controllers from the practical PID controller (proportional, integral, derivative) to optimal, adaptive, and robust controllers (Zhou et al. 1996; Narendra and Annaswamy 2005; Astrom and Murray 2008). However, the forms of control available today are not powerful enough to account for successful cases of sustained adaptability in biology (such as glycolysis; Chandra et al. 2011), engineered systems (such as the internet; Doyle et al. 2005), and human systems (such as Balinese water temple networks; Lansing and Kremer 1993). In these cases, multiple interacting and interdependent processes continuously re-adjust to each other as they cope with, and as they exploit, changing demands, contexts, and constraints (Meyers and Bull 2002). In these and other cases, complex systems, in the sense of networks with extensive and sometimes hidden interdependencies, adapt in the face of variation, but much more importantly, are able to sustain adaptability as the forms and sources of variation continue to change over longer cycles. In shorthand, the underlying architecture of these layered networks facilitates future adaptations as conditions change and challenge the fitness of past adaptations.

These successful cases can be contrasted with much less successful adaptive networks in complex settings where initially successful adaptations unwind over time—become stale, work at cross purposes, are unable to keep pace with change and cascades—and suffer sudden performance collapses (Scheffer et al. 2009; Vespignani 2010; Woods and Branlat

✉ David D. Woods
woods.2@osu.edu

[1] Department of Integrated Systems Engineering, The Ohio State University, 1971 Neil Ave, Columbus, OH, USA

2011). The failures to sustain adaptability stand out when we witness an ecosystem degrade in a tragedy of the commons (Dietz et al. 2003), a matching market unravel (Roth 2008), or a dramatic failure in safety–critical operations (Woods 2005).

The mystery of sustained adaptability refers to the finding that the complexity arises from changes to increase optimality and robustness of a network of adaptive units which leads to an emergent susceptibility to sudden performance collapses and failures (Carlson and Doyle 2000; Csete and Doyle 2002; Ormerod and Colbaugh 2006). The effort invested to improve fitness leads to systems "which are robust to perturbations they were designed to handle, yet fragile to unexpected perturbations and design flaws" (Carlson and Doyle 2000, p. 2529). The network will look more and more fit to its environment on some criteria, while, the same processes produce severe brittleness when events occur that challenge the design envelope. Ormerod and Colbaugh summarize simulation results: "as the connectivity of a network increases, we observe an increase in the average fitness of the system. But at the same time, there is an increase in the proportion of failure/extinction events which are extremely large." Nevertheless, there are cases of biological and human systems that are able to counter the brittleness that inevitably grows for layered networks with extensive interdependencies. These systems demonstrate the ability to continue to adapt to changing environments, stakeholders, demands, contexts, and constraints—in effect, sustained adaptability refers to the ability adapt how the system in question adapts (Woods 2015). For example, mechanisms that facilitate future adaptation continue to be found in biological systems (e.g., Meyers and Bull 2002; Beaumont et al. 2009).

Possible answers to the mystery of sustained adaptability should have some linkage and continuity with concepts about control systems. Control systems are in many ways a simple form of adaptive system, and theory specifies how to ensure stability (adequate adaptive performance) given well-defined targets and well-modeled disturbances (e.g., Narendra and Annaswamy 2005). But for layered networks of interdependent adaptive units, the architectural principles that will produce sustained adaptability over cycles of change remain to be fully worked out (Alderson and Doyle 2010; Doyle and Csete 2011; Seager et al. 2017). Rieger (2010) and Alderson et al. (2015) call the issue the problem of resilient control in a *layered network* of interdependent adaptive units. Research emphasizes how these layered networks have hidden interdependencies, operate at multiple varying tempos, with significant cross-layer interactions (e.g., the case in Mendonca and Wallace 2015). These factors lead me to use the more vivid label in this work of *Tangled Layered Networks* (TLNs).

Sustained adaptability is also a problem in coordination across multiple human roles in a network where the roles interact and adapt to each other vertically across multiple echelons or layers, as well as horizontally over different spatial and temporal scopes of responsibility (Park et al. 2013). The contrast between successful and unsuccessful cases of sustained adaptability in multi-role, multi-echelon (layered) human systems has been studied from the perspectives of organizational dynamics, anthropology, environmental systems, sociology, (macro-)cognitive systems, and experimental micro-economics. Following Ostrom, research on sustained adaptability in human systems has been labeled as the problem of poly-centric governance (Ostrom 2012). Combining results on sustained adaptability as a new kind of resilient control problem and on sustained adaptability as a coordination problem in poly-centric human systems requires bridging completely different forms of inquiry centered on (1) understanding empirically what makes the difference between successful and unsuccessful examples of sustained adaptability and, critically, (2) developing actionable tactics and strategies to design architectures for TLNs that will demonstrate sustained adaptability.

Cutting across these lines of inquiry are a set of common questions. What are the differences between the successful and unsuccessful cases? What factors are critical when adaptability is sustained? Can research extract general principles that tend to generate sustained adaptability, or even guarantee it? A growing number of empirical studies contrast successful and unsuccessful cases and together provide partial answers about underlying patterns. One setting where the contrast has been examined extensively is critical care medicine such as emergency departments and intensive care units, see for example, Cook (2006), Nemeth et al. (2007), Miller and Xiao (2007), Wears et al. (2008), Perry and Wears (2012), and Patterson and Wears (2015). For other empirical samples from disaster response, critical digital infrastructure, and military history see Mendonça and Wallace (2015), Robbins et al. (2012), and Finkel (2011).

Asking the above questions—questions about how adaptability can be sustained over time, over patterns of change, and as new challenges continue to arise—poses a deeper question, what are the fundamental principles of the *adaptive universe*, i.e., what are the basic universal rules that govern how adaptive systems behave? Or at the least, what are the rules that govern how systems that serve human purposes adapt. Attempts to uncover these fundamental principles have begun. Two that are particularly important for the theory of graceful extensibility are Ostrom's work on polycentric governance (e.g, Ostrom 2012) and Doyle and his colleague's work on control of layered networks (e.g., Doyle and Csete 2011), despite the fact that these lines of inquiry develop from very different disciplinary starting points.

1.2 Graceful extensibility

This paper presents a theory that explains the contrast between successful and unsuccessful cases of sustained adaptability. The theory is stated in the form of 10 proto-theorems, which introduce a new core concept—*Graceful Extensibility*. Graceful extensibility is the opposite of brittleness, where brittleness is a sudden collapse or failure when events push the system up to and beyond its boundaries for handling changing disturbances and variations. As the opposite of brittleness, graceful extensibility is the ability of a system to extend its capacity to adapt when surprise events challenge its boundaries (Woods and Branlat 2011; Woods 2015). All systems have an envelope of performance, or a *range of adaptive behavior*, due to finite resources and the inherent changing variability of its environment. Thus, there is a transition zone where systems shift regimes of performance when events push the system to the edge of its envelope (e.g., how materials under stress can experience brittle failure; see Bush et al. 1999, and Woods and Wreathall 2008, for analyses of brittleness for complex systems drawing on the analogy to material science).

Boundary refers to the *transition zone* where systems shift regimes of performance. This boundary area can be more crisp or blurred, more stable or dynamic, partially well-modeled or mis-understood. Brittleness and graceful extensibility refer to the behavior of the system as it transitions across this boundary area. The latter refers to system's ability to adapt how it works to extend performance past the boundary area into a new regime of performance invoking new resources, responses, relationships, and priorities (for example, see Wears et al. 2008 for description of how medical emergency rooms adapt to changing and high patient loads and to Chuang et al. 2018 for how emergency departments adapted during a mass casualty event).

With low graceful extensibility, systems exhaust their ability to respond as challenges grow and cascade. As the ability to continue to respond declines in the face of growing demands, systems with low graceful extensibility risk a sudden collapse in performance. With high graceful extensibility, systems have capabilities to anticipate bottlenecks ahead, to learn about the changing shape of disturbances/challenges prior to acute events, and possess the readiness-to-respond to meet new challenges (Woods and Wreathall 2008; Woods et al. 2013). As a result, systems with high graceful extensibility are able to continue to meet critical goals and even recognize and seize new opportunities to meet pressing goals.

Studies of graceful extensibility ask: what do systems draw on to *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 can 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. And all systems, however successful, have boundaries and experience events that fall outside these boundaries—model surprise (Woods 2015).

Surprise has regular characteristics as many classes of challenge re-cur even though the specific challenges are relatively unique (Caporale and Doyle 2013). Cascades of disturbances and friction in putting plans into time are two examples of generic classes of demands that require the ability to extend performance to avoid collapse due to brittleness (Woods and Branlat 2011; Chen et al. 2015). Simply focusing on continual improvements turns out to move boundary areas around due to constraints imposed by fundamental trade-offs. As a result, improvements change where and how the system is exposed to collapse due to brittleness (Csete and Doyle 2002; Woods 2006; Alderson and Doyle 2010; Hoffman and Woods 2011). Ironically, local successes in space and time shift the kinds of events that produce challenges and shift how events challenge fitness as the boundary zone moves as well. This process of change means that graceful extensibility is a dynamic capability.

Brittleness describes how a system performs near and beyond its boundary zone, separate from how well it performs when operating well within its boundaries. Descriptively, brittleness is how rapidly a system's performance declines when it nears and reaches its boundaries. 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 interdependencies change, and as other parts of the network adapt. Plus, 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 (Woods 2006). In other word, the theory proposes that graceful extensibility is a necessary ingredient for sustained adaptability and then specifies constraints on how systems can demonstrate (or lose) graceful extensibility.

The term graceful extensibility can be thought of as a blend of two traditional terms—graceful degradation and software extensibility. I coined graceful extensibility because adaptation at the boundaries is active and critical to how a system grows and adjusts in a changing world, not simply a softer degradation curve when events challenge base competencies. Graceful extensibility also is a play on the concept of software extensibility from software engineering. Software engineering emphasizes the need to design, in advance, properties that support the ability to extend capabilities later, without requiring major revisions to the basic architecture, as conditions, contexts, uses, risks, goals, and relationships change. Ironically, the best examples of extensibility come from biology (e.g., Kirschner and Gerhart 2005). Graceful extensibility is a core concept that integrates a variety of the ideas in the theory such as boundary zones, surprise, brittleness, saturation, varieties of adaptive capacity, and forms of adaptive system breakdown.

1.3 What must a candidate theory provide? Six desiderata

Previous work on the mystery of sustained adaptability has identified a variety of key ideas that appear fundamental. First, the pioneering studies all started with a small set of fundamental trade-offs and then examined the implications of that set of trade-offs as basic constraints on layered networks facing complex environments (Doyle 2005; Woods 2006; Hollnagel 2009; Alderson and Doyle 2010; Hoffman and Woods 2011). A candidate theory needs to identify which are most fundamental, show how others that have been proposed can be derived, and provide a basis for how fundamental trade-offs arise.

Second, a candidate theory needs to explain how the performance of networks of adaptive units breaks down and what capabilities are needed to minimize or mitigate the risk of such breakdowns. Are there a few basic forms of adaptive systems failure and, if so, what are they?

Third, a candidate theory should provide the conceptual means to monitor, anticipate, learn, and respond to brittleness in order to steer away from the potential for a collapse *proactively* (Hollnagel et al. 2006). Of particular note, the current definition of brittleness is descriptive and not actionable—a collapse is required to identify the brittleness of a system in question. A candidate theory needs to supply a means to reduce the risk of brittle collapse.

Fourth, a candidate theory needs to provide a positive means for a unit *at any scale* to adjust how it adapts in the pursuit of improved fitness (how it is well matched to its environment), as changes and challenges continue apace. And this capability must be centered on the limits and perspective of that unit at that scale.

Fifth, a candidate theory needs to specify how units in a neighborhood of a network can interact to build sustained adaptability without centralized or command signals coming from outside that neighborhood. To do this, a candidate will need to integrate findings from diverse fields. The resulting integration must point to a means to influence or 'design' these networks, i.e., to show how complexity can be outmaneuvered (e.g., Doyle and Csete 2011; Woods and Branlat 2011).

Finally, a candidate theory must provide some ability to account for unintended consequences of changes proposed or ongoing in the network in question (e.g., the introduction of drones into civil aviation). This is especially important for systems that serve human purposes because stakeholders propose and pursue changes intended to improve performance from their perspective. These changes to complex networks of adaptive units result in widespread unanticipated reverberations, both negative ones that offset or undermine the impact on desired goals as well as surprising opportunities that are seized upon from unexpected quarters in ways that transform the network in question (news cycles continue to provide examples of both regularly—in 2017 one is ransomware).

The theory of graceful extensibility provides an account that covers these six desiderata, at least in part.

1.4 Fundamentals

The paper articulates the theory of graceful extensibility beginning with two simple assumptions and then progressing step by step in the form of 10 proto-theorems. The 10 statements are grouped into three subsets: managing the risk of saturation, networks of adaptive units, and outmaneuvering constraints. The paper provides some basic expositions of how these ideas capture general patterns about how sustained adaptability is built and lost.

I have called the 10 statements 'proto-theorems' deliberately, though potentially provocatively. I am claiming that the set of 10 statements to follow are provably correct, not only empirically correct (law-like). Observations fit these basic rules because the rules capture the fundamental characteristics of the adaptive universe. Observations of systems that serve human purposes and how these systems adapt to cope with and even exploit complexities have led me to the set and to the claim that they are fundamental. By fundamental I mean the set integrates and accounts for patterns from control engineering, from infrastructures, from distributed cognitive systems in context, from coordination and interaction across people, as well as across people and machines, and from organizational dynamics. In all of these areas adaptation occurs, patterns about adaptation are noted, and explanations are attempted. But the scope of reference for the patterns and their explanations is limited within each disciplinary topic. Behind these patterns and partial explanations lie a deeper truth about how the adaptive universe works that is then manifested when a portion of that universe is observed closely. Of course, this means the diverse uses of language across these disciplinary approaches undermine attempts at uncovering the fundamentals. The noisy interchange leads some to jettison language all together and to assert only a purely formal mathematical approach can provide clarity. The logical flow across the 10 statements is a different attempt to escape the languages of each contributing topic and to organize the key ideas in a basic progression—as logical a progression as I am capable of expressing given my own paths of inquiry. Table 1 summarizes the 10 statements organized into three subsets; Table 2 provides key terms of reference.

For the moment, I ask the reader to suspend the legitimate demand for the formal proofs and a resolution of a definitive

Table 1 Theory of graceful extensibility

Assumptions: (A) All adaptive units have finite resources. (B) Change is continuous
Subset A: Managing risk of saturation
S1: The adaptive capacity of any unit at any scale is finite, therefore, all units have bounds on their range of adaptive behavior, or capacity for maneuver.
S2: Events will occur outside the bounds and will challenge the adaptive capacity of any unit, therefore, surprise continues to occur and demands response, otherwise the unit is brittle and subject to collapse in performance.
S3: All units risk saturation of their adaptive capacity, therefore, units require some means to modify or extend their adaptive capacity to manage the risk of saturation when demands threaten to exhaust their base range of adaptive behavior.
Subset B: Networks of adaptive units
S4: No single unit, regardless of level or scope, can have sufficient range of adaptive behavior to manage the risk of saturation alone, therefore, alignment and coordination are needed across multiple interdependent units in a network.
S5: Neighboring units in the network can monitor and influence—constrict or extend—the capacity of other units to manage their risk of saturation, therefore, the effective range of any set of units depends on how neighbors influence others, as the risk of saturation increases.
S6: As other interdependent units pursue their goals, they modify the pressures experienced by a UAB which changes how that UAB defines and searches for good operating points in a multi-dimensional trade space.
Subset C: Outmaneuvering constraints
S7: Performance of any unit as it approaches saturation is different from the performance of that unit when it operates far from saturation, therefore there are two fundamental forms of adaptive capacity for units to be viable—base and extended, both necessary but inter-constrained.
S8: All adaptive units are local—constrained based on their position relative to the world and relative to other units in the network, therefore there is no best or omniscient location in the network.
S9: There are bounds on the perspective of any unit—the view from any point of observation at any point in time simultaneously reveals and obscures properties of the environment—but this limit is overcome by shifting and contrasting over multiple perspectives.
S10: There are limits on how well a unit's model of its own and others' adaptive capacity can match actual capability, therefore, mis-calibration is the norm and ongoing efforts are required to improve the match and reduce mis-calibration (adaptive units, at least those with human participation, are reflective, but mis-calibrated).

formal status for each statement (as theorems, lemmas, corollaries, or definitions). I expect that further inquiry, and I hope that heated response, will sharpen this account, or re-align it more closely to what proves to be the ultimate fundamentals. Perhaps, the value of this attempt lies only in its weaknesses and how these weaknesses stimulate others to launch different expeditions to uncover the fundamentals.

2 The theory

The theory of graceful extensibility is presented as 10 statements that express fundamentals that govern TLNs of UABs.

2.1 Assumptions

The first two statements emerge from a simple starting point about the nature of the adaptive universe: (a) resources are always finite and (b) change is ongoing (the rhythms of change). As a result, uncertainty is never zero and varies; risk is never zero and varies.

A basic term of reference in what follow is *Units of Adaptive Behavior* or UABs (see Table 2). These are units that adapt their activities, resources, tactics, and strategies in the face of variability and uncertainty to regulate processes relative to targets and constraints. For human systems, adaptation includes adjusting behavior and changing priorities in the pursuit of goals. UABs exist at multiple nested scales (e.g., processes, individuals, roles, agents, teams, networks, groups, organizations, enterprises, societies). UABs are active, seeking to improve how well they 'fit' their environment given the activities of other nearby UABs. One can consider a UAB as a generalization from control engineering, i.e., any control mechanism that regulates a process to meet targets in the face of disturbances, as well as a generalization from classic characterizations of human skill and expertise, i.e., the ability to adapt behavior in changing circumstances to pursue goals.

The above assumptions are all that is necessary to establish that any and all units have bounds on their range of adaptive behavior (S1) and that events—surprises or challenge events—will occur outside those bounds (S2).

2.2 Subset A: risk of saturation

S1 The adaptive capacity of any unit in a network at any scale (UAB) is finite, therefore, all units have bounds on their range of adaptive behavior—*or, Boundaries are universal*.

S1.1 The location of boundaries to the ability to meet demands is uncertain.

There is a boundary on any unit's adaptive capacity or the ability to be in-control or stay

Table 2 Terms of reference

Adaptive unit or unit of adaptive behavior (UAB): Units in a network that adapt their activities, resources, tactics, and strategies in the face of variability and uncertainty to regulate processes relative to targets and constraints. For human systems, adaptation includes adjusting behavior and changing priorities in the pursuit of goals. Units of adaptive behavior (UABs) exist at multiple nested scales (e.g., processes, individuals, roles, agents, teams, networks, groups, organizations, enterprises, societies). UABs are active, seeking to improve how well they 'fit' their environment given the roles of other nearby UABs.
Fitness: Fitness refers to the match between an organism's capabilities and the properties of its environment. But this degree of match is a question, and answers to this question are always tentative, never complete. Current capabilities and activities are a temporary, local answer to the question of what is fitness. In the adaptive universe, adaptive units at all levels are able to constantly re-think their answer to the question—fitness, both in terms of changing capabilities and in terms of changing challenges and opportunities in their world.
Adaptive capacity: The potential for modifying what worked in the past to meet challenges in the future; adaptive capacity is a relationship between changing demands and responsiveness to those demands, relative to goals.
Range of adaptive behavior: For any adaptive capacity, that capacity generates a range of behavior that is adapted to the patterns of change ongoing and upcoming; thus, adaptive capacity has a range or a boundary over which it is capable of responding to changing demands. While range of adaptive behavior is a term that has been used to describe biological systems, it is passive in tone, whereas adaptive units are quite active so a better term is *Capacity for Maneuver* (*CfM*) and this capacity has limits.
Saturation: Exhausting a unit's range of adaptive behavior or capacity for maneuver as that unit responds to changing and increasing demands.
Risk of saturation: Inverse of remaining range or capacity for maneuver given ongoing and upcoming demands.
Brittleness [descriptive]: Rapid fall off (or collapse) of performance when situations challenge boundaries.
Graceful extensibility: How to extend the range of adaptive behavior for surprises at and beyond boundaries—to deploy, mobilize, or generate capacity for maneuver when risk of saturation is increasing or high. Graceful extensibility is the inverse of brittleness.
Brittleness [proactive]: Insufficient graceful extensibility to manage the risk of saturation of adaptive capacities.
Varieties of adaptive capacity: Minimally, there are 2—base and extensible.
Base adaptive capacity: The potential to adapt responses to be fit relative to the set of well-modeled changes. The extended form of adaptive capacity is referred to as *graceful extensibility*: how to extend the range of adaptive behavior for surprises at and beyond boundaries—to deploy, mobilize, or generate capacity for maneuver when risk of saturation is increasing or high.
Net adaptive value equals the total effective range of base plus extended adaptive capacities. [Note: adaptive value is an expression used in understanding biological systems.]
Surprise: Given bounds on adaptive capacity, there are events which will occur that fall near and outside the boundaries; thus, surprise is model surprise where base adaptive capacity represents a partial model of fitness.
Potential for surprise: information about what surprises may occur in the future. The potential for surprise is related to the next anomaly or event that will be experienced, and how that next event may challenge pre-developed plans and algorithms in smaller or larger ways—how well plans/models/automata fit particular situations to be handled.
[Tangled] Layered Networks: Rather than speak of systems or groups or organizations, the theory addresses layered networks as defined by Doyle in Alderson and Doyle (2010) and Doyle and Csete (2011). I added the descriptor 'tangled' to emphasize that the network interdependencies are very hard to map, messy, change,contingent, and often hidden from view.

in-control as variation, disruptions, and change occur. As a result, all UABs can run out of the capacity to adapt as the demands faced grow and change in difficulty.

S1.2 Given a finite range, there is a general parameter—*Capacity for Maneuver (CfM)*—which specifies how much of the range the unit has used and what capacity remains to handle upcoming demands.

- All UABs risk *saturation*, that is, running out of *CfM* as upcoming events present increasing challenges or demands [also referred to as the risk of saturating control]. Managing the risk of saturation then becomes the definition of what it means to be in-control.

S2 Events will occur that challenge boundaries on the adaptive capacity of any unit, therefore, surprise continues to occur and demands response, otherwise the unit is brittle and subject to collapse in performance—*or, Surprise occurs, continuously*.

S2.1 There are recurring patterns that characterize model surprise—how events challenge boundaries.

- Events will occur at some rate and of some size and of some kind that increase the risk of saturation—exhausting the remaining *CfM*.
- Descriptively and specifically, brittleness is how rapidly a unit's performance declines when it nears and reaches its boundaries (S1). Brittleness describes how a UAB performs near, at and beyond its boundaries, separate from how well it performs when operating far from its boundaries.

- The range of adaptive behavior of a UAB is a model of *fitness*; that model has boundaries (S1) and events occur which fall outside that boundary → model surprise.
- Events that occur near or outside a UAB's boundary increasees the risk saturation, and this occurs independent of how well that UAB matches responses to demands (the degree of fit) well within its range of adaptive behavior (or *competence envelope*).

S3 All units risk saturation of their adaptive capacity, therefore, units require some means to modify or extend their adaptive capacity to manage the risk of saturation when demands threaten to exhaust their base range of adaptive behavior—*or, Risk of saturation is monitored and regulated*.

S3.1 The work (effort/energy/resources) required to adapt and handle changing demands increases as *CfM* decreases, i.e., there is some function relating effort to be in-control to the risk of saturating *CfM*.

S3.2 As risk of saturation increases and *CfM* approaches exhaustion, UABs need to adapt to stretch or extend their base range of adaptive behavior to accommodate surprises. This extended form of adaptive capacity is graceful extensibility: how to deploy, mobilize, or generate capacity for maneuver when risk of saturation is increasing or high.

S3.3 The risk of saturating controls as demands grow and cascade creates systematic patterns in how adaptive systems break down. The first systematic pattern is **decompensation**, which is, exhausting the capacity to adapt as disturbances/challenges grow and cascade faster than responses can be decided on and deployed to effect.

- All UABs have some potential for adaptive response 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. Concepts about varieties of adaptive capacity can be integrated around the single parameter of Capacity for Maneuver (*CfM*) and how UABs adjust/regulate their adaptive capacities relative to the risk of saturating *CfM* as they respond to future challenges and opportunities. The struggle for fitness in the face of changing demands is ongoing and requires the potential to adjust adaptive capacities. This leads to a new operational and actionable definition of brittleness as the risk of saturating *CfM* and to the concept of graceful extensibility as the opposite of brittleness. Risk here becomes operationalized as some dynamic function of how *CfM* is being used and what remains relative to ongoing and possible future demands.

2.2.1 Exposition for subset A

The goal for the theory is to capture basic generalities about adaptive capacity that apply in all contexts and as change continues—what has to be true about sustained adaptability and this starts with bounds and surprise (S1 and S2).

The theory (or any set of theorems about networks of adaptive units) is not about how well a UAB meets its targets and constraints, and not about how they regulate processes. Rather, however a UAB regulates processes and however well it meets targets and constraints, (a) its capacity to do these things is bounded and (b) the environment will present events that fall outside its bounds. The UAB has to have some ability to continue to function when this happens, if not the system is too brittle and vulnerable to sudden collapse so that long-term viability declines. The viability of a unit in the long run requires the ability to gracefully move its capabilities as change continues to produce new challenges, surprises, and opportunities. *Viability requires extensibility*.

The hospital emergency department or ER is well studied as an exemplar of graceful extensibility that can serve to ground the theory (Miller and Xiao 2007; Wears et al. 2008; Perry and Wears 2012; Patterson and Wears 2015). All systems have some built-in capacity matched to handle regularities and variations and variation of variations. This defines their base adaptive capacity or competence envelope—what they can handle without risk of saturation—performance far from saturation. (To anticipate, note S10 though—reflective systems, always risk overestimating their base adaptive capacity/competence envelope.) Some systems are designed to be able to handle a range of changing demands. For example, ERs are able to adjust their resources to handle a range of patient problems and numbers of patients as these vary within and across every shift as well as over longer time frames. Each ER still has finite ability to handle surges in patient load. Situations occur that challenge the boundaries even of systems like the ER that are equipped to handle changing loads. And then the issue is what happens when situations challenge boundaries—when risk of saturation is high? ERs regularly have experience with situations that challenge their ability to respond.

ERs demonstrate a base competence which can be seen in the specific deployable capabilities present at any point in time—areas of expertise, staff with various levels of experience, space, equipment, supplies, plans for special situations (mass casualty), and others. And the staff in ERs regularly experience events that challenge this envelope so that they also exhibit some degree of graceful extensibility as they recognize increasing risk of saturation and adapt in a variety of ways (Wears and Woods 2007). As Wears et al. (2008) describes, personnel reconfigure their work, their patients, how they coordinate, how they utilize available equipment, and how they create additional effective space to manage increasing load before their ability to provide care is saturated. By studying what these people draw on to demonstrate resilient performance, the basis for graceful extensibility can be understood—capabilities such as anticipation, initiative, and reciprocity.

ERs illustrate how UABs have some degree of graceful extensibility in addition to a base capacity for handling demands. And the studies of how ERs adapt as load goes up also illustrate that there are limits on how much graceful extensibility a single UAB can exhibit near saturation. ERs can and do have breakdowns where triage goes astray, patients are mis-prioritized/under-monitored, and patient condition deteriorates faster than ER staff can recognize/respond (the decompensation form of adaptive system breakdown). Continued graceful extensibility requires other neighboring UABs in the network to recognize and adjust when the ER risks saturation (e.g., Stephens et al. 2015). For example, this should occur when other units in the hospital system adapt to assist the ER as should occur in mass casualty events (Chuang et al. 2018).

A key phrase used to characterize graceful extensibility in a setting like the ER or in general is—the ability to be prepared to be surprised. This appears on first blush to be a contradiction—if it is a surprise, one cannot be prepared. However, this is mistaken as in the ER case. First, many classes of demands or challenge re-cur. As per Statements 1 and 2, model surprise is a regular occurrence and thus can be tracked and characterized so that changes can be recognized. There are general forms of challenges that apply to all layered networks, however tangled. Basic examples are cascades of disturbances and friction in putting plans into time.

> However, many classes of environmental challenge re-cur. 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 (Caporale and Doyle 2013, p. 20).

Second, adaptive capacity is a potential for future action when conditions change.

Definition: Adaptive capacity is the potential for adjusting patterns of activities to handle future changes in the kinds of events, opportunities, and disruptions experienced, therefore, adaptive capacities exist before changes and disruptions call upon those capacities.

Responding to surprise requires preparatory investments that provide the potential for future adaptive action. Thus, biology finds "mechanisms that generate variation can adapt to a recurring nonuniform distribution of challenges" (Caporale and Doyle 2013, p. 21). Examples of architectures in biology that facilitate future adaptability continue to be uncovered (Meyers and Bull 2002; Beaumont et al. 2009).

The first portion of the theory also introduces the term *CfM* as a general parameter that characterizes all UABs whatever the scale (note: originally, the label used for this parameter was margin of maneuver Woods and Branlat 2010, 2011). That bounds and surprise are universal simply means a UAB risks saturating or running out of *CfM* as new events occur or could occur. Saturation refers to how much *CfM* has been used up to handle ongoing events, which then reduces the remaining *CfM* available to handle upcoming and future events. Nearing saturation, or increasing risk of becoming saturated, means little *CfM* remains available to handle upcoming and future events.

This subset of the theory captures how bounded capabilities to manage surprise can be integrated into a single parameter—*CfM*—and in particular, the risk of exhausting or of saturating *CfM which can be monitored and regulated. CfM* is a simple unifying concept that is theoretical yet is also applicable across diverse practical settings.

This subset captures how *CfM* is regulated to manage and reduce the risk of saturation. This subset asserts *CfM* is a control parameter, that is, units can monitor and act on information about the risk of saturation. Note this does not mean they always do this well. The stronger meaning is that, to produce sustained adaptability, a unit must be a good regulator of the risk of saturation (following Conant and Ashby 1970).

The risk of saturating *CfM* (or risk of saturation), operationalized as exhausting the capacity for maneuver as demands grow or new demands arise, is the central idea in the theory. *CfM* is a potential and if that potential gets too low, saturation risk is high. No matter what is to be controlled or managed, and no matter how well that is controlled or managed, things can and will change. When that change or new challenge occurs, some capacity has to be there to draw on to adjust to the change or challenge—otherwise the system is too brittle and the risk of collapse in performance relative to an important criterion or dimension is too high.

It is particularly important to note that the capacity for maneuver is a parameter that is defined by the relationship between events/variations in the environment that demand response and the capability of the UAB in that environment to respond to those demanding events by drawing on various resources. Demands refer both to the demands that have absorbed capacity to respond already and the upcoming demands that will absorb or exceed the remaining capacity to respond. The capability to continue to make effective responses in the face of changing streams of demands is the ability to stay "in-control."

The risk of saturation provides an operational and actionable definition of brittleness and provides the connection to assess the potential for failure. As a general parameter, it has scale properties, can be estimated (particularly how it is changing), and that estimate provides important information about how well a UAB can provide sustained adaptability (e.g., Fariadian et al. 2018).

Regulating *CfM* defines a needed adaptive capacity. Thus, research results can be organized around how *CfM* is built and lost, how it could be sustained in the face of competing pressures, what resources are needed, how that capacity should be adjusted to match information about changes in ongoing or upcoming challenges. Stakeholders can begin to ask what is the right size of that capacity and its associated resources. If the capacity is too small, the risk of saturating *CfM* is too high. If the capacity is too large, the resources required to produce that level of *CfM* will erode under the inevitable pressures for better performance on criteria that apply far from saturation. The latter pattern has been observed regularly where the capability for graceful extensibility degrades over time due to production pressures until a brittle failure occurs (e.g., Woods 2005, 2006).

The classic exemplar of regulating *CfM* as a general capacity is how some human operators and teams are able to "anticipate bottlenecks ahead" and adapt activities to generate the capacity to handle that bottleneck or challenge should it arise (e.g., from studies of expertise at anesthetic management during surgeries). To fail to anticipate and prepare for the bottleneck ends up putting the team in a situation where they have to generate the means to respond in the middle of the challenge event—greatly increasing the risk of decompensation—failing to keep up with the pace and tempo of events. Thus, decompensation, the risk of being slow and stale to respond to the pace of events, is the first of three basic failure modes for adaptive systems. The ability to decide on and deploy actions to effect as the pace of disturbances grow and cascade is critical to the capacities that make up *CfM* and the risk of saturating *CfM*. This pattern is easily seen in risky operational settings, but it exists at multiple scales for all forms of adaptive systems (Woods and Branlat 2011).

Sustained adaptability naturally leads one to ask: what does it mean to be always adapting? Always adapting, being *poised to adapt*, is a readiness or potential to change (Wears and Woods 2007; Woods and Branlat 2010, 2011). Always adapting does not mean you are changing what you do, shifting how you do things, or adjusting what you have planned *all the time*, but rather that you're able to recognize when it is adequate to continue the plan, to continue to work in the usual way, and when its not adequate to continue on, given the demands, changes, and context ongoing or upcoming. For example, studies of accidents have noted (Woods and Shattuck 2000), p. 242:

> In these studies, either local actors failed to adapt plans and procedures to local conditions, often because they failed to understand that the plans might not fit actual circumstances, or they adapted plans and procedures without considering the larger goals and constraints in the situation. In the latter type B problems the failures to adapt often involved missing side effects of the changes in the replanning process.

Adaptation can mean *continuing to work to plan*, but, and this is a very important *but*, with the continuing ability to reassess whether the plan fits the situation confronted—even as evidence about the nature of the situation changes and evidence about the effects of interventions changes. Adaptation is not about always changing the plan or previous approaches, but about the potential to modify plans to fit situations—being poised to adapt. Space mission control is the positive case study for this capability. See Woods and Hollnagel (2006, Chap. 8) and Patterson et al. (1999), Watts-Perotti and Woods (2009) for studies of how space shuttle mission control developed its skill at handling anomalies, even as they expected that the next anomaly to be handled would not match any of the ones they had planned and practiced for previously. Successful military organizations are another positive case study; see the contrasting cases in Finkel (2011).

The ability to stretch, extend, or change what you are doing/what you have planned, and the ability to recognize when this is needed, has to be there in advance of adapting, *even when there are no adjustments to behavior visible to outside observers. CfM* is meant to capture what has to be there to have the potential for future adaptive action. In particular, the risk of losing *CfM* is the signal to monitor that tells an agent or unit that it needs to adjust and adapt.

As a potential for future adaptive action, *CfM*, like any concept in biology or physics that is defined as a potential, has some interesting properties and challenges for developing sustainable regulatory mechanisms. For example, putting potential into action for some demands then consumes the potential to respond to other demands. The difficulties in handling this constraint over time becomes visible when observers examine how critical resources such as operating rooms and intensive care beds are managed (e.g., Cook 2006).

2.3 Subset B: networks of adaptive units

Next, given Statements 1–3 (boundaries, surprise, risk of saturation, and potential), graceful extensibility depends on how one adaptive unit (UAB) interacts with neighboring units in a network of interdependent units.

S4 No single unit, regardless of level or scope, can have sufficient range of adaptive behavior to manage the risk of saturation alone, therefore, alignment, coordination, and synchronization are needed across multiple interdependent units in a network—*or synchronization across multiple UABs in a network is necessary*.

S4.1 UABs exist in and are defined relative to a network of interacting and interdependent UABs at multiple scales → networks with multiple roles, multiple echelons.

- As risk of saturating the base adaptive capacity grows, additional adaptive capacity must be brought to bear, and this requires invoking other UAB that extend *CfM* beyond the remaining capacity of the unit at risk of saturation. To bring additional adaptive capacity to bear, requires alignment, coordination, and synchronization across multiple units and echelons.

S5 Neighboring units in a network can monitor and influence—constrict or extend—the capacity of other units to manage their risk of saturation, therefore, the effective range of any set of units depends on how neighbors influence others as the risk of saturation increases somewhere in that neighborhood of the network—*or, risk of saturation can be shared*.

S5.1 Misalignment and mis-coordination across UABs increases the risk of saturating control as demands grow and cascade. This creates a second form of adaptive system breakdown—**working at cross purposes** where one UAB responds to demands by managing its *CfM* in ways that reduce the *CfM* of UABs nearby or at a larger or finer scales. When this occurs it reveals a general pattern of responses that are locally adaptive (from one perspective), but globally maladaptive (from a different perspective). On the other hand, some UABs monitor the risk of saturating *CfM* in another UAB by monitoring signals associated with the increasing effort to stay in-control. When they recognize that the risk of saturating the *CfM* of the other unit is becoming too high, they respond in ways that have the effect of extending the capacity and behavior of the UAB at risk.

S6 As other interdependent units pursue their goals, they modify the pressures experienced by a UAB of interest. In response to changing experienced pressures, a UAB searches for better operating points in a multi-dimensional trade space—*or, pressure changes what is sacrificed when*.

In pursuing their goals, a Unit of Adaptive Behavior (UAB) generates pressure on neighboring UABs. As a result, the goals UABs pursue or prioritize are changed relative to the pressures they experience and the conflicts these pressures exacerbate or generate. As the pressures generated by other interdependent units change, the trade-offs a unit faces change. The pressures experienced influence the search for how to balance or prioritize across basic trade-offs, especially when trade-offs intensify (Woods 2006). This constraint poses the research question—what architectural properties of the network influence the way units in a network respond to varying pressures on trade-offs?

2.3.1 Exposition for subset B

Each UAB must have some capacity to adapt when risk of saturation is high (when trends on demands risk exhausting *CfM* relative to deployable response capability), yet no single UAB has sufficient capability to manage the risk of saturation completely, given bounds and surprise. As a result, a UAB exists in a network where the activities of nearby units affect the *CfM* of the target UAB, either constricting or expanding its *CfM* (either intentionally or unintentionally from the perspective of nearby units).

The *CfM* at any one unit is affected by the activities of interdependent (nearby) units across a network. When other unit's activities, relative to their own goals and relative to managing their own *CfM*, constrict the unit of interest, the units are working at cross purposes. This is the second general form of breakdown in adaptive systems (i.e., *locally adaptive but globally maladaptive* responses) which relates to how the responses of nearby UABs constrict, rather than extend, the *CfM* of other UABs, defined at the same scale or at larger or narrower scales. Results from a study of hospitals where patients get stuck in the emergency department (ER) for long periods of time illustrate this process across UABs in a network (Stephens et al. 2015). This study observed working at cross purposes in the interaction between the ICU and the ER as each group worked hard to achieve the local goals defined for their scope of responsibility, but each unit pursued their goals in such a way that their activities made it more difficult for other groups to meet the responsibilities of their roles. Plus the locally adaptive behavior for each unit undermined the global or long-term goals that all groups recognized as important (for this case, it is poor care to leave patients stuck in the ER for long time periods).

Mis-synchronization across roles and echelons can lead a set of UABs in a network to respond too slowly to a cascading problem, that is, decompensate. A case of "runaway" automation in financial trading, the Knight Capital collapse in August, 2012, illustrates how events can challenge coordination across roles so that decisions end up slow and stale (see https://michaelhamilton.quora.com/How-a-software-bug-made-Knight-Capital-lose-500M-in-a-day-almost-go-bankrupt and https://www.kitchensoap.com/2013/10/29/counterfactuals-knight-capital/). In this case, one part of the organization deployed new software in order to keep up with and take advantage of changes in the industry—and all changes regardless of type become changes in software for computerized financial trading. The rollout did not go as expected and produced anomalous behavior. The team then tried to roll back to a previous software configuration as is standard practice for reliability. But the rollback produced more anomalous behavior. The roles responsible for managing the digital infrastructure struggled to understand what produced the anomalous behaviors and the failure of normal attempts to recover to block or stop the cascade of effects. Meanwhile, automated trading continued.

The team felt it did not have the authority on its own to stop trading. By the time the team was able to decide to go to upper management and tell upper management that there was a problem, that they did not understand the problem, that they were unable to block the cascade of effects, and that the only action available was to stop trading, tens of minutes had gone by. When upper management approved and trading was stopped, it was already too late—so much automatic trading had gone on that the company was left holding an untenable position in the markets and was, for all practical purposes, bankrupt.

First, this case illustrates how small problems can interact and cascade quickly and surprisingly given the tangle of dependencies across layers inside and outside the organization. Second, as effects cascade and uncertainties grow, multiple roles struggle to understand anomalies, diagnose underlying drivers, identify compensating actions. Third difficulties arise getting authorization from appropriate roles to make non-routine, risky, or resource costly actions, especially while uncertainty remains. Fourth, all of the above take effort, time, and require coordination across roles/levels. Fifth, when critical decisions require serial communication vertically through the network to receive authorization, actions are almost always unable to keep pace with events, change, and challenge. Thus, the case illustrates a combination of the two basic failure modes for adaptive systems: inability to keep pace with events and working at cross purposes, in this case across the vertical layers.

Subset B addresses what is required for a layered network to sustain adaptability. Tangled layered networks require capabilities that will synchronize the responses across UABs when one or more of the units in that network start to run out of *CfM* as surprising challenges unfold. Thus, some UABs have to be capable of monitoring the risk of saturation of other UABs or of a region of the network. These UABs also need to be capable of synchronizing activities across units in new ways to extend the necessary readiness to respond when demands change, increase, and threaten saturation. In the above example, Knight Capital did not have the capability to shift to forms of coordination that could have kept pace with the cascading effects. Other more successful cases indicate the key properties that provide graceful extensibility. Among these key properties are the *expression of initiative* (Finkel 2011) and *reciprocity* (Ostrom 2003).

2.3.2 What governs the expression of initiative?

Note that initiative is a fundamental property of UABs. A unit with no or reduced initiative loses its ability to function as a UAB, as its contribution to producing graceful extensibility drops or disappears. Consistent with the empirical findings (e.g., Woods and Shattuck 2000; Finkel 2011) and with subset A of the theory, UABs have to possess some degree and form of initiative to contribute to graceful extensibility.

Initiative is a necessary capability for adaptation that has three parts: (a) the ability of a unit to adapt when the plan no longer fits the situation, as seen from that unit's perspective; (b) the willingness (even the audacity) of a unit to adapt planned activities to work around impasses or to seize opportunities in order to better meet the goals/intent behind plans or pursue new goals; (c) when taking the initiative, a unit begins to adapt on its own, using information and knowledge available at that point, without asking for and then waiting for explicit authorization or tasking from other units.

On the other hand, initiative can run too wide when undirected, leading to fragmentation, working at cross purposes, and mis-synchronization across roles in a different way. The question is—what governs the expression of initiative? The changing pressures generated by other units energizes or reduces initiative. The changing pressures also constrain and direct how the expression of initiative prioritizes some goals and sacrifices others goals when conflicts in the trade space intensify. Changes in the pressures experienced by a unit changes how that unit moves in the multi-dimensional trade space. In other words, changing pressure influences what goals are *sacrificed* as pressures, demands, and risk of saturation grow (Woods 2006).

The theory poses a key research question: What properties of a network of adaptive units organizes relationships across units so that the expression of initiative is regulated to produce and sustain graceful extensibility?

2.3.3 Reciprocity

Coordination across units in this theory is based on how nearby adaptive units respond when another UAB experiences increasing risk of saturating its *CfM*. Will the neighboring units adapt in ways that extend the *CfM* of the adaptive unit at risk? Or will the neighboring units behave in ways that further constrict the *CfM* of the adaptive unit at risk? Ostrom (2003) has shown that reciprocity is an essential property of networks of adaptive units that produce sustained adaptability.

It is important to note that all UABs can be in either position depending on events and context—the unit at risk of saturation in need of assistance from neighbors, or the neighbor with the potential to assist another at risk of saturation. In networks with high graceful extensibility, adaptive units demonstrate *reciprocity*—each unit can anticipate assistance from neighbors should saturation of their *CfM* loom larger, even though that assistance requires adaptations by the assisting units—adaptations that increases the costs and risks experienced by those assisting units. Reciprocity in the theory of graceful extensibility means that when an adaptive unit provides assistance that unit "anticipates" others will adapt to assist them in the future when they are at risk of saturation. In showing reciprocity, adaptive units take on costs and risks relative to their goals, for example, expending resources that could have gone to ensuring its own performance criteria are well met. This is a kind of sacrifice of unit specific performance criteria in order to ensure gains at neighborhood levels of performance through a form of synchronization across units. Reciprocity is a mechanism for alignment across units when crunches occur. Mechanisms like this for goal alignment and sacrifice in a network as conflicts arise or intensify are evident and needed in biological and in human systems (e.g., Meyers and Bull 2002 and; Ostrom 2012).

It is easiest to explain reciprocity in terms of the interaction between two roles. UABs 1 and 2 demonstrate reciprocity when UAB1 takes an action to help UAB2 that gives up some amount of immediate benefit relative to its scope of responsibility. The sacrifice by UAB1 relative to a narrow view of its role allows for a larger, longer run benefit for both UABs 1 and 2 relative to the broader goals of the network in which these two units exist. But in helping another unit manage its risk of saturation, UAB1 is relying on UAB2 to "reciprocate" in the future—when UAB1 needs help, UAB2 will be responsive and willing to take actions that will give up some benefit to that role in the short run in order to make both roles better off relative to common goals and constraints.

One UAB is donating from their limited resources now to help another in their role in order to achieve benefits for overarching goals. There are limited resources in terms of energy, workload, time, attention for carrying out each role. Diverting some these resources creates opportunity costs and workload management costs for the donating unit. On the other hand, a UAB can ignore other interdependent roles and focus their resources on meeting the standards set for performance in their role alone (especially if that unit is under 'faster, better, cheaper' pressure; Woods 2006), even though this can be quite short sighted and parochial.

Notice the potential instability arises because there is a lag between donating limited resources and when that investment will pay off for the donating unit. It could pay off in better performance on larger system goals—an investment toward common pool goals—but that effect may be quite difficult to reflect back on and improve matters for the donating unit. It could also "pay off" in the future when other units make donations of their limited resources to help the unit donating now when it experiences challenges. The investment is now, definite, and specific; the benefit is uncertain and down the road. Plus the receiving unit can act selfishly, exploiting the donating unit, by not being willing to reciprocate in the future. Aligning the multiple goals will always require relaxing (a sacrifice) some local short-term (acute) goals in order to permit more global and longer-term (chronic) goals to be addressed (Woods 2006). For graceful extensibility, interdependent UABs show a willingness to invest energy to accommodate other units specifically when the other units are at risk of saturation, rather than just performing alone, walled off inside its narrow scope and sub-goals.

The creation and decay of reciprocal relations have been explored extensively in social science and experimental micro-economics as in the Nobel prize winning works of Ostrom (2012) and Roth (2008). One important point here is that designing for resilient control in an engineering sense has to incorporate concepts found as fundamental in studies of human social systems.

In Subset B, the theory of graceful extensibility captures several basic processes that influence how adaptive units will act when a neighbor is at risk of saturation and whether units will act in ways that extend or constrict the *CfM* of the unit at risk. The interplay across adaptive units will exert pressures on all UABs. These pressures modify the expression of initiative and reciprocity in ways that extend or constrict *CfM* of the network of adaptive units. The theory challenges modelers and empiricists to delineate how varying pressures across units influence initiative, reciprocity, and other factors in ways that produce or undermine graceful extensibility.

2.4 Subset C: constraints on maneuver

Given the previous Statements,

S7 Performance of any unit as it approaches saturation is different from the performance of that unit when it operates far from saturation, therefore there are two fundamental forms of adaptive capacity for units to be viable—base and extended, both necessary but inter-constrained—*or, pressure for optimality undermines graceful extensibility*.

Managing the risk of saturation of *CfM*, at a minimum, requires two forms of interdependent adaptive capacity—'base' and 'extended.' Base refers to the mechanisms, resources, and performance characteristics of the UAB when it is operating far from saturation. Extended refers to the capability to monitor risk of saturation and adapt to mobilize or generate additional *CfM* when risk of saturation is high. To extend, adaptive capacity requires mechanisms that consume resources; investing in the resources that provide the extended adaptive capacity negatively impacts on base adaptive capacity. And the reverse holds—improving base adaptive capacity in isolation reduces the resources that underpin the capacity to extend response capability when risk of saturation is high. *Net adaptive value*, as a sense of fitness, includes both.

Adaptive value is a term often used in models of how biological and neurobiological systems increase their fitness to a changing environment (e.g., Bialek et al. 2007). The 'value' refers to the advantage in fitness gained for the unit in question when it adapts. Adaptive value is seen quite clearly in human sensory systems where adaptation is ubiquitous (see e.g., Attneave 1954; Brenner et al. 2000; Wark et al. 2007). Models of neurocomputation in sensory adaptation have to account for the relationship between behaviors that provide adaptive value and the availability and scarcity of the resources those processes require (Fairhall et al. 2001; Wark et al. 2009). Sensory systems use a variety of cues to move a finite range of discrimination power to maximize the fit between sensing capability and what is valuable to sense, and these systems do this while keeping pace with a changing stimulus world. Note how the processes of sensory adaptation illustrate risk of saturation (S1 to S3), the ability of some units to assess others' risk of saturation (S4 to S6), mechanisms that are poised to adapt when risk of saturation is high so as to maintain the capability to perform. In this case, the result is an expanded dynamic range multiple orders of magnitude greater than the base capability (e.g., brightness discrimination in human sensory systems).

The theory builds on this tradition and recognizes explicitly that there are two basic kinds of adaptive value—one far from saturation and another that operates near saturation. Operating far from saturation, when criteria are oriented toward optimality (that is, pressures for adding value to base adaptive capacity), gains come from achieving a reference level of performance from a reduction in resources (more efficiency or productivity). For graceful extensibility needed near saturation, adding adaptive value comes from expanding the performance possible from a reference level of resources. This leverages the adaptive value from a set of available resources to produce and sustain graceful extensibility. These different forms of value help reveal the trade-off between these two different, but both critical, forms of adaptive capacity. Reducing resources to improve performance far from saturation inadvertently targets resources that underlie graceful extensibility when risk of saturation is growing. This pattern is what has been seen in systems safety (e.g., the lead up to the Columbia space shuttle disaster; Woods 2005). On the other hand, neurobiology is one realm where systems perform well on both—despite the trade-off. Neurobiology provides existence proofs that systems can effectively pursue **net** adaptive value. Addressing the mystery of sustained adaptability should lead to the identification of the basic architectural properties that allow networks of adaptive units to perform well on net adaptive value despite the trade-off (e.g., Doyle and Csete 2011).

S8 All adaptive units are *local*—constrained based on their position relative to the world and relative to other units in the network, therefore there is no best or omniscient location in the network.

A UAB is embedded in a place relative to an environment and a set of relationships across a network of UABs. A UAB is responsible for goals relative to its local position in the network—responsible in the sense that that the UAB experiences that consequences that result from achieving or failing to achieve its goals. Different UABs in the network are differentially responsible for different subsets of goals that can interact and conflict.

S9 There are bounds on the perspective of any unit—the view from any point of observation at any point in time simultaneously reveals and obscures properties of the environment—but this limit is overcome by shifting and contrasting over multiple perspectives—*or, perspective contrast overcomes bounds*.

Each UAB in a network has a perspective where perspective consists of a point of observation (think of this as the position of a virtual camera) relative to a point of interest in a scene which defines a view direction and a field of view. The view from any point of observation simultaneously reveals and obscures properties of the environment. There is no best perspective. To see perspective requires another perspective (or a perspective shift). The capacity to shift and contrast perspectives is essential for adaptive action.

S10 There are limits on how well a unit's model of its own and others' adaptive capacity can match actual capability, therefore, mis-calibration is the norm and ongoing efforts are required to improve the match and reduce mis-calibration—*or, reflective systems continually risk mis-calibration*.

Note calibration is how well an agent can model and track its own capabilities and performance. Mis-calibrated agents usually overestimate their capability to perform a task and tend to underestimate the difficulties and demands to be faced when carrying out that activity.

S10.1 A UAB's model of itself and others will be mis-calibrated without mechanisms to shift and contrast perspectives. Mis-calibration risks include all of the parameters of networks of UABs defined previously (e.g., boundaries, risk of saturation, demands, perspective).

S10.2 Since risk of mis-calibration is omnipresent, effort must be invested to reduce risk of mis-calibration. In other words, since there is a bound on how well models of capability match actual capability, effort must be invested to improve the match.

S10.3 To fail to continue to check and adjust calibration means that learning will slow or stop. This learning breakdown defines the third basic form of maladaptive behavior: where models of adaptive capacity become stuck and outdated as a result of change. Given changes afoot, models of demands and models of effective responses to those demands, which had been adaptive in the past, become stale, are no longer effective and require revision.

S10.4 Boundary areas are *discovered* and known only through the *experience* of surprise and the experience of risk of saturation. Furthermore, changing to handle the risk of saturation produces change to the system adapting. These changes modify what is base adaptive capacity, and modifies what and when and where surprise occurs.

UABs have models of the their own adaptive capacity (i.e., are reflective) and models of the adaptive capacity other nearby and nested UABs (across both horizontal and vertical interdependencies in the network). A UAB has limits on its ability to model its own and other's ability to regulate *CfM* including the risk of saturating *CfM*. It tends to underestimate demands and how they change and to overestimate base adaptive capacity. When mis-calibrated, UABs are under-responsive to changes in demands and slow to learn and adopt new responses to handle the changes. As the location of boundaries are uncertain and dynamic, mis-calibration further limits a UAB's ability to explore boundary areas and update models. Thus, mis-calibrated UABs tend to act in ways that constrict the *CfM* of other units in the network.

2.4.1 Exposition for subset C: shifting the view of brittleness

Traditionally, brittleness refers to how performance changes when demands push the network in question near to boundaries, i.e., when a system's performance declines rapidly or dramatically near boundaries, it is brittle. In principle, brittleness is a phenomena that can be directly observed—though once it occurs, the system in question is changed and damaged. From the perspective of sustained adaptability, the question is—what can signal brittleness ahead of the losses and damage that result from experiencing a rapid performance decline?

The risk of saturating *CfM* provides a signal of approach to the decline that can be used to initiate changes to forestall or limit the potential damage from a rapid fall off in performance. The risk of saturating *CfM* provides an actionable parameter. When risk of saturating *CfM* is high it serves as a signal for the UAB at risk and for other nearby UABs to act in ways that expand *CfM* of the unit at risk of saturation. The changes initiated to expand performance are more than just providing graceful degradation, it is a positive set of actions that serve to extend the ability to respond even as challenges change and grow—graceful extensibility.

Statement 7 highlights processes that extend adaptive capacity when challenges arise. The risk of saturating control becomes the trigger for regulation of adaptive capacity. In effect, risk of saturation becomes an operational definition for stress in the context of the adaptive universe. Regulation of adaptation occurs in response to the stressor of the risk of saturation "the consequences of which vary according to the nature of the challenge to be met" (Caporale and Doyle 2013, p. 22). Regulation of adaptive capacity is triggered or induced by challenge, where challenge can be an impasse or an opportunity to meet pressures and goals in new ways or to new degrees. As a potential, regulation of adaptive capacity needs to operate in anticipation of and keep pace with challenges as they arrive and build. This requires the ability to read and track the changing patterns of challenge. Thus, extenders are linked to changing demands.

2.4.2 Net adaptive value

Statement 7 helps lead us to the new concept of net adaptive value. Net adaptive value for a UAB has two interdependent parts: performance far from saturation and performance as it approaches saturation. The former is the fitness of its base adaptive capacity to well-modeled events and variations—its competence envelope. The latter is how a UAB can extend or stretch when events challenge boundaries. This means that measures of base adaptive capacity are separate from the measure of brittleness. Two different (but interdependent) measures are needed to characterize a UAB.

However, well matched at one point in time to its environment, a UAB can be insensitive to and lag behind changes in its environment. It then needs some basic and ever-present plasticity or resilience as the potential to generate change in adaptive capacities. The UAB needs to re-adjust to continue to achieve fitness as it confronts new challenges and opportunities, as its environment changes and as its relationship to other UABs changes. This is the performance attribute graceful extensibility—how will a UAB behave when it confronts situations that fall outside its base range or competence envelope, that is, when risk of saturation is high?

Graceful extensibility captures the second part of net adaptive value. Systems with high graceful extensibility have capabilities to anticipate challenges ahead, to learn about the changing shape of disturbances and possess the readiness-to-respond to adjust responses to fit the challenges ahead (Finkel 2011; Woods et al. 2013). But there is a limit on how much any one UAB can extend its own *CfM*. Thus, the question becomes how do other nearby UABs adapt to expand the *CfM* of the UAB at risk of saturation. Effectively, a system is resilient, in one sense of this label, when graceful extensibility is high (relative to changing demands); a system is brittle when graceful extensibility is low. The key process is the ability to regulate *CfM* as risk of saturation varies.

Resilience as graceful extensibility asks the question: how does a system function and adapt when events produce challenges at and beyond its boundaries? 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. Hence, the empirical foundation for the theory comes from analyzing past cycles of adaptation to disrupting events and analyzing how the system stretched to accommodate or take advantage of the reverberations arising from those events.

Studies of how systems extend adaptive capacity to handle surprise have led to characterization of basic patterns in how adaptive systems fail, or their reverse, key capabilities that are needed to avoid these risks (Woods and Branlat 2011). These three basic patterns have emerged naturally from the theory. The first pattern is exhausting the capacity to deploy and mobilize responses as disturbances grow and cascade—decompensation. 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 (Scheffer et al. 2009; Dai et al. 2012). When the time to recovery increases (and/or there is a decrease in the level recovered to), this pattern indicates that a system is exhausting its ability to handle growing or repeated challenges. There are many other indicators of the risk of decompensation. Studies of systems that reduce the risk of decompensation provide valuable insight about where to invest to produce graceful extensibility. For example, Finkel (2011) 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. They provide a systems with the capability, in advance, to handle classes of surprises or challenges such as cascading events. Sources of resilience undergird this capability and providing/sustaining these sources has its own dynamics and difficulties that arise from fundamental trade-offs (Woods 2006). For example, mis-calibration can lead organizations to undermine, inadvertently, their own sources of resilience as they miss how people step into the breach to make up for shortfalls in adaptive capacity (Stephens et al. 2015).

The two parts of net adaptive value, then, are robust optimality and graceful extensibility. Each operates according to a performance/resource relationship (P/R ratio), but different ones though interdependent ones. Improving UAB performance far from saturation consumes resources and capabilities in ways that change how that UAB acts near saturation. And the inverse operates over temporal scales as well: allocating resources to improve performance near saturation undermines measures of and pressures on performance (efficiency) far from saturation. In other words, there are two cost/benefit, or performance/resource, curves that capture systems (near and far from saturation); two measures are needed to characterize adaptive systems; and the two measures are interdependent and trade-off. Net adaptive value captures how both robust optimality and graceful extensibility are balanced for a system (e.g., the concept of robust yet fragile; Csete and Doyle 2002). Note a system can be improving on robust optimality yet decreasing on graceful extensibility, as well as gaining on graceful extensibility while performing lower relative to criteria on robust optimality. The larger question for sustained adaptability is what architectures can continue to shift this trade-off closer to hard limits for the system in question, i.e., boost both contributors to net adaptive value (Doyle and Csete 2011).

One way to see the interaction between robust optimality and graceful extensibility is to look at the resources required to improve performance near and far from saturation. Reducing the resources that support base adaptive capacity leads to better scores on aspects and indicators of robust optimality. There are resources needed to support graceful extensibility as well. The difficulty arises because graceful extensibility represents a potential for future adaptive action—adaptive capacity that must be present before disrupting events call upon that capacity. This means that the resources that support graceful extensibility will be seen as valuable only if the disrupting events are experienced regularly or are tangible to the different UABs in the network (i.e., in high potential for surprise situations). Prior to visible disrupting events, UABs at broader echelons can see the resources that support graceful extensibility as underused or excess, and these become targets for resource reductions to improve robust optimality. Thus, pursuit of improving base adaptive capacity (pressure to be faster, better, cheaper) leads to increased brittleness as the resource efficiencies will also reduce the resources and capacities needed to support graceful extensibility [the acute-chronic trade-off; see Woods (2006) for descriptions of this dynamic in action, and Woods (2005) provides a tangible tragic case description for the lead up to the Columbia space shuttle accident].

On the other hand, when the need for better performance near saturation is salient (usually after a sudden collapse or failure has already occurred) a large set of extra resources become available to support many different capabilities related to graceful extensibility, for a time. Investing resources may increase performance on graceful extensibility but the ratio of performance to resources is poor if the resources are large, smeared all about, and imprecisely targeted. When the ratio of performance to resources is poor, the investment in improving graceful extensibility will prove unsustainable and the system in question will slide back into a brittle state (see the history of NASA's failures and responses and the discussion in Woods 2005, 2006). Reserves are required for graceful extensibility but, as in military history, these need to be targeted. Reserves need certain capabilities designed to handle the regularities about surprise, anticipating surprise, and how to flexibly respond to surprise (Finkel 2011; Chuang 2018). The characteristics of reserves that support and create the capability for graceful extensibility remain to be worked out. Sustained adaptability requires a network architecture that can continue to find an effective balance between improving robust optimality while readjusting capabilities for graceful extensibility as pressures, resources, and demands change.

2.4.3 Mapping a system in terms of UABs?

As in all forms of systems analysis, mapping a system as a (tangled) layered network of UABs is a matter of perspective and purpose. For some outside perspective driving an analysis, what criteria can be used in this process?

To decide whether a node in a network functions as a unit of adaptive behavior, ask—does the provisional unit in question have some capability to continue to extend its performance when risk of saturation is high? All UABs require some capacity to extend capacity for maneuver in the face of risk of saturation. If the unit has no ability to extend then it may be a control agent but it does not really rise to the level of the unit of adaptive behavior. If a provisional unit has no such capability it is important to re-define the UAB with a broader scope that draws in neighboring units which do introduce some ability to extend responses in the face of risk of saturation. The constraint is no single UAB by itself regardless of scale has sufficient ability to continue to adapt in the face of risk of saturation. Nevertheless, every UAB needs some capability of its own, but this capability, in and of itself, is always incomplete, in principle.

UABs cover multiple echelons. Some "upper" echelons operate more distantly from the physical processes at work, whereas lower echelons operate close to points of action in the world. Coordination vertically across echelons is needed and the form of coordination vertically changes relative to the risk of saturation as specified by the theory. For examples of this vertical interplay both successfully and unsuccessfully see the analysis of contrasting military cases in Finkel (2011).

What is the role of upper echelons in sustained adaptability? Like the stress response system in physiology, upper echelons monitor the relationship between upcoming demands and response capability to continuously assess the risk of saturation. When risk of saturation is high or increasing, the upper echelon UAB acts to increase capacity for maneuver by changing relationships, invoking new processes, and bringing to bear new resources. This means S5 and S6 are quite essential to successful sustained adaptability. If you ask what should management do in some human system, i.e., an upper echelon UAB, it should act to regulate the *CfM* of other UABs based on monitoring the risk of saturation. Regulating *CfM* comes from adjusting the interdependencies in the network to remove potential constrictions and to enhance relationships that expand *CfM* for the UAB at risk of saturation. In doing this, the upper echelon UAB redefines the composition of the network as a set of interacting UABs.

Often upper echelon UABs (and those who would select or design an architecture for a network of UABs) assume the network can be regulated by a simple switching mechanism transferring control from one lower echelon UAB to another. The common example is where the system either works in automatic mode or switches to backup human manual mode (the system is intended to operate in just one or the other). This is an extremely crude and not particularly effective architecture. Despite a long record of not being effective, designers almost universally select this limited architecture as a starting point. However, it is unstable from the point of view of sustained adaptability. This architecture always is subject to a significant shortfall in adaptive capacity that invokes a stopgap response from responsible human roles in the system. That this common choice turns out badly highlights S5. Some upper echelon UAB needs to monitor the risk of saturation of other parts of the network and, when that risk is too high, it needs to act to change the relationships across the network and change the portfolio of resource investments to support the potential for adaptive action.

Note this process happens at a broad temporal scale where upper echelon units learn proactively how, where, when extra *CfM* is needed for graceful extensibility. As a result, upper echelon units can generate in advance the readiness to respond in the form of what capabilities are ready to deploy or are mobilizable, when future surprise events occur. The upper echelons are making, modifying, and sustaining investments in graceful extensibility as part of balancing net adaptive value.

One example of this comes from Deary's study of how a large transportation firm had learned to reconfigure relationships across roles and layers to keep pace with unpredictable demands. In particular, Deary was able to observe how the organization used these techniques during Hurricane Sandy in the fall of 2012 (Deary et al. 2013). To adapt effectively, the organization had to re-prioritize over multiple conflicting goals, sacrifice cost control processes in the face of safety risks, value timely responsive decisions and actions, coordinate horizontally across functions to reduce the risk of missing critical information or side effects when replanning under time pressure, control the cost of coordination to avoid overloading already busy people and communication channels, and push initiative and authority down to the lowest unit of action in the situation to increase the readiness to respond when new challenges arose. New temporary teams were created quickly to provide critical information updates (weather impact analysis teams). They stood up temporary local command centers where key personnel from different functions worked together to keep track of the evolving situation and re-plan. The horizontal coordination in these centers worked to balance the efficiency-thoroughness trade-off (Hollnagel 2009) in a new way for a situation that presented surprising challenges and demanded high responsiveness. In the case of disruptions, this highly adaptive firm was able to synchronize different groups at different echelons even with time pressure, surprises, goals conflicts, and trade-offs. The firm had developed mechanisms to keep pace with cascades and expand/speed coordination across roles, though these sacrificed economics and standard processes, because this firm's business model, environment, clientele, and external events regularly required adaptation even though critical events or periods occurred less often (Deary 2015).

Biology also speaks to the processes by which units in a network change and develop over cycles of challenge and adaptation (Bonner 1998). During an evolutionary transition, for example, from single cells to multicellular organisms, the new integrated unit gains its emergent properties by virtue of cooperative interactions among the local units. Only means for synchronization transfers fitness from the set of local units to the new integrated unit (though this transition transfers some costs to the local level). As integration of units creates new levels of potential fitness, it also creates conflict between local and integrated levels. Hence, the need for architectural principles that guide development to align goals, balance sacrifices, and manage basic trade-offs so that the changing network will be able to continue to adapt and evolve as changes continues. The discussion of reciprocity and governing the expression of initiative provide exemplars of some of the factors at work.

The capability of some upper echelon UABs to expand the *CfM* of others is a defining characteristic of architectures that can sustain adaptability over multiple cycles of change. The basic policy is to empower decentralized initiative at lower echelons, reward reciprocity across units, and then to coordinate their activities and relationships over emerging trends to meet changing priorities. This policy is demonstrated in Finkel's 2011 analysis which shows how organizations create the basis for adapting successfully to surprise in military operations. The policy is also illustrated in neurobiology at a far different scale (Brenner et al. 2000; Wark et al. 2007). Ostrom's work on how some human systems are able to avoid the tragedy of the commons (Dietz et al. 2003; Ostrom 2003, 2012) refers to the capability of some UABs to expand the *CfM* of other units in poly-centric governance. Understanding how to build and sustain this capability over change is central to architectures for sustained adaptability. It is quite important to re-state the basic policy: empower decentralized initiative at lower echelons, reward reciprocity across units, and then provide means to synchronize their activities and relationships over emerging trends to meet changing priorities among goals.

UABs are nested and composable over scales. This means the outside analyst has degrees of freedom in how they aggregate and decompose candidates for UABs when mapping a network. The starting scope for mapping a network may draw the nesting of UABs too narrowly or too broadly. If one draws the scope of the UAB too narrowly they may find the provisional UAB has very limited capability to extend its *CfM* when risk of saturation is high; as a result, the analyst should aggregate the UAB to include additional nodes in the network which provide some capability to extend *CfM*.

An example of drawing UABs too broadly occurs with an apparently autonomous vehicle. In this case, the UAB is defined based on the physical platform. But this is much too broad hiding many important functions and relationships that effect its adaptive capacities. Instead, mapping the network of UABs requires decomposing the vehicle as it serves as a platform that is made up of many different and interacting on-board sensors, control and information processing computations and algorithms (Woods 2016). In such cases, the analyst should decompose the original unit into a network that makes explicit the interactions among UABs that are on-board and the interactions with various off-board UABs, both human roles and other machines, when disrupting events and surprises occur. In the cases I have mapped, apparently simple autonomous units turn out, when in contact with real-world variability, to be decomposed into dozens of UABs interacting over 2–3 echelons and to need assistance from multiple off-board units much more often than designers anticipated. The appropriate network map changes when anomalies occur and become visible. Far from saturation, a more aggregated, small set of UABs is sufficient to characterize a network. As anomalies occur, the few aggregated UABs need to be decomposed to see a more fine-grained map of units, interactions, and interdependencies.

Mis-calibration, S10, leads to the requirement that UABs possess some reflective capability in order to produce sustained adaptability. Reflective capability is operationalized as the ability of a UAB to monitor its own risk of saturation. Mis-calibration also addresses the ability of a UAB to monitor the risk of saturation in other UABs. Monitoring the risk of saturation requires a model, either of itself as a UAB or other UABs, and as always is the case, models are at risk of mis-calibration. The constraint, as for all agents that function via models, is that UABs must have some ability to invest energy in *revising* or updating their models as information comes in or is searched for (Woods 2017). This monitoring function re-assesses the model's ability to capture what is changing in the world, not simply the ability to monitor the world itself. While models are necessary, they will be surprised.

The limits on models expressed in the theory go even further. All models become stale as soon as they are embodied and deployed in some form to guide behavior, i.e., a model's fitness declines as soon as it makes contact with the variability of the world and the adaptive behavior of other UABs in the neighboring network.

2.5 Logic of expression

The logic of expression in the exposition of the theoretical ideas flows in several ways given how I have structured the set of 10 statements. First, each statement is a logical claim, e.g., given bounds are universal, surprise is ongoing. Second, there are logical claims expressed implicitly in the flow from one statement to others. For example, the assumptions lead to S1 and S2, then S1 and S2 lead to S3. Third, some of the logical force derives from the set as a whole (or chunks of it). For example, the logic of the whole set reveals that boundaries are dynamic and uncertain, yet estimates of the boundaries are misplaced and overconfident. As a result, sustained adaptability requires a drive to explore for boundaries and to discover how they shift in an indeterminate "borderland."

Fourth, combinations of the 10 are inter-linked in interesting ways. S4 and S5 arise when the assumptions are re-applied to S1 to S3. S1 expresses a limit; S2 expresses that this limit matters as challenge events will occur; then S8, S9, and S10 expand on the limits of any unit at any scale. S3, S5, S6, and S9 specify the means available to adapt despite the constraints captured in the set. These four express how limited units embedded in neighborhoods of additional adaptive units can act productively despite the limits, if the network exhibits the right properties. Actually the point is much stronger: the adaptive capabilities emerge because of the limits. Following Kirschner and Gerhart (2005), imposing a constraint (of the right kind and at the right place) forces adaptation relative to the constraint, and thus produces the paradoxical effect of releasing new capabilities. The capabilities released then have the effect of extending performance (deconstrain) beyond the limits inherent in the constraints. The set of 10 identifies constraints that lead to the varieties of emergent adaptive behavior which characterize the rules for generating sustained adaptability and overcoming the risk of adaptive breakdowns.

3 General discussion

The theory arises from a fundamental reframing. Instead of seeing the world as linear or usually operating in the linear region so that stakeholders need only to deal with complexity as a special case, the theory is based on reversing the framing to make complexity the universal condition and starting point. Linear islands are special cases carved out temporarily, and these islands take extra energy to maintain relative to the omnipresent complexities nipping at their heels. Analogous to entropy—inevitable push toward disorder in the absence of energy investments—the theory, with the trade-offs and the laws that result, asserts that complexities will grow and dominate performance in the absence of continuing energy investments to regulate the varieties of adaptive capacity in the face of changing risks of saturation.

3.1 Architectures that balance net adaptive value

Attempts to sidestep the constraints in the theory are commonplace. Pursuing continuous improvements, deploying new computational advances (e.g., the current rhetoric about the potential of machine learning technologies), adding factors to correct linear approaches for complexities, all of these have been claimed to make the constraints captured above moot. But the point of S1 and S2 is that there is no place to hide from the constraints that every thing has bounds and there are events happening outside those bounds—model surprise. It does not matter what approach is taken, these constraints hold; hence, any and all approaches can saturate. This means there are two parallel regimes: one far from saturation with one kind of performance measures and one near saturation with different, but interdependent, performance measures.

Any system needs processes that function near saturation to provide graceful extensibility—otherwise it will prove to be too brittle to be viable in the long run—*viability requires extensibility*. Now those processes that function near saturation can be modeled using various formal machinery too. Let us take one as an example: anomaly recognition. Anomaly recognition can be modeled in different ways. For example, with statistical methods for recognition of a change from previous data. There are biological systems which perform this function and which can be investigated to uncover the underlying functions. For example, the brain does anomaly recognition (Wark et al. 2009), and one general brain process that contributes to this capability is adaptation level (a concept that is quite old in neuropsychology; Attneave 1954). There are various attempts emerging to capture how neurocomputations might perform anomaly recognition (Bialek et al. 2007). These become starting points to develop formal and general models, though these have to work at multiple and different scales that go beyond an explanation at just one, such as the level of neurological function.

In the end, behavior near saturation is different from behavior far from saturation. Modeling behavior near saturation can result in quite different explanatory models than those that capture behavior far from saturation. Both are necessary forms of adaptation. Interestingly, this has been noted in general as a capability of the brain for a long time—the ability to continually improve performance to frequently encountered stimuli, while at the same time remaining sensitive to what is new (being able to recognize what is different from previous or expected) and being able learn and change to handle these. Plus, those who have described these two forms of adaptation have also understood that there is a potential negative interaction between them—when one dominates the other, it leads to increased vulnerability to failures due to either under- or over-adaptation.

The theory works in part because it provides a minimal set of concepts that logically express how adaptive systems have these two basic capabilities. The theory also reveals the interdependence between these capabilities and therefore that there is a trade-off (e.g., robust yet fragile emerges naturally from the first few statements; Doyle and Csete 2011). Many approaches to improve performance far from saturation have the unintended consequence of degrading performance near saturation. Several dynamic patterns have been observed that produce this trade-off in human systems and the patterns seem to be tied to a lack of tangible experience with surprise. The sense of precariousness that can accompany these surprise experiences provides a mechanism for reducing mis-calibration about boundaries and the risk of saturation.

Some capability is required for graceful extensibility in any adaptive unit. This requirement means there are at minimum two performance measures—one for far from saturation and one for near saturation. And the second measure cannott be zero (if the second measure is zero, then the item of interest is not an adaptive unit; rather, it is only a piece of a larger adaptive unit). However, the graceful extensibility of any single adaptive unit is limited and needs to be supported by responses from other nearby units in the network. The concept of net adaptive value captures how sustained adaptability requires systems that can achieve a suitable score and balance on both types of performance measures in parallel.

Good architectures are able to continue to re-balance the two forms of adaptive capacity to sustain performance over the long term. Hence, the search for the key architectural properties that will produce sustained adaptability. Among the questions to consider are as follows:

- How to assess whether there is enough graceful extensibility and whether the right kind is present?
- How to sustain graceful extensibility in architectures that can re-balance the two forms of adaptive capacity?

Graceful extensibility involves interactions across neighbors in the network:

- How can those interactions change to increase extensibility when part of the network risks saturation?
- How do other forms of interaction across neighbors reduce graceful extensibility when part of the network risks saturation?

3.2 Expanding the base competence envelope is not sufficient

The usual assumption is that improving base adaptive capacity far from saturation consumes a larger and larger share of the variability to be accommodated, leaving the residual variability to be rare events and "unknown unknowns" that can be safely downplayed or disregarded. When these rare events occur, in hindsight many signals can be seen which indicated this risk was present and growing. In other words, the previously discounted rare event turns out to be part of a class of events that are not so rare after all (S2). Nevertheless, the standard assumption leads to claims of, "if we just invest in this or another technology, computational mechanism, or automata then we will expand the competence envelope to encompass this challenge event, expanding the envelope so that the remaining challenges and variabilities become sufficiently rare again."

This assumption is mistaken, and the theory makes this clear. Improving performance far from saturation does not consume a larger and larger share of the variability to be accommodated. There are always events and changes occurring outside the current envelope that challenge and fall outside of base capacity to handle. Why? Because the potential for surprise to occur is omnipresent, the potential for events to push base adaptive capacity to near saturation is ongoing. There is always some risk of saturation and thus there is always the need for some capacity for graceful extensibility as events signal or increase the risk of saturation. Without this second capability, the system will be brittle in the face of surprise, risking collapses in performance and threaten long-term viability. Again, improving performance far from saturation does not change the need for adaptive capacity that comes into play near saturation to produce graceful extensibility. Hence, the need to invoke net adaptive value as a measure that combines both forms of adaptive capacity and how they trade-off.

But the situation is even worse with respect to the standard assumption of ever-expanding base competence. The continuing need for graceful extensibility also arises as a result of improving base adaptive capacity far from saturation. This is the fundamental trade-off between robust optimality and sustained adaptability (or robust yet fragile).

Continual improvement only operates far from saturation. It does not continually reduce the potential for surprise. But such improvements do change facets of the potential for surprise leading to the need to re-adjust the response capabilities (and required resources) needed to handle the new forms of surprise. Let us say this again—continual improvement does not guarantee a reduction in the potential for surprise, but continual improvement does change what contributes to the potential for surprise and therefore what response capabilities are needed to support graceful extensibility. Thus, some capacity for graceful extensibility is always present and needed; the nature of that capacity, and the resources required, moves around as the world changes and as the base capacity changes.

> A Reminder: The potential for surprise asks the question—what is the likelihood that the next event or next period of operation will present a challenge to the base capacity (Woods and Hollnagel 2006)?

Thus, as base adaptive capacity changes, the risk of saturation remains real. Changes in base adaptive capacity affect what threatens saturation and affect where and how graceful extensibility is needed to produce sustained adaptability. Doyle makes the case formally in a series of analyses with multiple colleagues, and I make the empirical case in the "essentials" chapter of the first Resilience Engineering book in 2006. There is some match or fitness between the response capabilities of a system and the range of situations and variations it experiences. The set where this is well-matched defines the range of adaptive behavior of a system. This range has limits, events occur outside this range—potential for surprise. Hence, there is a need for a second range of behavior capable of extending response capabilities when events challenge or go beyond the boundaries of the first range. Pursuing performance improvements far from saturation in isolation is likely to produce unintended side effects that undermine graceful extensibility—reducing net adaptive value. The theory highlights how both forms, base and extensible, are necessary. Architectures capable of sustained adaptability are able to achieve and balance both graceful extensibility (near saturation performance) and robust optimality (performance far from saturation) to achieve net adaptive value.

As a result, the theory of graceful extensibility reframes optimality. The pursuit of optimization is a form of pressure on units that arises from other units/levels in the network. It is one form of pressure arising within TLNs and experienced by units in TLNs (Subset B and S6). The need to respond to changing pressures given conflicting goals is omnipresent in networks of adaptive units. Computations, however justified, miss the reality of being caught in a squeeze between conflicting goals as pressures ramp up. The assumption is the right computations use a policy that provides a best solution to the conflict and pressures—the end of the story. But adaptive cycles are always stories about how pressures and conflicts stimulate adaptations. As pressures ramp up, squeezes intensify; one critical question is what goals get sacrificed first and which last (Woods 2006). This process of re-prioritization gets lost because all revision related activities are eliminated or subsumed in the computations. The theory, really any theory of the adaptive universe, flips the starting point—adaptive capacities are about the potential to shift—revise—what worked previously as change continues—being poised to adapt.

3.3 Escaping from the simplification of central command

Whenever stakeholders consider TLNs, there is a common and almost overwhelming temptation to believe in the need for a central authority or dominating command node in a hierarchy. Decentralization, and especially pushing initiative closer to points of action in the world, appears too uncertain. So responsible units, far from points of action, try to dominate uncertainty and variability. The harder they squeeze to guarantee their intentions are fulfilled by other parts of the tangled layered network they are part of, the more the performance they seek slips out of their hands (then they blame the unintended effects on human error).

The remedy begins with S8—every adaptive unit is local. It is interesting that I had to make this explicit as a statement in the theory, and, when this is explicit as a fundamental condition, it causes a great deal of trouble for conventional thinking. One powerful implication of the bounds on perspective (S9) is that it directly rejects the possibility of any dominating command node in any tangled layered network. Then S10 adds that the risk of mis-calibration is so great, the base state is one of misunderstanding, with continuous effort needed to keep the degree of mis-calibration under control (for a setting where this is clear see Woods 2017). Every unit is local, only sees part of the world that matters, and its model of the world is always limited (and therefore in danger of being wrong in important ways). But all of this is really derives from the first two statements (bounds and surprise), given the simple two premises of finite resources and change.

The idea that adaptive units are reflective is potentially controversial. The theory as constructed requires UABs to be reflective, i.e., they have models about their capabilities and fitness relative to the world and nearby units. But this result is not as strong a commitment as it might seem—see the classic Conant and Ashby's 1970 paper on how every controller is a model. However, the theory produces an important shift to this classic result—being reflective requires continual effort to reduce mis-calibration.

At this point the objection shifts: "how can any network of adaptive units be designed to work well without a comprehensive view or dominating command view?" The escape route is *CfM* as a control parameter—units can monitor and regulate *CfM* and the risk of saturation (yes its a trick, but biology shows it's a good trick for TLNs). The ability to control *CfM* and the risk of saturation really exists, but this ability is limited for each adaptive unit thus requiring extension by neighboring units in the network of UABs, assuming the network possesses the right constraints on how they interact. To escape requires the ability to control *CfM* and the risk of saturation! This is part of why I claim the set of 10 statements could be shown to be formal theorems, not just empirical findings that some networks are able to do this (existence proofs). As in all forms of control theory, this account provides the demarcation between good and poor control that is a requirement for such theories.

Resilient control is local and limited yet must be able to balance mechanisms that improve robust optimality when conditions are far from saturation combined with the capacity for graceful extensibility when conditions are near saturation. It is important to note that adaptive units in a tangled layered network are able to continue to demonstrate graceful extensibility in the face of change and surprise through two layers of local action (an idea with roots in Ashby 1956): first there is a UAB adapting at some level or scale and second there is another layer of local action going on but on top of the first—a neighborhood in a network relative to the unit of interest. Thus, the escape from the need for a omniscient designer/commander is through coordination across these two layers of local action—a UAB of interest and a neighborhood of interdependent units horizontally and vertically that includes the unit of interest. The contrast and interplay across these two kinds of perspectives—local and neighboring—is a critical process. Different architectures attempt to characterize how this relationship produces or fails to produce sustained adaptability over cycles of change. The search for architectural principles for TLNs focuses on learning what are the cross level constraints between these two layers that provide the flexibility for sustained adaptability—what architectural constraints stimulate the continued capacity to adapt to continuing change. For example, history strongly shows empirically that moving information vertically across layers as a requirement for adapting plans and activities to changing situations fails to keep pace with the tempo of events. When the process for revision only runs vertically through layers, the risk of decompensation failures is high—adaptations will be too slow and stale (e.g., Woods and Branlat 2011; Finkel 2011).

3.4 Escaping Archimedes trap

The bound on perspective, S9, seems to limit our ability to understand the adaptive universe. To talk about the adaptive universe—to comment on or reflect on how any TLN works, one has to adopt a perspective outside it, but the theory says this cannot be done cleanly. The apparently outside perspective still is from some place in some neighborhood subject to the constraints in the theory. In other words, to see the limits and possibilities of one perspective requires another one to see what is obscured in the first one—perspective contrast. This is why moving the point of observation is so powerful at revealing the structure of the world in human perception (see Morison et al. 2015 for concrete illustrations of this for people using robots with sensors to explore distant scenes). While each perspective is limited, contrast across perspectives provides a way to overcome the limits of each. Perspective contrast occurs over different units each local and at risk of mis-calibration, as well as perspective shifts that occur over time for a single unit. The outside perspective is not really outside but rather part of a network that overlaps with and influences the original neighborhood of interest.

But the combination of S8, S9, and S10 goes much further in constraining the search for good architectures. To model a TLN appears to require adopting a perspective outside that network. As Archimedes is attributed to have said, "give me a place to stand and, with a lever, I will move the earth." Similarly for the adaptive universe, modelers assume or assert they occupy a position outside the system/network they model. In this way they can see the model as optimal if one organizes the activities of that network to match the model—not the actual world where the network exists. The trap for Archimedes is that there is no place to stand outside of the Earth. The trap for modeling is:

> The act of modeling expands the network modeled to encompass the modeler.

In modeling, the modeler becomes part of the network in the adaptive universe, expanding the network to include the modeler and a region of units related to the modeler. Modelers when setting up algorithms to optimize or learn or control do not model themselves as part of the system and think they exist outside the system of interest—optimization assumes a nearly omniscient modeler. One might say, like Archimedes, they are asserting there is a virtual "omniscience" point outside any system where the modeler resides—an Archimedes Point. For sustained adaptability of TLNs, the goal is to understand how adaptation to changing constraints (including opportunities) can go on without recourse to any outside modeler, designer, or commander.

The paradox is bold: modeling appears to require a perspective outside the region of the network to be modeled, but using that model to imply or assert changes to the network means that the modeler becomes a unit in the network of interest subject to the limits expressed in the theory. The goal is defined usually as building the right model, or building a better model than previously, but bounds, change, and surprise, eventually will lead to events that challenges the limits of the model (surprise is model surprise). The third subset of theory (mis-calibration S10 plus locality and bounds on perspective S8/S9) points out the critical capability in the adaptive universe is the ability to *revise or update the model in pace with change* (Woods 2017).

Thus, the theory contains the route for escape from the trap. Its not the place we stand nor where we position our view that matters; what matters is the contrast that emerges when we shift perspective. The constraints on perspective end up eliminating any possibility in the real world of a dominating omniscient perspective from which one can see all, know all, and do all. In the adaptive universe it is not how well you have performed, its the ability to change from what has worked previously. Thus, the logic of the theory returns to the need for some graceful extensibility at the boundaries of previous function, the constraints on graceful extensibility, and the need to anticipate the effect of changes on net adaptive value.

3.5 Buffering is insufficient

The biggest driver of the need for graceful extensibility is past success—to put it concisely, *adaptive behavior hijacks success*. This was originally captured as the Law of Stretched Systems (Woods and Hollnagel 2006; Hoffman and Woods 2011). Success, defined as improvements on some attributes over some subset of the TLN, creates opportunities for other levels or parts of the TLN. The result is not extra room relative to boundaries (bigger margins or larger buffers)—at least not for long. Instead, the follow-on adaptations result in the improvement being consumed, leaving the unit originally gifted to operate again under pressure to meet new demands with barely adequate resources that force trade-offs. Instead of reducing the risk of saturation for the unit in question, the change serves only to move around when, where and how the unit risks saturation. Relaxing pressure or increasing the ratio of resources relative to demands is not stable or sustainable over longer horizons regardless of how any improvement relaxes the ratio of performance relative to resources in the short run. In other words brittleness will grow again, as in the adaptive universe, pressures reassert from within and from without the neighborhood of the network (S6). In the end, regardless of improvements, risk of saturation remains even if the specifics of how saturation occurs change. Units in the TLN have to have the separate capability to manage risk of saturation; the general form of this capability remains the same even as the detailed expression varies to meet specific threats specific to the network at each stage of its evolution. As in biology, specialization changes the environment in ways that eventually produce new challenges to the viability of the more specialized (or more optimal) unit.

The key then is understanding how UABs interact when at least one of them is at risk of saturation.

- How do they interact to produce graceful extensibility?
- How does the unit at risk signal others and recruit their involvement in ways that extend the *CfM* of that unit or of the inter-related set within the TLN relative to the demands ongoing and ahead?
- How do other nearby units recognize that one is at risk of saturating *CfM* and modify their activity or relationship to extend the *CfM* of the unit at risk or the related set of units?
- How do nearby units act to constrict *CfM* of other unit when they are at risk of saturation or to enable additional *CfM* when another interdependent unit is at risk of saturation?

3.6 The network indeterminacy principle

The above points that emerge from the theory of graceful extensibility pack one more surprise—the Network Indeterminacy Principle: networks of adaptive units are fluid, changing as they are modeled and changing as adaptation occurs (e.g., Ormerod and Colbaugh 2006). What moves in the landscape will provide adaptive value, change. The value of changes is modified (a) as other units move, (b) as the network changes, (c) as the environment changes, and (d) as units' understanding of the network they are part of changes. Archimedes trap is one facet—modeling changes the network modeled. The Law of Stretched Systems is another facet—adaptive behavior hijacks past successes leading to new pressures and relationships. But current approaches to understand complex networks violate the Network Indeterminacy Principle when they assume nodes and links are fixed rather than fluid and evolving.

4 Future directions

The theory is an integration of results from control theory, distributed cognitive systems, and human organizations, with some hints from biology, as systems that serve human purposes increasingly exist in a tangled and layered network of interdependent and adaptive units. The result is the theory of graceful extensibility which sets out basic rules about how networks of adaptive units continually search for answers to the question of what is fitness.

The theory of graceful extensibility provides a structure that explains formally many observed empirical patterns. This paper is not intended to cover the explanatory linkage from the theory to observed phenomena of brittleness and resilience in action. This paper is needed first as it lays out the theory's key ideas and logic. In this process, I have made reference to a variety of empirical findings as they have played a role in developing the key ideas. Future work can then propose and test the explanatory power. Even more important though is the power of the theory to influence architecture: how does it help discover architectural principles for regulating networks in complex settings so that adaptive power can continue to search for fitness over multiple cycles of change (Alderson and Doyle 2010).

The theory of graceful extensibility provides an account that addresses the six desiderata set out at the beginning of this paper. Important trade-offs emerge directly from the theory (e.g., robust yet fragile), while others were surprisingly necessary as base statements in themselves (e.g., bounds on perspective). The three forms of how adaptive systems fail identified by Woods and Branlat (2011) emerge naturally from the theory. The theory provides the concept of capacity for maneuver (*CfM*)/risk of saturation as the scalable, positive means for control for all adaptive units and for neighborhoods of interacting units. Empirical evidence indicates that human systems do monitor and regulate *CfM* so as to reduce the risk of saturation, and new kinds of control systems can be developed that utilize this construct for specific but diverse settings (Fariadian et al. 2018). This concept of capacity for maneuver leads to a new operational definition of brittleness. This provides criteria for desirable network properties, those that extend rather than constrict capacity for maneuver when a unit is at risk of saturation. As a result, some network properties emerge as critical, e.g., reciprocity, which also have been identified in the empirical literature. Reciprocity illustrates how the theory integrates some social science findings—Ostrom's work, with results in control theory—saturation, and generates a new and more actionable operational definition of this basic concept. The theory provides a starting point for developing new ways to assess unintended reverberations in TLNs—one notable pattern, how pursuit of optimization by some units can increase pressures on other units in ways that reduce reciprocity and therefore reduce the graceful extensibility needed to handle surprise events. The theory shows that distant roles will overestimate what the base envelope can handle competently and underestimate the potential for surprise leading to reduced graceful extensibility and lower net adaptive value. The theory shows that this relationship is more than an occasional empirical occurrence, but rather it is emergent system property independent of specific people or organizations that requires architectural safeguards.

The theory is a starting point. How valuable is this one account? Is it only provisional, or does it capture a portion of the fundamentals? One hope is that attempts like this one define a baseline that serves as a floor for other potential integrations and comprehensive explanations. Today there is too much noise generated as there are so many lines of inquiry starting from so many different backgrounds with very different purposes for different stakeholders (IRGC 2016). Another hope is this paper does express a minimal set of basics that can produce explanations for a very wide set of regularities and patterns—and can lead inquiry to uncover new patterns as well. Third, and perhaps most important, the potential value of this set depends on whether it can serve to guide the development of TLNs that lean toward architectures that sustain adaptive capacities and away from architectures that undermine them.

At the heart of the theory of graceful extensibility is the fundamental concept of managing risk of saturation via regulating the capacity for maneuver—both at the level of an adaptive unit and at the level of a network where neighboring adaptive units interact as risk of saturation increases. This idea is put forward as a general and actionable control mechanism for TLNs of adaptive units of all types and over all scales. Initial work already demonstrates the potential for this concept to lead to new types of resilient control mechanisms (Fariadian et al. 2018).

Are the statements theorems or empirical generalizations? Are they the fundamental bedrock or do they only mark out way posts in the search for the fundamental driving properties of the adaptive universe? Do the characteristics highlighted only apply to the sphere of human adaptive systems or do they cover all networks of adaptive systems regardless of scale or inclusion of human roles and organizations? Whatever paths of inquiry reveal in the future, this attempt at capturing fundamentals offers a provisional comprehensive statement for debate as researchers search for the bedrock of how the adaptive universe works.

Acknowledgements The development of the theory has benefited from reactions, both positive and negative, from many people engaged in debates about complexity, safety, and resilience. I would like to thank the Resilience Engineering communities and all of those who have invited me to discuss and speak on the theory for stimulating inter-disciplinary lines of inquiry to make risky systems less brittle and more resilient. This research is supported in part by funding from National Science Foundation, Grant No. 1549815 and Department of Transportation DTRT13-G-UTC47.

References

Alderson DL, Doyle JC (2010) Contrasting views of complexity and their implications for network-centric infrastructures. IEEE SMC A 40:839–852
Alderson DL, Brown GG, Carlyle M (2015) Operational models of infrastructure resilience. Risk Anal 35:4. https://doi.org/10.1111/risa.12333
Ashby (1956) An introduction to cybernetics. Chapman & Hall, London
Astrom KJ, Murray RM (2008) Feedback systems. Princeton University Press, Princeton
Attneave F (1954) Some informational aspects of visual perception. Psychol Rev 61:183–193
Beaumont HJ, Gallie J, Kost C, Ferguson GC, Rainey PB (2009) Experimental evolution of bet hedging. Nature 462(5), 90–93. https://doi.org/10.1038/nature08504
Bialek W, de Ruyter van Steveninck, RR, Tishby N (2007) Efficient representation as a design principle for neural coding and computation. Quant Biol. arXiv:0712.4381 [q-bio.NC]
Bonner JT (1998) The origins of multicellularity. Integr Biol 1:27–36
Brenner N, Bialek W, de Ruyter van Steveninck, R (2000) Adaptive rescaling maximizes information transmission. Neuron 26:695–702
Bush SF, Hershey J, Vosburgh K (1999) *Brittle system analysis*. http://arxiv.org/pdf/cs/9904016.pdf downloaded 15 June 2006
Caporale LH, Doyle JC (2013) In Darwinian evolution, feedback from natural selection leads to biased mutations. Annals of the New York Academy of Science special issue on evolutionary dynamics and information hierarchies in biological systems. Ann Rep 1305:18–28
Carlson JM, Doyle JC (2000) Highly optimized tolerance: robustness and design in complex systems. Phys Rev Lett 84(11):2529–2532
Chandra F, Buzi G, Doyle JC (2011) Glycolytic oscillations and limits on robust efficiency. Science 333:187–192
Chen Y-Z, Huang Z-G, Zhang H-F, Eisenberg D, Seager TP, Lai Y-C (2015) Extreme events in multilayer, interdependent complex networks and control. Sci Rep 5:17277. https://doi.org/10.1038/srep17277
Chuang S, Woods DD, Ting D, Cook RI, Hsu J-C (2018) Coping with Mass Casualties: Adaptations in a hospital without adequate capacity after the Formosa Fun Coast Dust Explosion. under review.
Conant RC, Ashby WR (1970) Every good regulator of a system must be a model of that system. Int J Syst Sci 1:89–97
Cook RI (2006) Being bumpable: consequences of resource saturation and near-saturation for cognitive demands on ICU practitioners. In: Woods DD, Hollnagel E (eds) Joint cognitive systems: patterns in cognitive systems engineering. Taylor & Francis/CRC Press, Boca Raton, pp 23–35
Csete ME, Doyle JC (2002) Reverse engineering of biological complexity. Science 295:1664–1669
Dai L, Vorselen D, Korolev K, Gore J (2012) Generic indicators for loss of resilience before a tipping point leading to population collapse. Science 336(6085):1175–1177. https://doi.org/10.1126/science.1219805
Deary DS (2015) Sources of organizational resilience: Sustaining production and safety in a transportation firm. Unpublished dissertation, Ohio State University, Columbus OH
Deary DS, Walker KE, Woods DD (2013) Resilience in the Face of a Superstorm: A Transportation Firm Confronts Hurricane Sandy. In Proceedings of the human factors and ergonomics society, 57th Annual Meeting (pp. 329–333). Human factors and ergonomics society, Santa Monica, CA. https://doi.org/10.1177/1541931213571072
Dietz T, Ostrom E, Stern PC (2003) The struggle to govern the commons. Science 302(5652):1907
Doyle JC, Csete ME (2011) Architecture, constraints, and behavior. Proc Natl Acad Sci USA 2011 108(Suppl. 3):S15624–S15630
Doyle JC et al (2005) The “robust yet fragile” nature of the Internet. Proc Natl Acad Sci USA 102:14497–14502
Fairhall AL, Lewen GD, Bialek W, Van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412:787–792
Fariadian AB, Annaswamy AM, Woods DD (2016) Towards a resilient control architecture: a demonstration of bumpless re-engagement following an anomaly in flight control. In: Choi J-K, Anctil A (eds) Sustainable conoscente network. Proceedings of the international symposium on sustainable systems and Technologies (ISSN 2329-9169)
Finkel M (2011) On flexibility: recovery from technological and doctrinal surprise on the battlefield. Stanford Security Studies, Palo Alto
Hoffman RR, Woods DD (2011) Beyond Simon’s slice: five fundamental tradeoffs that bound the performance of macrocognitive work systems. IEEE Intell Syst 26(6):67–71
Hollnagel E (2009) The ETTO principle: efficiency-thoroughness trade-off: why things that go right sometimes go wrong. Ashgate, Farnham
Hollnagel E, Woods DD, Leveson N (eds) (2006) Resilience engineering: concepts and precepts. Ashgate, Aldershot
International Risk Governance Center (2016) Resource Guide on Resilience. Ecole Polytechnique Federale de Lausanne and International Risk Governance Center (IRGC). v29-07-2016, Lausanne. https://www.irgc.org/
Kirschner M, Gerhart J (2005) The plausibility of life. Yale University, New Haven
Lansing JS, Kremer JN (1993) Emergent properties of Balinese water temples. Am Anthropol 95:97–114
Mendonça D, Wallace WA (2015) Factors underlying organizational resilience: the case of electric power restoration in New York City after 11 September 2001. Reliab Eng Syst Saf 141:83–91
Meyers LA, Bull JJ (2002) Fighting change with change: adaptive variation in an uncertain world. Trends Ecol Evol 17(12):551–557
Miller A, Xiao Y (2007) Multi-level strategies to achieve resilience for an organization operating at capacity: a case study at a trauma centre. Cognit Technol Work 9:51–66
Morison A, Murphy TB, Woods DD (2015) Human-robot interaction as extending human perception to new scales. In: Hoffman RR, Hancock PA, Scerbo M, Parasuraman R, Szalma JR (eds) Handbook

of applied perception research, vol 2. Cambridge University Press, New York, p. 848–868
Narendra KS, Annaswamy AM (2005) Stable adaptive systems. Dover Publications, Mineola
Nemeth CP, Nunnally M, O'Connor M, Brandwijk M, Kowalsky J, Cook RI (2007) Regularly irregular: how groups reconcile cross-cutting agendas and demand in healthcare. Cogni Technol Work 9:139–148
Ormerod P, Colbaugh R (2006) Cascades of failure and extinction in evolving complex systems. J Artif Soc Soc Simul. http://jasss.soc.surrey.ac.uk/9/4/9.html
Ostrom E (2003) Toward a behavioral theory linking trust, reciprocity, and reputation. In: Ostrom E, Walker J (eds) Trust and reciprocity: interdisciplinary lessons from experimental research. Russell Sage Foundation, New York
Ostrom E (2012) Polycentric systems: multilevel governance involving a diversity of organizations. In: Brousseau E, Dedeurwaerdere T, Jouvet P-A, Willinger M (eds) Global environmental commons: analytical and political challenges in building governance mechanisms. Oxford University Press, Cambridge, pp 105–125
Park J, Seager TP, Rao PSC, Convertino M, Linkov I (2013) Integrating risk and resilience approaches to catastrophe management in engineering systems. Risk Anal 33(3):356–367. https://doi.org/10.1111/j.1539-6924.2012.01885.x
Patterson MD, Wears RL (2015) Resilience and precarious success. Reliab Eng Syst Saf 141:45–53
Patterson ES, Watts-Perotti JC, Woods DD (1999) Voice loops as coordination aids in space shuttle mission control. Comput Support Coop Work 8:353–371
Perry S, Wears R (2012) Underground adaptations: cases from health care. Cogn Technol Work 14:253–260. https://doi.org/10.1007/s10111-011-0207-2
Rieger CG (2010) Notional examples and benchmark aspects of a resilient control system. In: Proceedings of the IEEE, 3rd international symposium on resilient control systems (ISRCS); 2010. p. 64–71
Robbins J, Allspaw J, Krishnan K, Limoncelli T (2012) Resilience engineering: learning to embrace failure. Commun ACM, 55(11):40–47. https://doi.org/10.1145/2366316.2366331
Roth A (2008) What have we learned from market design? Hahn Lecture. Econ J 1818(March):285–310
Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V et al (2009) Early-warning signals for critical transitions. Nature 461(7260):53–59
Seager TP, Clark SS, Eisenberg DA, Thomas JE, Hinrichs MM, Kofron R, Jensen CN, McBurnett LR, Snell M, Alderson DL (2017) Redesigning resilient infrastructure research. In: Linkov I, Palma Oliveira J (eds) Resilience and risk: methods and application in environment, cyber and social domains, Springer, Dordrecht
Stephens RJ, Woods DD, Patterson ES (2015) Patient boarding in the emergency department as a symptom of complexity-induced risks. In: Wears RL, Hollnagel E, Braithwaite J (eds) Resilience in everyday clinical work. Ashgate, Farnham, pp 129–144
Vespignani A (2010) Complex networks: the fragility of interdependency. Nature 464(7291):984–985
Wark B, Lundstrom BN, Fairhall A (2007) Sensory adaptation. Curr Opin Neurobiol 17:423–429
Wark B, Fairhall A, Rieke F (2009) Timescales of inference in visual adaptation. Neuron 61:750–761. https://doi.org/10.1016/j.neuron.2009.01.019
Watts-Perotti J, Woods DD (2009) Cooperative advocacy: a strategy for integrating diverse perspectives in anomaly response. Comput Support Coop Work 18:175–198
Wears RL, Woods DD (2007) Always adapting. Ann Emerg Med 50:517–519
Wears RL, Perry SJ, Anders S, Woods DD (2008) Resilience in the Emergency Department. In: Hollnagel E, Nemeth C, Dekker SWA (eds) Resilience engineering perspectives 1: remaining sensitive to the possibility of failure. Ashgate, Aldershot, pp 193–209
Woods DD (2005) Creating foresight: lessons for resilience from Columbia. In: Starbuck WH, Farjoun M (eds) Organization at the limit: NASA and the Columbia disaster. Blackwell, Malden, pp 289–308
Woods DD (2006) Essential characteristics of resilience. In: Hollnagel E, Woods D, Leveson. N (eds) Resilience engineering: concepts and precepts. Ashgate, Aldershot, pp 21–34
Woods DD (2015) Four concepts for resilience and their implications for systems safety in the face of complexity. Reliab Eng Syst Saf 141:5–9. https://doi.org/10.1016/j.ress.2015.03.018
Woods DD (2016) The risks of autonomy: Doyle's catch. J Cognit Eng Decis Mak 10(2):131–133
Woods DD (ed) (2017) STELLA report from the SNAFU catchers workshop on coping with complexity. SNAFU catchers consortium, October 4, 2017, downloaded from stella.report 10/05/2017
Woods DD, Branlat M (2010) Hollnagel's test: being 'in control' of highly interdependent multi- layered networked systems. Cogn Technol Work 12:95–101
Woods DD, Branlat M (2011) Basic patterns in how adaptive systems fail. In: Hollnagel E, Pariès J, Woods DD, Wreathall J (eds) Resilience engineering in practice. Ashgate, Farnham, pp 127–144
Woods DD, Hollnagel E (2006) Joint cognitive systems: patterns in cognitive systems engineering. CRC Press/Taylor & Francis Group, Boca Raton
Woods DD, Shattuck LG (2000) Distant supervision—local action given the potential for surprise. Cognit Technol Work 2:242–245
Woods DD, Wreathall J (2008) Stress–strain plot as a basis for assessing system resilience. In: Hollnagel E, Nemeth C, Dekker SWA (eds) Resilience engineering perspectives 1: remaining sensitive to the possibility of failure. Ashgate, Aldershot, pp 145–161
Woods DD, Chan YJ, Wreathall J (2013) The stress–strain model of resilience operationalizes the four cornerstones of resilience engineering. In: Proceedings of the fifth international symposium on resilience engineering, resilience engineering association. Download from The Knowledge Bank. Columbus OH; http://hdl.handle.net/1811/60454 June 2013. p. 25–7
Zhou K, Doyle JC, Glover K (1996) Robust and optimal control. Prentice Hall, Englewood Cliffs