ARL-TR-10082 ● APR 2025

DEVCOM ARMY RESEARCH LABORATORY

The Measurement of Cyber Resilience in Context

by M. A. Thomas

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

ARL-TR-10082 ● APR 2025

The Measurement of Cyber Resilience in Context

M. A. Thomas
DEVCOM Army Research Laboratory

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
April 2025	Technical Report	Apr 2024	Feb 2025

4. TITLE AND SUBTITLE
The Measurement of Cyber Resilience in Context

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
M. A. Thomas

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
DEVCOM Army Research Laboratory
ATTN: FCDD-RLA-ND
Aberdeen Proving Ground, MD 21005

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-10082

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES
ORCID: M. A. Thomas, 0000-0002-9141-4043

14. ABSTRACT
In the 2000s, research on the resilience of engineered systems gave rise to a research agenda on "cyber resilience," which is the ability of cyber systems to recover from stress. The development of resilient cyberspace systems was seen as a complement to traditional cybersecurity, which focused on hardening systems to prevent attacks. As an initial step, researchers sought to define and measure cyber resilience. This report considers both history and metrology to understand how this literature can mature. Materials scientists first measured the resilience of materials to compressive stress in the 19th century. The word "resilience" was then applied more widely by analogy and become the name of a broad, unmeasurable concept. While it may be possible to measure some system response to some stimulus for a given system, such a measurement does not necessarily capture other system responses to other stimuli and thus cannot capture the whole of the concept of resilience. Because there are many systems, possible stimuli, and possible responses, measurement alone may not contribute significantly to knowledge. As research on this topic matures, the value of a proposed measurement or measurement approach should be demonstrated by its service to a research question or decision-making model.

15. SUBJECT TERMS
cyber resilience; cyber resiliency; measurement; resilience; resiliency; Network, Cyber, and Computational Sciences

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	C. THIS PAGE		
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UU	32

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
M. A. Thomas	(410) 278-5573

STANDARD FORM 298 (REV. 5/2020)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

Acknowledgments

The author thanks Dr. Sidney Smith (ARL) for his comments on the report.

1. Introduction

"Cyber resilience" emerged as a specific focus in the late 2000s in the context of a larger discussion of how best to ensure the resilience of engineered systems and critical infrastructure. There is no agreed definition of cyber resilience; however, the National Institute of Standards and Technology (NIST) defines cyber resiliency as "the ability to anticipate, withstand, recover from, and adapt to adverse conditions, stresses, attacks, or compromises on systems that use or are enabled by cyber resources" (Ross et al. 2021). The focus on developing and choosing resilient cyberspace systems was seen as complementary to traditional cybersecurity, which focuses on system hardening.

To better develop and assess such systems, researchers sought to measure cyber resilience (Kott and Linkov 2021). As of this writing, a Google Scholar search shows more than 12,000 results addressing the measurement of cyber resilience. But it is not always clear what these measurements capture, how they compare, when they should be used, or what they contribute. The inconsistent use of the word "measurement" further muddies the waters.

This report places the literature on the measurement of cyber resilience in the wider context of history and metrology. Materials scientists first measured the resilience of materials to compressive stress in the 19th century. The word "resilience" was then applied more widely by analogy and become the name of a broad, unmeasurable concept. While it may be possible to measure some system response to some stimulus for a given system, such a measurement may not be generalizable to other system responses to other stimuli and thus cannot capture the whole of the concept of resilience. Because there are many systems, possible stimuli, and possible responses, measurement alone may not contribute significantly to knowledge. As research on this topic matures, the value of a proposed measurement or measurement approach should be demonstrated by its service to a research question or a decision-making model, and the circumstances under which it is appropriately used should be explored.

2. The Resilience of Physical Bodies: From One Measurement to Several

The measurement of resilience began in the early 19th century and quickly evolved from one measurement to several distinct measurements. The word "resilience" comes from the Latin "resilire," meaning "to jump back" ("Resilience" [date unknown]). The word was used to refer to a measurement of the ability of materials to withstand compressive force. Emerson (1768) did not use the word but is credited

with having published experiments on the ability of bodies to withstand compressive force in 1748. He posed as a problem the calculation of the weight that, falling from 185 ft, would break a bar of steel, given knowledge of the weight that would break it when hung from its middle.

Materials scientists began to explore resilience as a criterion for the choice of materials for specific applications. Young (1807, p. 50), drawing on existing scientific works, offered what he labeled as a "definition" of resilience: "The resilience of a beam may be considered as proportional to the height from which a given body must fall to break it." The following decade, Tredgold (1818) published a paper titled "On the Transverse Strength and Resilience of Timber," reporting the results of his experiments on six types of timber that could be used for shipbuilding. He defined "resilience" as "the power of resisting a body in motion" (Tredgold 1818, p. 216). He dropped different weights on beams of different types of timber, noting the weight that would break beams of each type.

Tredgold did not explicitly propose a measure of resilience, although implicitly the weight that caused breakage was the measure in his experiment. But there was a strong incentive to propose a measure. As the mechanical engineer Robert Thurston (1874, p. 20) wrote, "For many purposes, a metal having double the resilience of another is worth more than double-price." Accordingly, Thurston invented a machine whose purpose was to provide more reliable measures of resilience.

Thurston's focus was on the resilience of materials to torsional (rather than compressive) force. He defined "resilience" as "a measure of the capacity of the material to resist shock, and its value as equal to the amount of energy expended, or the 'work' performed in producing distortion or rupture" (Thurston 1879, p. 231). His patented Autographic Torsion Testing Machine produced a penciled curve, where the x axis reflected the magnitude of the stretch or distortion of the material and the y axis reflected the load or force applied to the material (Thurston 1874; Cornell University Library, Division of Rare and Manuscript Collections c2025). He then proposed to measure resilience as the area under the curve (AUC) drawn by his machine. This measure was put forward in 1878 by a U.S. Government board for testing metals, crediting Thurston (U.S. Board Appointed to Test Iron, Steel and Other Metals 1881, p. 305). AUC measures would go on to be one of the most proposed approaches to the measurement of resilience in engineering.

Researchers then began considering the resilience of materials to multiple types of stress. For example, Johnson (1890, p. 94) addressed the resilience of cast iron, proposing a definition and a measure:

> The resilience of a body is the work it does in resisting distortion. The elastic resilience is the work it may exert in resisting distortion, up to the elastic limit of the material, or up to the point where a part of the distortion becomes permanent. The total resilience of a body is the work it is capable of performing in resisting distortion up to the point of rupture. Resilience must therefore be measured in foot-pounds, or inch-pounds.

He distinguished resilience by type of stress, as either "tensile resilience," "transverse resilience," or "cross breaking resilience." Following Thurston, he plotted the extension of the body under stress for each type, producing what he called a "strain diagram," where the x axis represented the stretch per inch in length (or relative stretch) and the y axis represented stress in pounds per inch (Johnson 1890). Resilience was calculated as the area under the strain curve.

Similarly, Ferry (1903, p. 119) stated, "Corresponding to the different types of strain are different types of resilience, as tensile resilience, flexural resilience, torsional resilience, etc." Ferry (1903) gave an equation for resilience, *R*, in terms of load in dynes multiplied by the length of the rod in centimeters, which he divided by volume in cubic centimeters or the mass of the rod in grams. He distinguished different types of resilience by giving *R* a subscript.

3. Measuring and Comparing Resilience

In metrology (i.e., the science of measurement), measurement allows comparison of two quantities of the same kind (JCGM 2012, p. 4; ISO 2022, p. 2). A "unit of measurement" is a "real scalar individual quantity, defined and adopted by convention, with which any other quantity of the same kind can be compared by ratio, resulting in a number" (JCGM 2012, p. 4). The archetypical example of measurement is the measurement of a length *y* expressed in terms of the length of a rod *x*, the unit of measurement. A set of measurements is a weakly ordered set of scalars, measuring quantities of the same kind, using the same unit of measurement.

Campbell (1928) distinguished "fundamental measurement," the measurement of attributes that could be directly observed and measured, from "derived measurement," mathematical relationships discovered experimentally between fundamental measures or calculated based on fundamental measures. Similarly, the International Vocabulary of Metrology recognizes seven "base quantities"—length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity—as well as "derived quantities" calculated from base quantities (JCGM 2012, p. 3–4).

The measures proposed by early resilience researchers, such as "foot-pounds" or "stretch per inch," were derived measurements, calculated based on directly observable quantities. However, researchers did not consider them all to be of the same kind and therefore comparable. This prompted new taxonomies of resilience measures (Johnson's stress-type prefix to the word "resilience" and Ferry's subscripts). Tredgold had considered resilience as the response of various types of timber to the same compressive stress, but the word "resilience" now had a wider meaning. "Resilience" was no longer a measurement but a category of measurements. It was no longer a scalar but a vector. By the early 20th century, it was not possible to measure, compare, or rank the resilience of materials without further specifying the type of resilience under discussion.

4. The Resilience of Engineered Systems

Over the next century, the term "resilience" was applied beyond materials and across disciplines to both engineered and natural systems. It became the name of a construct and a concept. Some researchers struggled to find an all-inclusive, cross-disciplinary definition of "resilience" and a generalizable approach to its measurement, while others argued that this was impossible and proposed specific, purpose-scoped measurement approaches.

Schlink (1919) published an article titled "The Concept of Resilience with Respect to Indicating Instruments." Schlink sought to reduce the variance in measurement that came from the physical operation of measuring instruments such as scales and thermostats. He hypothesized that this variance was a function of the quantity of the thing to be measured and proposed plotting the quantity measured against the value reported by the instrument. Thurston (1879) had originally proposed this AUC approach to measure resilience to torsional stress. However, Schlink (1919, p. 168) argued that the performance of a measuring instrument, a mechanical system capable of storing and restoring energy, could also be measured as the AUC, a metric that was "exactly analogous to the resiliency in the case of other quasielastic bodies."

Although Schlink consciously analogized to the work on the measurement of resilience in materials science, the awareness of the use of an analogy would soon be lost as researchers went on to apply "resilience" to other systems. Systems engineers discussed the resilience of tires, power grids, computer systems, transport systems, infrastructure systems, subsystems, and systems of systems as the recovery of some aspect or functionality of the system after degradation or disruption (Healey 1924; Cottam et al. 2019). The term "resilience" was also used in the discussion of natural systems, including ecology (Holling 1973), biology, and astronomy, and in

behavioral systems in psychology (Lee Kum Sheung Center for Health and Happiness 2025), economics, politics, and international relations.

The problem space exploded. The set of systems is much larger and more diverse than the set of physical materials. A "system" is a subjective, contextual grouping of elements by function, purpose, or interaction to facilitate analysis and discussion. System boundaries segregate elements that are the subject of the speaker's analysis or discussion from those that are not, which are considered to be in the "environment." Accordingly, "systems can be either physical or conceptual, or a combination of both" (Sillitto et al. 2019, p. 3). They "exist in four domains: physical, information, cognitive, and social" (Bodeau et al. 2018a, p. 13). The International Council on Systems Engineering defines a system as "an arrangement of parts or elements that together exhibit behavior or meaning that the individual constituents do not" and an engineered system as "a system designed or adapted to interact with an anticipated operational environment to achieve one or more intended purposes while complying with applicable constraints" (Sillitto et al. 2019, p. 3).

A single system presents multiple possible states, capabilities, functionalities, or subsystems that could conceivably be degraded or disrupted by a stress. Moreover, these could be assessed according to different criteria, such as against an agreed standard, in terms of recovery to the pre-stress level, or the ability to use the system for a specific purpose under specified conditions.

Finally, there are many possible "stresses," now understood more broadly as any type of change that affects a system characteristic or functionality of interest in a way that a speaker considers to be negative. The effects of changes on the system may be proximate or distal, deterministic or probabilistic.

Broad, interdisciplinary adoption of the term "resilience" led to multiple, competing definitions of "resilience" used for different kinds of research, some definitions specific to disciplines or problems, some generic or very broad. One definition of "resilience" in psychology is "the maintenance or quick recovery of mental health during and after times of adversity" (Kalisch et al. 2021); in ecology, the "measure of the persistence of systems and of their ability to absorb change and disturbance" (Holling 1973). Cottam et al. (2019, p. 11) provides a systematic literature review of definitions of "resilience" in the engineering discipline and proposes, "An engineered resilient system is a system that is able to successfully complete its planned mission(s) in the face of disruption(s) (environmental or adversarial), and has capabilities allowing it to successfully complete future missions with evolving threats." This mission-focused definition implicitly ties the definition of "resilience" to the decision-maker's intended system use, making the definition subjective.

The literature ultimately proposed so many definitions of "resilience" that some questioned whether the word has any meaning at all. As Smith (2023, p. 379) reported, "A literature review searching for a consensus of the definition of resilience, in general, and cyber resilience, specifically, uncovered that the only consensus on the definition of resilience is that there is no consensus on the definition of resilience." Reflecting this evolution, Merriam-Webster offers two definitions of "resilience": "1. the capability of a strained body to recover its size and shape after deformation caused especially by compressive stress;" and "2. an ability to recover from or adjust easily to misfortune or change" ("Resilience" [date unknown]).

5. Measuring the Resilience of Engineered Systems

Research on system resilience had very different premises and purposes, with implications for measurement. One distinction was whether researchers conceived of "resilience" as a single property of a system under study or as a name for conceptually similar responses to change.

In the field of psychology, psychological resilience was seen as a "construct," a single attribute that is not directly observable but that has multiple, observable consequences (Cronbach and Meehl 1955; Windle et al. 2011). Measurements of constructs are derived mathematically from measurement of their observable consequences. Proposed measurements of constructs are hypotheses, validated by comparing the measure with the construct definition and considering the correlation of the measure with other observables as predicted by the theory of the construct. Measurements of constructs seek to capture the whole of the definition and nothing extraneous. If this can be done, then measurements of constructs are quantities of the same kind and therefore comparable.

By contrast, in engineering, there was no claim that systems, however defined, had a single, unobservable attribute of "resilience." Instead, "resilience" is the name of a concept: the idea of system responses to negative changes. Although specific system responses to specific stresses may be measurable, such measurements are not necessarily of quantities of the same kind and comparable. They may not even be measurable in the same units (e.g., system physical distortion in response to compressive stress vs. the latency of a computer network under heavy load). This is why Haimes (2009) argued that the resilience of an engineered system is a vector, meaning a set of values that reflect the resilience of different subsystems to different stresses. Haimes (2009, p. 500) argued that no common scale of resilience could be used "unless we pretend to assume that these different systems will be subjected to

the same exact threats and the same exact levels of such threats with the same exact probabilities."

Some engineering researchers proposed quantitative measurements of system resilience (Hosseini et al. 2016). The resilience of engineered systems is often depicted as a graph of performance of some functionality of the system over time (e.g., before a disruptive event or condition, from the beginning of the disruptive event or condition until the lowest level of system performance, during performance recovery, and after recovery). Following Thurston, a principal measurement approach continues to be the measurement of the area under this performance curve (Yodo and Wang 2016, p. 111408-3; Cottam et al. 2019, p. 25). This measurement substitutes the variable "time" for "force" or "load" used in the original strain curves and so does not distinguish response based on stress intensity. Another approach is the use of a ratio of predisruption and postrecovery performance (Yodo and Wang 2016, p. 111408-5).

These are better described as approaches to measurement, rather than measurements, because they cannot be applied without further specification. The use of these approaches requires selecting a type of system, some aspect of system performance, a quantitative measure of that performance, a stress, perhaps a measure of that stress, and a set of assumed environmental conditions. Unless the selections are identical, measurements produced following these measurement approaches would not necessarily be of the same kind and so would not be comparable. They are not measurements, and they do not and cannot capture the entirety of the concept of resilience.

Measurements of specific responses of specific systems to particular stresses may nevertheless be valuable if they contribute to answering a basic or applied research question; the research question is then the point of departure. Basic research seeks contributions to knowledge about the world in the form of generalizable findings about the subject of study and relationships among variables of interest. Findings are particularly prized if they are highly generalizable, counterintuitive, or novel. Descriptive research, including the observation of a measurement, can be an initial step in the search for such findings, allowing identification of patterns and formation of hypotheses.

Applied science seeks to solve practical problems. The value attached to solutions reflects the significance of the problem, and the generalizability, novelty, and practicality of the proffered solution. Research on the resilience of engineered systems is usually framed as a response to the practical problem of a decision-maker who must decide whether to acquire, develop, modify, or use a system, according to their specific objectives (mission or business process), preferences, and

constraints. The research question is how to develop, populate, and use a model as a decision aid. Some research on the resilience of engineered systems has proposed decision models, although decision science is a field in its own right (Specking et al. 2019; Azhar et al. 2021). The decision model determines what data is required; specific measurements of resilience are of interest to populate the model. The measurement may not be useful to other decision models for other decision-makers, decisions, systems, or stresses.

Rather than specific measurements, some researchers have proposed definitions of resilience or approaches to the measurement of resilience that are based on decision-maker criteria (Cottam et al. 2019; Specking et al. 2019). However, again, these are best thought of as approaches to measurement rather than as measurements because application of these approaches requires further problem specification and there is no assurance that two measures generated using the same approach would be "of the same kind" and therefore comparable. The approaches themselves may be appropriate in some circumstances and not others. For example, depending on their needs, a decision-maker may be interested in pre- or post-stress performance; the minimum performance of a functionality while under stress; the duration of degraded performance; or the time required for recovery of a functionality after stress. The units of measurement will depend on the specific functionality under assessment (e.g., speed, load supported, resolution). One size does not fit all, and the utility of the approaches can only be assessed in the context of an actual problem or class of problems. This history of the measurement of resilience provides a broader context for understanding the challenges of measuring cyber resilience.

6. Implications for the Measurement of Cyber Resilience

The discussion of "cyber resilience" arose in the 2000s as part of the broader discussion of the resilience of critical infrastructure (Tzavara and Vassiliadis 2024). In 1998, Benjamin et al. (1998) called for an improvement in "[information technology] resilience" in response to cybercrime. In 2005, the Homeland Security Advisory Council was directed to establish a Critical Infrastructure Task Force. The task force proposed making critical infrastructure resilience a top strategic priority, citing the dictionary definition of "resilience" as "an ability to recover from or adjust easily to misfortune or change" (HSAC 2006, p. 4). Critical infrastructure included both physical and cyber elements. "Cyber resilience" then became a topic of its own, seen as a complement to cybersecurity, which focuses on preventing system compromise (Figure 1). Researchers sought to assess and measure cyber resilience just as they had sought to assess and measure system resilience more broadly (Linkov et al. 2013).

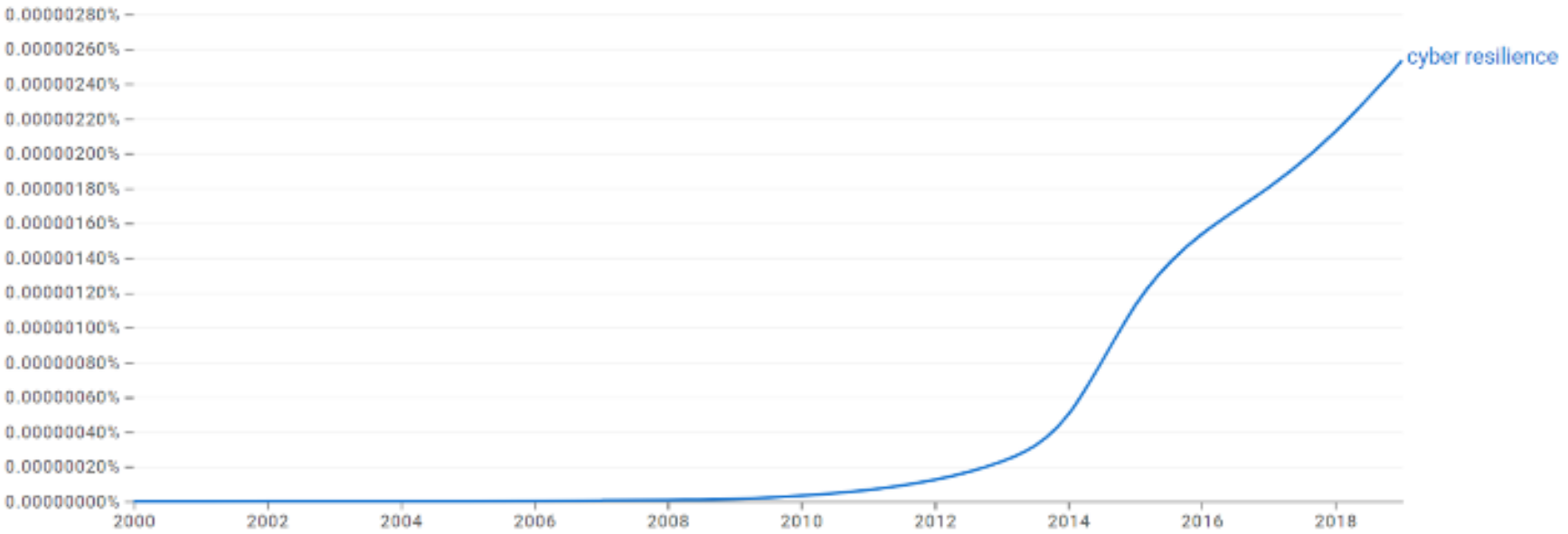

Figure 1. References to "cyber resilience" from 2000–2011 in the Google Books corpus using Google Books Ngram Viewer (Michel et al. 2010) show increasing attention to the term beginning in the late 2000s.

There is no agreed upon definition and scope of "cyber resilience." Just as there are many definitions of "resilience," there are multiple definitions of both "cyber" and "cyber resilience." Table 1 provides several definitions of these terms. In discussions of cyber resilience, the adjective "cyber" may be used to describe the system under stress or the stress itself. For example, NIST applies "cyber" in "cyber resiliency" to describe the system under stress (Ross et al. 2021). Similarly, Smith (2023, p. 5) defined cyber resilience as "the ability of a cyber system to recover from stress that causes a reduction of performance," noting that "the definition of cyber resilience must be restricted to cyber systems, but need not be restricted to cyber stress." Others have applied the adjective "cyber" to the stress. For example, AlHidaifi et al. (2024, p. 110446) define cyber resilience as "a system's ability to prepare, absorb, recover, and adapt its performance to pre-cyber-attack levels."

Table 1. Definitions of cyber and cyberspace

Source	Term	Definition
Merriam-Webster	Cyber	"Of, relating to, or involving computers or computer networks (such as the Internet)" ("Cyber" [date unknown])
Committee on National Security Systems	Cyberspace	"The interdependent network of information technology infrastructures that includes the Internet, telecommunications networks, computers, information systems, industrial control systems, networks, and embedded processors and controllers" (CNSS 2022)
DoD Dictionary of Military and Associated Terms	Cyberspace	"A global domain within the information environment consisting of the interdependent networks of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers" (DoD 2024)

In the more general study of the resilience of engineered systems, the system responses and stresses are those relevant to decision-maker requirements, making the choice subjective. Some studies define cyber resilience with respect to the general function of the system. For example, Cybenko (2016) and Kott et al. (2022) proposed measuring cyber resilience as a function of system key performance parameters (KPPs) and thresholds set for their performance as specified in DoD contracts. Other studies define cyber resilience in terms of a specific use and purpose for the system. Smith et al. (2022, p. 8) advanced a definition of cyber resilience as "the ability of systems to resist, absorb, and recover/adapt to a cyber compromise after the cyber compromise occurs, during execution of a mission," stating that "as currently defined, cyber resilience is a subset of mission resilience." NIST noted that "cyber resiliency is intended to enable mission or business objectives that depend on cyber resources to be achieved in a contested cyber environment" (Ross et al. 2021).

Bodeau et al. (2018b) emphasized the specificity of cyber resilience metrics and therefore proposed a methodology for selecting and incorporating relevant metrics from the catalog to meet decision-maker needs. The article stated that metrics are "strongly situated in an assumed context" and that "evaluation of a metric which seeks to represent a wide range of contexts may be infeasible, except when evaluation involves modeling and simulation (M&S), which perforce encodes assumptions about the system and its operational and threat environments" (Bodeau et al. 2018b, p. ix). It rejected the idea of producing a single metric for resilience, cybersecurity, or cyber resiliency, arguing that "any single metric will either obscure the complexity of the problem domain or require a large number of input measurements, which can vary so much in quality (e.g., timeliness, accuracy) that the resulting figure is highly uncertain" (Bodeau et al. 2018b, p. 11–12). It also cautioned that most metrics will not be comparable unless the same organization, mission, or business function is tracked over time.

Similarly, Cybenko (2016, p. 1) advanced "a concrete notion of cyber resiliency that can be tailored to meet specific needs of organizations that seek to introduce resiliency into their assessment of their cyber security posture." The paper defined "cyber resilience" as "an information processing system's ability to return to some level of desired performance after a degradation of that performance" (Cybenko 2016, p. 1). The paper acknowledged the subjectivity of the definition of resilience, noting that "concepts of 'performance' are specific to the missions of the enterprise" and so the aspect of system performance to be evaluated and the criteria for assessment of the acceptability of performance are to be "determined by the system operator based on mission requirements" (Cybenko 2016, p. 1, 3).

The literature on the measurement of cyber resilience is also shaped by the software community's inconsistent use of the terms "metrics" and "measures." A "software quality metric" was defined in 1993 by the Institute of Electrical and Electronics Engineers (IEEE) and the American National Standards Institute as "a function whose inputs are software data and whose output is a single numerical value that can be interpreted as the degree to which software possesses a given attribute that affects its quality" (IEEE 1993, p. 3). However, software metrics and their uses were criticized for lack of rigor (Abran et al. 2004; Abran et al. 2012). In 2017, international metrology standards were incorporated in ISO/IEC/IEEE 15939:2017, a standard adopted by the International Organization for Standardization, International Electotechnical Commission, and the IEEE (ISO/IEC/IEEE 2017). The standard defines a "measurement method" as a "logical sequence of operations, described generically, used in quantifying an attribute with respect to a specified scale," which includes both objective and subjective measures (ISO/IEC/IEEE 2017, p. 3).

Where a distinction is made, "measurements" are defined as a subcategory of "metrics," but "measurement" may be used to refer to objective data broadly, rather than given its metrological definition. For example, NIST's Software Quality Group explains, "We use measure for more concrete or objective attributes and metric for more abstract, higher-level, or somewhat subjective attributes . . . that are hard to define objectively. Measures . . . are bases for metrics" (NIST 2025). Bodeau et al. (2018b, p. 5) published a catalog of more than 500 cyber resiliency metrics dealing with systems, practices, and organizations and defined "cyber resiliency metrics" as

> the result of a process or method for measuring, evaluating, or comparing similar objects. Metrics can take a variety of forms (including quantitative, qualitative, semi-quantitative, and nominal); types (including measurements; evidence or observables; metrics computed or derived from measurements or evidence; and expert judgments); and relationships to intended effects (ranging from direct representations to indirect indications).

In this definition, measurements are a type of metric. Metrics tagged as "measured" in the catalog do not satisfy the metrological definition, but, consistent with NIST's definition, appear to include any type of objective, numeric data, such as percentages (e.g., of services that could be relocated to another machine, of users whose privileges could be modified dynamically), Booleans (e.g., whether data validation includes certain fields), and counts (e.g., number of dedicated enclaves defined).

In response to the calls for more rigorous, quantitative measurements of cyber resilience (Abran et al. 2012; Kott and Linkov 2021), researchers borrowed measurement approaches from the literature on the measurement of the resilience of engineered systems. They also faced the same challenges of generalizability and metrology.

Most research on the quantitative measurement of cyber resilience defines "cyber resilience" in terms of the generally expected functions of an engineered system or a decision-maker's specific use of that system for a business process or mission. Research then proposes measurement approaches, often illustrated with a specific use case. For example, researchers may compare the difference in system performance under stress over time or under different intensities of stress; they may use AUC approaches or a simple pre- and post-stress comparison. Such measurement approaches are sufficiently generic that they must be further specified in application, with the consequence that two measurements produced using the same approach may not be comparable. They therefore differ from "measurement

methods" as defined in the ISO/IEC/IEEE 15939:2017 standard (ISO/IEC/IEEE 2017).

Cybenko (2016) identified several possible measurement approaches to quantify and measure cyber resilience that could be relevant to a system operator with resilience requirements, such as the maximum periods of time that the system is allowed to operate below the threshold performance level or below the objective operational performance level (Cybenko 2016). Thresholds and objectives for operational performance are defined for each KPP of a system as part of the DoD acquisitions process, describing the generally expected functionality of a system. The intervals of time would be measured "with respect to some operator-defined class of failures and attacks that is known to the designer or vendor" (Cybenko 2016, p. 98250R-5). Application of these approaches requires further specification of the system KPPs, the operator requirements, and a group of failures and attacks, which means that measures developed following one of these approaches would not necessarily be comparable. Cybenko (2016) did not explore the circumstances under which these approaches would be useful, other than positing a system operator who needed them.

In some work, the subjectivity of the definition, research question, and contribution to knowledge are unclear. A measurement approach with or without a specific use case is offered without an explicit research question or without consideration of the circumstances under which the approach is useful or applicable. These gaps (and sometimes the title of the work) may give the impression that the claim is that the proposed approach is generalizable, that measurement approaches are equivalent to measurements,[*] and the proposed approaches or measurements capture the whole of the concept of cyber resilience.

For example, Hossain-McKenzie et al. (2018, p. 766), drawing on the language of Presidential Policy Directive 21, "Critical Infrastructure Security and Resilience," (White House 2013) defined "resilience" as the ability "to prepare for and adapt to changing conditions and withstand and recover rapidly from disruptions" and defined "cyber resilience" as resilience in the case of disruption from cyber threats. Hossain-McKenzie et al. (2018, p. 766) advanced performance-based metrics for the cyber resilience of industrial control systems, applying resilience metrics previously used for physical threats against infrastructure (Vugrin et al. 2011),

[*] Metrology recognizes "measurement methods," "generic description[s] of a logical organization of operations used in a measurement," typically descriptive, and "measurement procedures," "detailed description of a **measurement** according to one or more **measurement principles** and to a given m**easurement method**, based on a **measurement model** and including any calculation to obtain a **measurement result**" (JCGM 2012, p. 18; emphasis retained). The measurement approaches found in the literature are typically more detailed than a measurement method but do not satisfy the definition of a measurement procedure.

"which can be used to analyze and measure resilience of these systems to selected cyber threats." The proposed approach measured the difference between targeted and actual system performance over time from the onset to the conclusion of the disruption, as well as the effort involved in system recovery. System performance and recovery may have multiple factors. For each, the factors are weighted and summed.

Hossain-McKenzie et al. (2018) illustrated this measurement approach with a case study of a microgrid disrupted by a worm that propagates across the system to map it, and a decision-maker who seeks to evaluate the effectiveness of proposed security tools ("ADDSec") and the comparative benefits of their modes of operation. The work advanced several case-specific decision-maker questions, such as "Does deployment of ADDSec provide a measurable resilience benefit against the specified reconnaissance attack?" (Hossain-McKenzie et al. 2018, p. 768).

The application of the approach to the test case requires a detailed description of the system, its condition and operation (e.g., "ICMP requests are rate-limited by the devices in our experiment to 1 request per second as a standard protection against denial of service (DoS) attacks"), and of the cyberattack (Hossain-McKenzie et al. 2018, p. 770). The system impact is framed as "the fraction of network hosts infected by ScanWorm and the rate at which ScanWorm infects the hosts" (Hossain-McKenzie et al. 2018). Recovery effort factors include the "average round trip time for a packet to be acknowledged" and "fraction of sent packets that are dropped" (Hossain-McKenzie et al. 2018, p. 770).

The paper concluded the following:

> This paper introduces a set of metrics for quantifying resilience against cyber attacks. Application of the metrics in the presented case study demonstrates how the metrics can be used to evaluate the benefits of resilience technologies, such as [moving target defense], and how the technologies can be best deployed to maximize benefits. (Hossain-McKenzie et al. 2018, p. 773)

But the circumstances under which the proposed measurement approach should be used are not explored.

Jacobs et al. (2018, p. 39) stated that "cyber resilience efforts aim to ensure essential operations; maintain critical function levels; and rapidly recover" and sought to compare quantitative measurement of cyberattacks in terms of cybersecurity or cyber resilience. The authors used the Infrastructure Resilience Analysis Methodology (IRAM) to measure the cyber resilience of a control system for critical infrastructure. IRAM includes measures of Systemic Impact (the degradation of performance from the attack), the Total Recovery Effort (the cost

and effort for restoration of performance), and the Recovery Dependent Resilience index (a combination of the two). The Systemic Impact is a function of the AUC of systemic error. They used this approach to evaluate the impact on resiliency of different cyberattacks on a notional controller system as a way of prioritizing response (Jacobs et al. 2018, p. 45). They provided an example of a load frequency control under different cyberattacks (no attack, denial of service, signal jamming).

Again, the application of the approach to a specific case requires substantial specification of the system under test, its state, and cyberattacks as granular modeling choices. The paper acknowledges that the scenarios are ad hoc, "as categorizing cyber events and their impact on a control system is a difficult problem and such discussion is beyond the scope of this work" (Jacobs et al. 2018, p. 45). Regarding the comparison of measurements of cybersecurity versus cyber resilience, the paper concludes that "various measures and approaches have their places in a comprehensive assessment of a system yet each on their own fail to capture the entire picture" (Jacobs et al. 2018, p. 45). But the paper does not consider whether the same could be said of measures of resilience or the circumstances under which the proposed resilience measurement approaches are useful.

Kott et al. (2023, p. 1) emphasized the importance of measurement:

> A key challenge in the field of cyber resilience is quantifying or measuring resilience. Indeed, no engineering discipline achieved significant maturity without being able to measure the properties of phenomena relevant to the discipline.

The paper defined "cyber resilience" as "the ability of a system to resist and recover from a cyber compromise that degrades the business task-relevant performance of the system" (Kott et al. 2023, p. 1). It used the term "measure" to describe its approach, which is "quantitative and not qualitative, experimental, using physical quantities to the extent possible, business task focused, theoretically and empirically grounded" (Kott et al. 2023, p. 1). This is distinct from "metrics," which they defined as qualitative assessments of subject matter experts.

The report then described an approach to measurement, given a system under test, a set of representative business tasks, a set of cyberattacks, and system performance functions. (Kott et al. 2023). The system performance functions are aggregated into a single number by, for example, creating a weighted sum. Data to calculate the single measure is gathered for baseline (no cyberattacks) and then random cyberattacks. The relative performance of the system is calculated by dividing the number for system performance under attack by the number for baseline performance at each point of time of interest.

Resilience is calculated as the area of the difference between baseline performance and performance under attack for the time period of interest (for example, the period of attack). The effectiveness of malware and "bonware" (defined as "everything that resists malware" [Kott et al. 2023, p. 8]) is calculated as the difference in functionality compared with baseline over a time period.

The definition of "measurement" does not correspond to the definition in metrology, and the use of the methodology requires further specification such that measures produced by it are not necessarily of the same kind and comparable. The report does not consider the circumstances under which the proposed measurement approach would be useful and, by advancing it as an answer to the challenge of measuring system resilience, implies that it is generally applicable. However, decision-maker needs can vary. Decision-makers may require minimum performance for multiple system functionalities simultaneously, which cannot be captured by a weighted sum; or they may be interested in system response to different intensities of stress rather than stress over time.

Other work on cyber resilience has proposed various measurement approaches or illustrated the measurement of some system response to some stress for a specific use case. For example, Cybenko (2016) and Kott et al. (2022) proposed measuring resilience as a weighted sum of KPPs and the thresholds for their performance. Smith et al. (2022) demonstrated the AUC approach to measurement of cyber resilience by evaluating the area under the curve of vehicle speed over time, specifying an attack target, a type of cyberattack, a mission, and environmental conditions.

This work illustrates the use of empirically grounded quantitative assessments of aspects of cyber resilience as opposed to subjective assessments. However, none of the approaches can capture the whole of the concept of cyber resilience, leaving open the question of when each approach should be applied.

7. Conclusion

Since the first research on the measurement of resilience, the definition of "resilience" and the context have changed. The measurement of resilience was initially the measurement of an inherent property of a material, its response to a single type of physical stress under given experimental conditions. However, the word "resilience" quickly took on other meanings. It was extended to refer to responses to different stresses in materials science. It was then applied by analogy to natural and engineered systems. It has become the name of a concept: the responses of systems to some (subjectively negative) stimulus. This, in turn, changed the meaning and utility of measurement in what has become a vast and

diverse problem space. While it is still possible to measure to make some observation about a system behavior under given circumstances, the size of the problem space means that unmotivated, nongeneralizable measurements are not likely to be significant contributions to knowledge.

The literature on the measurement of cyber resilience continues to mature. As in the broader literature on the resilience of engineered systems, researchers have borrowed and proposed general approaches to measurement and demonstrated that it is possible to measure some cyber system response to a specific stress over time. However, the research question and the applicability and limitations of the proposed measurement approach are often unclear. In some cases, this can give the erroneous impression that the proposed approaches are universally generalizable. Finally, the word "measurement" is not always given its metrological definition, but is instead used to describe objective, numerical data or general approaches that, when applied, do not necessarily yield comparable results. Future work should adopt definitions of "measurement" that are consistent with metrological standards and make the contribution to knowledge clear by situating a measurement approach in the context of a research question and describing the applicability and limitations of the approach.

8. References

Abran A, Desharnais J, Cuadrado-Gallego JJ. Measurement and quantification are not the same: ISO 15939 and ISO 9126. J Software Evolu Process. 2012;24(5):585–601. https://doi.org/10.1002/smr.496

Abran A, Sellami A, Suryn W. Metrology, measurement and metrics in software engineering. Proceedings of the Fifth International Workshop on Enterprise Networking and Computing in Healthcare Industry (IEEE Cat. No.03EX717); 2004 Sept 5; Sydney, Australia. IEEE; c2004. p. 2–11. https://ieeexplore.ieee.org/document/1232451

AlHidaifi SM, Asghar MR, Ansari IS. Towards a cyber resilience quantification framework (CRQF) for IT infrastructure. Comput Netw. 2024;247:110446. https://doi.org/10.1016/j.comnet.2024.110446

Azhar NA, Radzi NA, Ahmad WSHMW. Multi-criteria decision making: a systematic review. Recent Adv Electr Electron Eng. 2021;14(8):779–801. https://doi.org/10.2174/2352096514666211029112443

Benjamin R, Gladman B, Randell B. Protecting IT systems from cyber crime. Comput J. 1998;41(7):429–443. https://doi.org/10.1093/comjnl/41.7.429

Bodeau DJ, Graubart RD, McQuaid RM, Woodill J. Cyber resiliency metrics, measures of effectiveness, and scoring: enabling systems engineers and program managers to select the most useful assessment methods. Mitre; 2018a. https://www.mitre.org/sites/default/files/2021-11/prs-18-2579-cyber-resiliency-metrics-measures-of-effectiveness-and-scoring.pdf

Bodeau DJ, Graubart RD, McQuaid RM, Woodill J. Cyber resiliency metrics catalog. Mitre; 2018b. https://www.mitre.org/news-insights/publication/cyber-resiliency-metrics-catalog

Campbell NR. An account of the principles of measurement and calculation. Longmans, Green and Company, Limited; 1928.

[CNSS] Committee on National Security Systems. Committee on National Security Systems glossary. Committee on National Security Systems; 2022. CNSSI No. 4009.

Cornell University Library, Division of Rare and Manuscript Collections. Robert H. Thurston's autographic torison testing machine. Cornell University Library; c2025. https://digital.library.cornell.edu/catalog/ss:549797

Cottam BJ et al. Defining resilience for engineered systems. Eng Manage Res. 2019;8(2):11–29. https://doi.org/10.5539/emr.v8n2p11

Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychol Bull. 1955;52(4):281–302. https://doi.org/10.1037/h0040957

Cybenko G. Quantifying and measuring cyber resiliency. Proceedings Volume 9825, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security, Defense, and Law Enforcement Applications XV; 98250R; 2016 May 12. https://doi.org/10.1117/12.2230586

Cyber. Merriam-Webster; c2024 [accessed 2024 Apr 26]. https://www.merriam-webster.com/dictionary/cyber

[DoD] Department of Defense. DoD dictionary of military and associated terms. Department of Defense (US); 2024.

Emerson W. The doctrine of fluxions: not only explaining the elements thereof, but also its application and use in the several parts of mathematics and natural philosophy. J. Bettenham; 1768. https://archive.org/details/TO0E037877_TO0324_PNI-1116_000000

Ferry ES. A course of practice physics for students of science and engineering. LaFayette; 1903. https://www.google.com/books/edition/A_Course_of_Practical_Physics_for_Studen/xZgAAAAAMAAJ?hl=en&gbpv=1

Haimes YY. On the definition of resilience in systems. Risk Anal. 2009;29(4):498–501. https://doi.org/10.1111/j.1539-6924.2009.01216.x

Healey A. The tyre as part of the suspension system. Proc Inst Auto Eng. 1924;19(1):26–128. https://doi.org/10.1243/PIAE_PROC_1924_019_013_02

Holling CS. Resilience and stability of ecological systems. Annu Rev Ecol. 1973;4(1):1–23. https://www.jstor.org/stable/2096802

Hossain-McKenzie S, Lai C, Chavez A, Vugrin E. Performance-based cyber resilience metrics: an applied demonstration toward moving target defense. IECON 2018 44th Annual Conference of the IEEE Industrial Electronics Society; 2018 Oct 21–23; Washington, DC. IEEE; c2018. p. 766–773. https://ieeexplore.ieee.org/abstract/document/8591764/

Hosseini S, Barker K, Ramirez-Marquez JE. A review of definitions and measures of system resilience. Reliabil Eng Syst Saf. 2016;145:47–61. https://doi.org/10.1016/j.ress.2015.08.006

[HSAC] Homeland Security Advisory Council. Report of the Critical Infrastructure Task Force. Department of Homeland Security (US); 2006 Jan. https://www.dhs.gov/xlibrary/assets/HSAC_CITF_Report_v2.pdf

[IEEE] IEEE 1061-1992. IEEE standard for a software quality metrics methodology. Institute of Electrical and Electronics Engineers; 1993. https://standards.ieee.org/ieee/1061/1548/

[ISO] ISO 80000-1:2022. Quantities and units — part 1: general. International Organization for Standardization; 2022. p. 2. https://www.iso.org/standard/76921.html

[ISO/IEC/IEEE] ISO/IEC/IEEE 15939:2017. Systems and software engineering — measurement process. International Organization for Standardization/International Electrotechnical Commission/Institute of Electrical and Electronics Engineers; 2017 [accessed 2024 Dec 10]. https://www.iso.org/standard/71197.html

Jacobs N, Hossain-McKenzie S, Vugrin E. Measurement and analysis of cyber resilience for control systems: an illustrative example. 2018 Resilience Week (RWS); 2018 Aug 20–23; Denver, CO. IEEE; c2018. p. 38–46. https://doi.org/10.1109/RWEEK.2018.8473549

[JCGM] Joint Committee for Guides in Metrology. JCGM 200:2012: international vocabulary of metrology – basic and general concepts and associated terms (VIM). 3rd ed. Joint Committee for Guides in Metrology; 2012. https://www.bipm.org/documents/20126/2071204/JCGM_200_2012.pdf

Johnson JB. Cast-iron—strength, resilience, tests and specifications. Trans Am Soc Civ Eng. 1890;22(1):91–120.

Kalisch R et al. The frequent stressor and mental health monitoring-paradigm: a proposal for the operationalization and measurement of resilience and the identification of resilience processes in longitudinal observational studies. Front Psychol. 2021;12:710493. https://doi.org/10.3389/fpsyg.2021.710493

Kott A et al. A methodology for quantitative measurement of cyber resilience (QMOCR). DEVCOM Army Research Laboratory (US); 2023 Sep. Report No.: ARL-TR-9672.

Kott A, Linkov I. To improve cyber resilience, measure it. Computer. 2021;54(2):80–85. https://doi.org/10.1109/MC.2020.3038411

Kott A, Weisman MJ, Vandekerckhove J. Mathematical modeling of cyber resilience. MILCOM 2022 - 2022 IEEE Military Communications Conference (MILCOM); 2022 Nov 28–Dec 2; Rockville, MD. IEEE; c2022. p. 849–854. https://doi.org/10.1109/MILCOM55135.2022.10017731

Lee Kum Sheung Center for Health and Happiness. Resilience measurement. Harvard T. H. Chan School of Public Health; c2025. https://hsph.harvard.edu/research/health-happiness/resilience-measurement/

Linkov I et al. Resilience metrics for cyber systems. Environ Syst Decis. 2013;33:471–476. https://doi.org/10.1007/s10669-013-9485-y

Michel J-B et al. Quantitative analysis of culture using millions of digitized books. Science. 2010;331(6014):176–182. https://doi.org/10.1126/science.1199644

[NIST] National Institute for Standards and Technology Information Technology Laboratory/Software and Systems Division, Software Quality Group. Metrics and measures. National Institute for Standards and Technology; 2025 Feb [accessed 2025 Mar 31]. https://www.nist.gov/itl/ssd/software-quality-group/metrics-and-measures

Resilience. Merriam-Webster; c2024 [accessed 2024 July 17]. https://www.merriam-webster.com/dictionary/resilience

Ross R, Pillitteri V, Graubart R, Bodeau D, McQuaid R. Developing cyber-resilient systems: a systems security engineering approach. NIST SP 800-160 vol. 2 rev. 1. National Institute for Standards and Technology; 2021 Dec. https://csrc.nist.gov/pubs/sp/800/160/v2/r1/final

Schlink FJ. The concept of resilience with respect to indicating instruments. J Franklin Inst. 1919;187(2):147–149. https://doi.org/10.1016/S0016-0032(19)90430-3

Sillitto H et al. Systems engineering and systems definitions: version 1.0. International Council on Systems Engineering; 2019 Jan 8. https://www.incose.org/docs/default-source/default-document-library/incose-se-definitions-tp-2020-002-06.pdf

Smith SC et al. Quantitative measurement of cyber resilience: a tabletop exercise. DEVCOM Army Research Laboratory (US); 2022. Report No.: ARL-TR-9380. https://apps.dtic.mil/sti/citations/trecms/AD1158532

Smith SC. Toward a scientific definition of cyber resilience. Proc of the 18th International Conference on Cyber Warfare and Security. 2023;18(1):379–86. https://doi.org/10.34190/iccws.18.1.960

Specking E et al. Assessing engineering resilience for systems with multiple performance measures. Risk Anal. 2019;39(9):1899–1912. https://doi.org/10.1111/risa.13395

Thurston RH. On the strength of American timber. J Frankl Inst. 1879;108(4):217–235. https://www.google.com/books/edition/On_the_Strength_of_American_Timber/NWgDAAAAYAAJ?hl=en&gbpv=0

Thurston RH. On the strength, elasticity, ductility and resilience of materials of machine construction, and on various hitherto unobserved phenomena. Merrihew & Son; 1874. https://openlibrary.org/books/OL24983009M/On_the_strength_elasticity_ductility_and_resilience_of_materials_of_machine_construction_and_on_vari

Tredgold T. 1818. On the transverse strength and resilience of timber. The Philosophical Magazine and Journal. 51(37):214–216.

Tzavara V, Vassiliadis S. Tracing the evolution of cyber resilience: a historical and conceptual review. Int J Inf Secur. 2024;23:1695–1719. https://doi.org/10.1007/s10207-023-00811-x

US Board Appointed to Test Iron, Steel and Other Metals. Report of the United States board appointed to test iron, steel and other metals in two volumes. Government Printing Office (US); 1881. https://www.google.com/books/edition/Report_of_the_United_States_Board_Appoin/DvJYAAAAYAAJ?hl=en&gbpv=1&dq=Report+of+the+United+States+Board+Appointed+to+Test+Iron,+Steel+and+Other+Metals&printsec=frontcover

Vugrin ED, Warren DE, Ehlen MA. A resilience assessment framework for infrastructure and economic systems: quantitative and qualitative resilience analysis of petrochemical supply chains to a hurricane. Process Saf Prog. 2011;30(3):280–290. https://doi.org/10.1002/prs.10437

White House. Presidential Policy Directive/PPD-21: critical infrastructure security and resilience. White House Office of the Press Secretary; 2013 Feb 12 [accessed 2024 Dec 9]. https://obamawhitehouse.archives.gov/the-press-office/2013/02/12/presidential-policy-directive-critical-infrastructure-security-and-resil

Windle G, Bennett KM, Noyes J. A methodological review of resilience measurement scales. Health Qual Life Outcomes. 2011;9:8. https://doi.org/10.1186/1477-7525-9-8

Yodo N, Wang P. 2016. Engineering resilience quantification and system design implications: a literature survey. J Mech Des. 138(11):111408–111421. https://asmedigitalcollection.asme.org/mechanicaldesign/article-abstract/138/11/111408/384079

Young T. A course of lectures on natural philosophy and the mechanical arts. J Johnson; 1807. http://archive.org/details/lecturescourseof02younrich

List of Symbols, Abbreviations, and Acronyms

AUC	area under the curve
DoD	Department of Defense
DoS	denial of service
ICMP	Internet Control Message Protocol
IEEE	Institute of Electrical and Electronics Engineers
IRAM	Infrastructure Resilience Analysis Methodology
KPP	key performance parameter
NIST	National Institute of Standards and Technology

1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	DEVCOM ARL FCDD RLB CI TECH LIB