Moving Beyond "Safety": Revising the Concept to Upgrade the Art and Science of Safety Management

By

Paul R. Schulman

Research Associate, Center for Catastrophic Risk Management, University of California, Berkeley

Abstract

The concept of "safety" is grossly underspecified. It has too often been ambiguously defined and understood as a noun describing only vaguely a technical, organizational or societal state. This has given rise to confusions and contradictions that have actually undermined progress in safety management.

But safety management can be significantly improved by revising our concept of safety to be a verb describing an ongoing multi-dimensional *process* and not a single technical or organizational end state.

As a process, safety can be understood, managed, assessed and measured by decomposing it into the following ongoing activities:

- hazard analysis, assessment and management,
- vulnerability analysis and mitigation,
- error definition, identification and management
- reliability defined as a process of continually tending to the integrity and effectiveness of all of the above, and less as the invariance of system outputs.

Clarity in defining, understanding, and measuring these processes will be a necessary foundation for to pursue "safety" in the midst of increasingly complex, rapidly changing and uncertain problem environments.

1. Introduction

Everywhere there is concern about improving "safety": system safety, worker safety, public safety, and societal safety. A variety of regulatory agencies regulate for safety, with special departments and with professional safety analysts. Many companies have safety goals, safety officers and safety slogans such as "Safety is Job One."

It's been easy to objectify safety as a noun -- (as in "We are safe.") -- something (a condition or state) an organization either does or doesn't possess. Sometimes safety can be used in its adjectival form as in or "we strive to err on the safe side." Of course, this usage doesn't specify

safe from what?

An argument to be pursued here is that ambiguity in the concept of "safety" has, ironically, probably undermined the clarity and effectiveness of many efforts to achieve it.

But from its conceptual foundation we don't have a clear enough concept of what is actually possessed to treat "safety" as a measurable, scalable or comparative variable. We are left with few agreed-upon or reliable ways to determine "how safe" a technical system, organization or society might actually be in real time.

2. The Definition of "Safety"

How do we typically define "safety"? The Oxford dictionary definition is: "the condition of being protected from or unlikely to cause danger, risk, or injury."

But isn't "protected from" quite different fromn "unlikely to cause"? "Danger, risk or injury" are themselves unspecified. Does "safety" mean an identical condition of protection or unlikelihood applying across all dangers, risks or injuries?

Another safety definition, offered by engineer Nancy Leveson, is: "Safety can be defined as freedom from unacceptable losses (accidents)." (Leveson, [18])

But the term "unacceptable losses", is ambiguous in two ways: Are these losses addressed only to physical harm? In many corporate organizations "losses" include major financial and equipment losses. Also, whose "unacceptability" is at issue: managers, shareholders or the public?

In the face of these ambiguities, how is "safety" to be determined and measured? It is not surprising that many managers and regulators operate with proxies for "safety".

An inspector in one regulatory agency stated simply that if "utilities are following the rules and regulations, they're safe." In this formulation Compliance = Safety. But are the rules and regulations themselves up to date, complete and actually effective? There is a difference between regulating for reliability and safety and *reliable regulation*. (Macrae, [21]; Roe and Schulman [32]).

Goven these uncertainties, many organizations and their regulators simply define and assess safety as "intervals without accidents" a condition that can only be retroactively determined. Assessing these and other definitional ambiguities, James Reason has concluded that:

"Safety is defined and measured more by its absence than its presence." (Reason, ([27]).

What Reason is saying here, in effect, is that there is much more clarity in understanding and identifying the *causes* of accidents than there is in understanding the *prevention* of them.

Take the idea of safety as “intervals without accidents”. In this perspective, measuring safety for both organizations and their regulators must always be a *lagging* and retrospective indicator. An organization is safe until it isn’t.

An extreme example of this safety concept is a statement by lawyers for the Pacific Gas and Electric Company in a court filing against a charge by the U.S. Justice Department of criminal negligence in its gas pipeline management, a condition leading to the 2010 explosion of a pipeline in a San Bruno, CA neighborhood that killed 8 people:

"The San Bruno pipeline explosion was a terrible accident which devastated many people and harmed an entire community. A pipe with a faulty weld was placed in service in 1956, *where it performed safely* [italics added] for 54 years. Suddenly, it failed catastrophically." (S.F. Chronicle, 9/4/14).

Is this really a sensible concept of “safety”?

This retrospective focus has meant that many managers and regulators have really had few ways to distinguish reliably and proactively a “safe organization” from one with little or no safety management commitment, but that possibly through luck has so far been “failing to fail.” But the problem with failing to fail is that the organization may eventually “fail to fail to fail” and actually “succeed” in having an accident!

The question must be asked: given these conceptual ambiguities how likely is it to deduce clear and consistent propositions pertaining to the management and real-time assessment of safety? A case can be made that the conceptual ambiguity and weakness of the “safety” concept has undermined efforts at both performing and assessing safety management.

The concept of safety has been vastly underspecified in two dimensions:

-- Safety with respect to *what* ? (specific hazards and harms?) and

-- Safety with respect to *how ? (*how does the “safety” actually work?*)*

3. Alternate Definitions of Safety.

Fortunately, some alternative perspectives have emerged that can both enlarge the concept and clarify the definition of safety.

With respect to *what* we have differentiations between Safety I, II and III -- individual and system safety (Hollnagel, [11]; Leveson, [19]); and organizational vs societal safety (Olsen, et.al. [25]).

With respect to the *how* of safety an international group of aviation regulators have defined safety not as a condition or state but as a *process*, and a continuing one:

"Safety is more than the absence of risk; it requires specific systemic enablers of safety to be maintained at all times to cope with the known risks, [and] to be well prepared to cope with those risks that are not yet known and to address the natural 'erosion' of risk controls over time." (Safety Management International Collaboration Group, [34]).

Echoing this perspective, psychologist Karl Weick has called safety "the continuous production of dynamic non-events." (Weick, [43]). But these non-events, are not the *absence* of something, they consist of a variety of activities and actions across many dimensions.

"Process Safety" has in fact become an important feature of the Chemical Processing Industry after a series of major accidents beginning with the Bhopal Toxic Chemical Release in 1984, the more recent BP Texas City Refinery Explosion in 2005 and other chemical plant explosions in 2010 and 2013. These accidents led eventually to the evolution by the U.S. Occupational Safety and Health Administration of a regulatory Process Safety Management Standard for Highly Hazardous Chemicals beginning in 1992 (OSHA, [40]). The Chemical Industry itself has established a Center for Chemical Process Safety that promotes process safety standards among its members in the "chemical, oil, petroleum, energy, food, pharmaceutical, fine chemicals. mining, minerals, metals, pipeline, specialty products, recycling-batteries, software and AI industries (CCPS, [5])."

From these perspectives safety is not a static structure and a fixed capacity but a process that must be ongoing, and that means acting and operating *safely* across a variety of dimensions and activities. As one organization's safety manager phrased it: "Our regulator wants to certify safety as something that we have, but it's really something that we do."

"Safety" as a process should actually be treated as a verb and not a noun. Then a safety statement could be something like: "We safety our way around specific hazards as well as through risks." This opens up the focus on "how" this process is to be specified.

4. Key Process Activities and Indicators.

In fact, a number of Key Performance Indicators (KPIs) are available as a way to assess safety processes in organizations, depending on their presence and degree of maturity. Many of these can be prospectively specified, identified and even measured. These processes are:

- hazard analysis, assessment and management,
- vulnerability analysis and mitigation,
- error definition, identification and management
- reliability defined as a process of continually tending to the integrity and effectiveness of all of the above, and less as the invariance of a system state.

Lapses or drifts away from these can also be identified before an accident or failure.

These processes have been observed in research on high reliability organizations (HROs): nuclear aircraft carriers, air traffic control centers and nuclear power plants. (LaPorte and Consolini, [15]; LaPorte[16]; Rochlin, et.al. [31] and Roberts, [29]; Schulman, [36]).

4.1 Ongoing hazard identification, assessment and management processes

In HROs safety management is organized around worst-case hazards they wish to preclude as "never" events, not simply probabilistically but deterministically, e.g.:

- a reactor meltdown and radiation release into the environment;
- a collision of commercial aircraft under air traffic control;
- a loss of aircraft carrier launch and retrieval capacity under wartime or threat conditions.

The first two of these events were accidents that could harm or kill people in large numbers outside of the organization. They were regarded by the public with dread (Weart [42].

This dread proved to be an important foundation for shielding HROs from wider market and efficiency pressures (LaPorte and Consolini [15]). Also, it offered political and regulatory support for pilots, ATC and nuclear control operators to refuse to operate under conditions they considered to be too dangerous or uncertain. It also supported highly competent and proactive regulatory safety oversight (Nævestad, [23]).

4.2 Hazards Assessment vs Risk Assessment

Hazards analysis has some important differences from typical risk assessment. The mathematical concept of risk links a specifically estimated consequence (loss or harm) to a specific probability that it will occur. The loss estimate is discounted by its multiplication with the probability estimate. But often there may be significant uncertainty surrounding both of these estimations.

Risk assessments often treat a risk as a single value of loss rather than a range of possible values reflecting these uncertainties, assessing a given risk as an "expected value". Often risk may be the measured harm of a set of possible foreseen outcomes (in lives lost or costs associated with damage and recovery) with each outcome *multiplied* by its foreseen probability. The results are then summed. An overall risk is then assessed as the average of the separate expected values (really losses).

But such an averaging process often averages out the extreme worst-case consequences potentially at issue. One analyst has termed this underestimation "The Flaw of Averages" (Savage, [35]).

Risk analyses typically don't account for or weigh in public risk tolerances based on public dread of the hazard. Nor do the mathematics of probability apply well to the character of dreaded events. A one-hundred-year storm, for example, could occur in consecutive years and

would not violate the validity of its probability calculation, which mathematically asserts only a likely distribution of the event over a very large if not infinite number of years. This "frequentist" approach also does not cover a uniquely consequential event with hazards previously unknown (Aven [2]).

It has also been argued that assessing uncertainties in both consequences and likelihoods may require analysis of variables beyond those in typical risk assessments to include organizational variables and societal vulnerabilities in a sociotechnical perspective (Aven and Ylonen [3]).

A recent article has argued that while there is "risk science" of the frequentist sort still being researched and practiced, there is increasingly research *on* risk science findings that occur across disciplines at the boundaries of risk – research in fields such as psychology, anthropology, sociology, and organization theory. In the intersection of these fields, risk science research can produce some new and larger insights concerning the process of risk assessment, perception, communication and management (Antonsen, Aven and Haavik [1]).

In the meantime, the seeming precision of risk calculations and risk assessments often gives the illusion and even self-delusion to organizational managers that they understand more than they do about hazards they confront, and that they are appropriately managing for safety because of their accounting for these as risks. But precision and accuracy are two quite different properties.

For example, recent efforts have been made to compute "risk/spend efficiencies" – a kind of risk/benefit analysis used to compare the worth in safety promotion of different safety investments by an organization. This type of analysis has even been required by some state regulatory agencies in the United States that have the authority to regulate rates that can be charged by private utilities for the provision of what may be monopoly services within a state. Utilities seek to get rate increases to offset the risk/spend safety investments they make.

One such "risk/spend efficiency" analysis based on probability assessments was taken to the supposed precision of a hundredth of a Serious Injury or Fatality (SIF) saved per year per $ spent. Among the assumptions here is that safety is fungible and can be decomposed down to that level of precision. (Apparently safety can be bought by the fraction of a SIF just like we can buy ground beef by the pound!)

This is a framework of false precision. It invites errors of misestimation and misunderstanding – not a good foundation for the pursuit of effective safety management.

In this respect it has also been argued relative to such risk assessments that quantitative evaluations are often based on other "unrealistic assumptions that are often unstated, such as that accidents are caused only by errors … testing is perfect … and a system is designed, constructed, operated, maintained, and managed according to good engineering standards." (Leveson[19]).

Risk analyses may also overlook the interactivity of one risk with another in consequences, likelihoods or both (Lindhout, Kingston-Howlett, Hansen and Reniers, [20]). They then become static elements in safety management, ignoring drift into conditions of shifted consequences or likelihoods over time.

But hazard analysis focuses on safety beyond the conceptual confines of risk. Hazard analysis must be sensitive to the shifting nature of hazard with shifting environmental conditions, new technological upgrades, enlarging technical and social interdependencies, as well as organizational complacency. That is why hazard analysis should be closely linked to another important safety process, that of vulnerability analysis.

4.3 Ongoing vulnerability analysis and mitigation.

Vulnerability is about more than the physical ability of an external event to harm. It involves the susceptibility to *be* harmed. Vulnerability is a condition in which a person, organization or a society itself can make or leave itself open to a specific harm. This can be quite a dynamic process.

In relation to both wildfire risk, and flood risk, for example, people can be agents in their own vulnerability, literally putting themselves in harm's way.

"In a 2022 study fire scientists and eco-hydrologists mapped out where vegetation is creating the highest fire risks across the western U.S., then compared that map to where in the region people have been moving into the wildland-urban interface, and were surprised to discover that the fastest rate of population growth by far has been in the areas with the highest fire risk. This includes several areas in California, Oregon, Washington and Texas" (Radeloff, [26]).

Also, nationally, the number of Americans living in 100-year floodplains in the year 2000 increased about 14% beyond population estimates for the same geographic areas that were made in 2016 (Macaig, [20]).

4.3.1 Vulnerability Imports and Exports

At the same time, on a societal level, the internet has become an important infrastructure with simultaneously great social dependency and perhaps the highest vulnerability to attack of any technical system humans have ever created. Every new element added to its coverage or capacity introduces additional vulnerabilities to cyber-attack across multiple dimensions. (Iftikhar, [12]; Bambauer, [4]).

Also, a "common theme" of vulnerability that can be applied to critical "lifeline" infrastructures such as electricity, telecommunications, water supply, and many forms of transport in relation to flood, wildfire and other natural events, is *the growing vulnerability of critical infrastructures to the failure of other infrastructures* (Kasmalkar, et.al. [14]). This vulnerability in particular pertains to new interdependencies introduced into the requirements for critical infrastructures to quickly and effectively engage in the restoration of service.

There are obvious interconnections between infrastructures under normal operations but in natural disasters latent interconnectivities can be revealed (Anonymous [29]). A water purification plant may become dependent upon a highway department for key road clearances that are necessary for its personnel to get to the plant. A power grid may be dependent upon the restoration of tele-communications before it can identify line breakages as well as control generators and switches through its Supervisory Control and Data Acquisition (SCADA) tele-communications system. Infrastructures each can become suddenly dependent on another for their own restoration of service.

These infrastructures can also export vulnerability to a society at large when it's populations become closely dependent on infrastructure services such as electricity, water, transportation or telecommunications for which there may be few substitutions. They are called "lifeline" infrastructures because people can die when these infrastructure services disappear for long periods.

So one challenge for safety management is to anticipate and mitigate, if possible, this vulnerability. It becomes a critical safety *management* role to plan for and achieve, an *infrastructure of inter-infrastructure resilience*, including ongoing personnel contacts between infrastructure personnel, a robust inter-infrastructure communication system with compatible data sharing software as well as the existence of back-up storage centers for spare parts and an ability to share resources for emergency transport and generation. It thus becomes a safety management responsibility across infrastructures to facilitate, in the face of actual inter-infrastructure disruptions, *infrastructure service restoration as rapidly as possible* (Anonymous [33]).

4.5 Safety as the Ongoing Management of Error.

The accident literature is filled with examples of cognitive errors of misperception, mis-specification, misestimation and basic misunderstanding as causes and compounding factors in fatal accidents. Trying to detect and prevent these errors are important safety processes. In high reliability organizations (HROs) priority was given to the management of these kinds of errors (LaPorte [15]; Roberts, [29]; Rochlin, [31]; Anonymous, [36]).

HROs identified a "precursor zone" of those conditions, states or practices, both technical and organizational, that increase the likelihood of cognitive errors that could ultimately lead to precluded events, those that simply must not happen and that need to be precluded deterministically, not probabilistically.

These precursor zones could be operating in unexperienced or unstudied conditions outside of prior analysis, performing a maintenance task without prior alerts and authorizations, excessive cognitive load or time pressure in the performance of important tasks or having too many people or too much noise in a control room.

In HROs there was also a "safety culture" that included widely shared values, attitudes and even personal identities among organizational personnel that supported precursor safety management through identifying and reporting any incursions into precursor zones. In addition, organizational members shared a commitment to improvement. (The plant manager at the Diablo Canyon nuclear plant in California was quoted as saying: "If we're not getting better, we're getting worse.") Among other things, improvement meant rethinking the precursor zone and often adding new precursors, both technical and organizational, to monitor.

Another prevalent set of attitudes and identities was the assumption of responsibility by many personnel for system reliability and safety and the prevention of system errors.

This meant that these personnel had a larger focus on errors and their precursors than simply the tasks in their job description (Anonymous [32]).

4.6 *Safety as the Management of Organizational Drift*

What also existed in HROs was an understanding that a key to maintaining the effectiveness of the processes noted above was a recognition that they are potentially perishable. Like the drift toward precursors, a lapse in close attention to micro-level detail or system-level attention under the power of job routine, or an erosion in inter-departmental communication and trust, could occur. The key to reliability and integrity in safety management processes was not the assumption or striving for their invariance but a management of their fluctuations (Anonymous [36]).

What existed was a special form of organizational resilience much of it centered around the recognition of drift into precursors, lapses in attention or communication and individuals speaking up about them triggering a rapid response to move away from these conditions. As one department manager in the Diablo Canyon nuclear plant said: "When I see evidence of a slack in care, I know its time to renew the fervor."

The processes noted above are important descriptors in a concept of safety. They can be observed, and even metrics applied to them as a means of assessing the ongoing safety of organizations. They constitute functions of safety management. As such they have implications for the role of safety managers in future organizations.

5. The Future of Safety Management

It is important to recognize that the technological, political and social characteristics of high reliability organizations are not typical for many organizations. In the future there may not be many, if any, HROs (LaPorte [16]; Anonymous [35]).

HRO's operated mature, well-understood technologies. They were not experimental or R&D organizations developing new, untried technical systems. Instead they provided repetitive outputs and services which allowed a long experience base to build up with respect to their

technical cores.

Because of their dreaded hazards, the events HRO's were attempting to preclude were not in dispute for prevention, and were widely accepted both within and beyond the organization. This was a strong political foundation for high reliability and safety expenditures and management.

5.1 *New Challenges to Reliability and Safety Management.*

But there are emerging challenges to safety management (including the HRO model): These include:

- More complicated technical systems especially in software and control systems. Many new systems now incorporate elements of AI in their design. The Boeing 737 MAX 8 Maneuvering Characteristics Augmentation System (MCAS) exceeded the ability of Federal Aviation Authority (FAA) inspectors to actually assess its safety implications, leading the agency to cede its certification authority to Boeing itself (N.Y. Times [24]).
- Rapid technological change (reducing experience among operators; and even a decline in preventive maintenance in anticipation of new technical updates)

"We are building systems for which we cannot anticipate or guard against unintended behavior. Complex systems today have so many potential states that exhaustively testing them before use is simply not feasible. The result is an increasing number of what engineers call the "unknown unknowns," i.e., unanticipated things that happen during use of a system. We can no longer assume that design errors will be detected and removed before use of a product or system." (Leveson, [19]).

In fact many organizations are in effect beta testing the technical systems they have adopted.

- The evolution of organizational networks and rising interconnected vulnerability:

"Many of today's complex systems require integrating parts designed and built by separate contractors or organizational entities. Even if each contractor or group takes steps to build quality into its own components, combining subsystems into a system introduces new paths to hazards that are not apparent when viewing the parts separately." (Leveson, [19])

- An inter-organizational network is also unlikely to have consistent safety commitments and management across the organizations in the network, especially in multi-organizational supply chains (Anonymous [38]).
- More turbulent and uncertain "problem environments":
    a. Global problems in scale and scope (e.g. climate change and pandemics) have more variables, complexity and can exceed managerial frameworks for safety management
    b. More rapid propagation of problems (speed of wildfires, floods, and pandemics) can exceed the appraisal and response capacities of human organizations.

c. More intersecting problems among traditionally separate policy and management domains (e.g. public health safety programs now address issues in drug abuse, gun control, etc.) also pose a challenge to safety management.

6. A New Challenge for Safety Management and A New Role for Safety Managers?

The new safety management challenges are likely to be ones of earlier high reliability organizations in precluding and preventing unwanted events from happening, but instead challenges of being prepared for uncertainty and likely surprise in the unfolding events of failures and disasters, and being resilient in responding to them. This is well illustrated in the challenges of safety management in the domain of critical infrastructures.

As noted earlier one of the most interesting and challenging features of modern critical infrastructures is their major vulnerability to the failure of other infrastructures. Unlike normal operations natural events such as storms, floods and earthquakes can reveal shifting interconnectivity and latent dependencies, of one upon another for the restoration of service. suddenly made manifest, among critical infrastructures. These are conditions under which safety and reliability as performance standards need to be supported *by resilience as an inter-organizational process* (Zuiderwijk, D., [44]; Anonymous [28]) and safety management will increasingly have to be an inter- organizational activity.

There is a difference, for example, between modelling and mitigating failure possibilities among infrastructures and the ability to put failed infrastructures back into operation. For this latter function it may be increasingly important to rely on personnel who operate and maintain the systems and have nuts and bolts experience with them, as well as ingenuity to create options to put broken infrastructure systems back together.

Interviews conducted with a number of infrastructure managers, operators and maintenance personnel have indicated that for many of them, in major natural events, a clarity can emerge at their levels in *perceptions of urgency*, *functional needs* and *specific actions* required for service restoration (Anonymous [33]).

Here, the experience, imagination and ingenuity of key players can be crucial, as is their collaborative capacity to achieve *mutual* clarity and coordinating action across those infrastructures that may now have *overlapping dependencies for service restoration* as a result of the emergency. Facilitating these processes can be an important part of safety management (Anonymous,[33]).

To do so safety managers will have to address the need for the development of an *infrastructure of inter-infrastructure resilience.* This could include making sure that there personnel across the infrastructures know and trust one another and that there are back-up communications capabilities and compatible communication and data sharing software among the

infrastructures.

Safety managers will have to defend preparations and expenditures that support this resilience through the supply of contingent resources to create options that facilitate coordinated inter-infrastructural improvisations for service restoration.

They will have to do so in relation to the higher-level authority and accountability valued within many emergency management frameworks. Within the U.S., for example, the Federal Emergency Management Agency (FEMA) has a National Incident Management System (NIMS, [9]) and the states have a federally mandated standardized Incident Command Center framework, under which improvisational actions taken at lower levels among safety managers, control operators and maintenance personnel may be seen as a drift away from clear accountability and an infringement of higher-level responsibilities at the Incident Command level to set priorities in light of their "bigger picture." But this bigger picture may well be a fuzzy one in that it may lack the clarity of the higher resolution available at the operational and maintenance level. Here people may have experience and skills in knowing how to put broken systems back together again.

7. Possible changes in "safety" as a performance standard?

a. Policy-making for safety under uncertainty becomes an R&D process that includes trial and error learning (e.g. COVID management and mask effectiveness).

b. Policy-making and management for safety will have to address a wider scope of variables across different policy and organizational domains, some that affect safety either as contributors to a hazard or as potential sources of hazard control (e.g. congested housing and international travel as sources of COVID-19 spread and school and public event closings as COVID controls).

c. Safety management must increasingly be a process beyond the prevention of events. It will have to include resilience in restoration and recovery from unpreventable events (D'Oliviero and Mario, [7]).

d. Finally, safety management will increasingly have to include the design and re-design of both technical and organizational elements.

9. Conclusion: A new role for design in safety management?

An early engineering argument for the role of design in safety was offered in 1891.The first paper dealing with the safe design of machinery was presented to the American Society of Mechanical Engineers by John H. Cooper in 1891:

"It is an easy task to formulate a plan of accident-preventing devices after the harm is done, but the wiser engineer foresees the possible weakness, as well as the dangers ahead, which are

involved in his new enterprise, and at once embodies all the necessary means of safety in his original design." (Cooper, [6])

The U.S. NRC now requires nuclear plants to analyze possibilities for both design-basis accidents but also *non-design-basis* accidents (USNRC [41]). This suggests new possibilities for *design* resilience with respect to safety.

9.1 A process of Sociotechnical Design?

Imagine a new approach to design with designers considering, in collaboration with social scientists, an enlarged set of organizational and social science variables. A sociotechnical design process could entail designing for situations and response actions that might actually be "outside of design" (Roberts, [30]). How can this be?

Sociotechnical design could be a process that could actually add to adaptive capacity, in part by designing in order to promote conditions that facilitate what might well be termed processes of *contingent re-design* in emergencies, facilitating adaptive actions, even improvisational ones to address currently unspecified events and conditions.

Design could be a backbone for safety resilience with respect to critical infrastructures. Addressing such things as:

--design strategies for infrastructure interoperability in equipment across organizations, such as robust and compatible communications and data software systems, compatible tools, pump and pipe fittings, and geographically distributed duplicate spare parts

-- designs can also support inter-infrastructure resilience, such including collaborative projects in contingency planning, table-top exercises and emergency simulations, and even personnel sharing on some maintenance and repair jobs. These could allow personnel in different infrastructures to meet, know and trust one another.

9.2 *A new approach to design errors.*

Sociotechnical design standards could also identify a set of socio-technical *design process errors* such *as:*

-- Designing a technical system without prior consultation with likely users.

-- Designing a technical system whose design basis for safe operating conditions and actions is so limited, complex or precise as to impose margins of error beyond the capacity of human organizations to fulfill.

-- Designing a technical system whose operation requires a transaction speed faster than a human can engage in and/or understand.

-- It is also a design error to deny control operators or pilots "claw back" capacities for manual control should newly developed automated systems fail. (A major problem with the Boeing 737

MAX 8's MCAS system.)

-- Finally, it is also a design error not to recognize that in emergencies:

"AI controllers differ from human operators and cannot be expected to contribute to adaptive capacity in identical ways. In particular, they are unable to come up with truly novel actions because they rely entirely on the *history* of a system's process states." (Gazos, et.al. [10]).

Avoiding errors like these can be a further way that design can enhance safety. But the starting point has to be the clarification of the concept of safety away from treatment as a noun and accepting it as a verb which will then allow the specification (and ultimate measurement) of processes and functions critical to adaptation and resilience. This requires that we come to recognize safety as a process of both prevention and recovery.

References

1. Antonsen, S., Aven, T. and Haavik, T. (2025). "Reflections on Research in Risk Science and On Risk Science." *Risk Analysis*.
2. Aven, T. (2025). Risk science and politics: What is and should be the relationship? *Risk Analysis*, 45(2), 274-282
3. Aven, T. and Ylonen, M. (2015). "A risk interpretation of sociotechnical safety perspectives." *Reliability Engineering & System Safety*, v. 175, July 18: 13-18.
4. Bambauer, D. (2014). "Ghost in the Network." *University of Pennsylvania Law Review*, v. 162; 5 (April), 1011-1091.
5. Center for Chemical Process Safety, (2024). *2024 CCPS Annual Report*.

(https://www.aiche.org/sites/default/files/docs/pages/ccps24annualreportweb20250204.pdf)

6. Cooper, J.H., (1891). "Accident-prevention devices applied to machines." *Transactions of the ASME*, 12:249-264.
7. D'Olivera, T.C. and Mario, C. (2010). "Organizational Safety: From Safety Culture to Resilience." *Proceedings of the 29th Conference of the European Association for Aviation Psychology*. (https://www.researchgate.net/publication/267926244_Organizational_safety_From_safety_culture_to_organizational_resilience).
8. Federal Emergency Management System (2017). *The National Incident Management System*. (https://training.fema.gov/nims/docs/nims.2017.instructor%20student%20learning%20materials.pdf)
9. Gazos, A., et.al. (2025). "Organising AI for safety: Identifying structural vulnerabilities to guide the design of AI-enhanced socio-technical systems." *Safety Science*, v. 184 (April).

10. Hollnagel, E. (2014). *Safety I and Safety II: The Past and Future of Safety Management* (Ashgate).

11. Iftikhar, S. (2024). "Cyberterrorism as a global threat: a review on repercussions and countermeasures." *PeerJ Computer Science*. (doi: 10.7717/peerj-cs.1772)

12. Institute of Nuclear Power Operators (INPO), 2012. *Traits of a Healthy Nuclear Safety Culture* (https://www.nrc.gov/docs/ml1303/ml13031a707.pdf)

13. Kasmalkar, I. et.al. (2020). "When floods hit the road: Resilience to flood-related traffic disruption in the San Francisco Bay Area and beyond." *Science Advances*, v. 6: 1-8.

14. LaPorte, T.R. and Consolini, P. (1991). "Working in Practice But Not in Theory: Theoretical Challenges of High Reliability Organizations." *Public Administration Research and Theory*, VC. 1 (1): 19-047.

15. LaPorte, T.R. (1996). "High Reliability Organizations; Unlikely, Demanding and At Risk." *Journal of Contingencies and Crisis Management*, v. 2 (2): 60-71.

16. Lederer, Jerome, (1986). "How far have we come? A look back at the leading edge of system safety eighteen years ago." *Hazard Prevention*, pp. 8-10, May/June.

17. Leveson, N., Dulac, N., Marais, K. and Carroll, J. (2009). "Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems. *Organization Studies* 30, 227–249.

18. Leveson, N. (2020). "System III: A Systems Approach to Safety and Resilience." *M.I.T Systems Engineering Lab* (July). (http://sunnyday.mit.edu/safety-3.pdf).

19. Lindhout, P., Kingston-Howlett, J., Hansen, F.T. and Reniers, G. (2020). "Reducing unknown risk: The safety engineers' new horizon." *Journal of Loss Prevention in the Process Industries*, 68 (104330): 2-11

20. Macaig, M. (2018). "Risky Waters". *Governing*, v. 31, 11 (August), 32-39.

21. Macrae, C. (2025). "Regulating Reliably: Building High-Reliability Regulators in Healthcare" *Journal of the Royal Society of Medicine*, Vol. 118(1), 11-15.

22. Nævestad, R-O, Hesjevoll, O.S. and Elvik, R. (2021). "How can regulatory authorities improve safety in organizations by influencing safety culture? A conceptual model of the relationships and a discussion of implications." *Accident Analysis and Prevention*, v. 159. (https://doi.org/10.1016/j.aap.2021.106228).

23. New York Times (2019). "The Roots of Boeing's 737 MAX Crisis: A Regulator Relaxes its Oversight". (https://www.nytimes.com/2019/07/27/business/boeing-737-max-faa.html).

24. Olsen, O.E., Kruke, B.I and Hovden, J. (2007). "Societal Safety: Concepts, Borders and Dilemmas." *Journal of Contingencies and Crisis Management*, v. 15, 2 (June).

25. Radeloff, V.C. et.al. (2022). “Rapid growth of the US wildland-urban interface raises wildfire risk.” *The Conversation* (https://doi.org/10.1073/pnas.1718850115

26. Reason, J. (2010). “Safety Paradoxes and Safety Culture.” *Injury Control and Safety Promotion*, v. 7 (1): 3-14.

27. Rentschler, J. et.al. (2022). “Rapid urban growth in flood zones: Global trends in exposure since 1985.” (World Bank, 2022) (“https://blogs.worldbank.org/en/developmenttalk/rapid-urban-growth-flood-zones-global-trends-exposure-1985).

28. Roberts, K. (1993). *New Challenges to Organizations*. Macmillan.

29. Roberts, V. L. (1984). “Defensive design.” *Mechanical Engineering*, (September), pp. 88-93.

30. Rochlin, G.I., et.al. (1987). “The Self-Designing High Reliability Organization: Aircraft Carrier Flight Operations at Sea.” *Naval War College Review*, v. 40, no. 4: 76-90.

31. Roe, E. and Schulman, P., (2016). Reliability and Risk (Stanford University Press).

32. Anonymous (2025). Details omitted for double-anonymized viewing.

33. Safety Management International Collaboration Group (2013). *Measuring Performance Guidelines for Service Providers*. (https://skybrary.aero/sites/default/files/bookshelf/2395.pdf)

34. Savage, S. (2012). *The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty.* (Hoboken, NJ: John Wiley)

35. Schulman, P. (1993). “The ‘Negotiated Order” of High Reliability,” Administration and Society, v. 25, no. 3 (November).

36. Anonymous (2022). Details omitted for double-anonymized viewing.

37. Anonymous (2023). Details omitted for double-anonymized viewing).

38. Sheard, S. (2015). “Aircraft Systems: Three Principles for Mitigating Complexity.” *Software Engineering Institute, Carnegie Mellon University*. (https://www.sei.cmu.edu/blog/aircraft-systems-three-principles-for-mitigating-complexity/).

39. U.S. Occupational Health and Safety Administration (OSHA) (1992). *Process Safety Management (PSM) standard for highly hazardous chemicals (HHCs)*. 29 CFR 1910.119. (https://www.ecfr.gov/current/title-29/subtitle-B/chapter-XVII/part-1910/subpart-H/section-1910.119).

40. U.S. Nuclear Regulatory Commission (2015*). Mitigation of Beyond Design Basis Events*. (https://www.regulations.gov/document/NRC-2011-0189-0040).

41. Weart, S. (2012). *The Rise of Nuclear Fear*. Harvard University Press).

42. Weick, K. (2011). “Organizing for Transient Reliability: The Production of Dynamic Non-Events.” *Journal of Contingencies and Crisis Management*, 19 (1) (March), 21-27.

43. Zuiderwijk, D. (2024). *Reliable beyond Design: The Role of Operational Actors in Dealing with Increasing Uncertainty in Complex Temporary Technical Organizations*. Ph.D. Dissertation, Vrije Universiteit Amsterdam