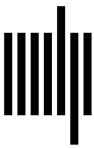
Engineering a Safer World

Systems Thinking Applied to Safety

By: Nancy G. Leveson

Citation:

Engineering a Safer World: Systems Thinking Applied to Safety
By: Nancy G. Leveson
DOI: 10.7551/mitpress/8179.001.0001
ISBN (electronic): 9780262298247
Publisher: The MIT Press
Published: 2016



The MIT Press

12 Controlling Safety during Operations

In some industries, system safety is viewed as having its primary role in development and most of the activities occur before operations begin. Those concerned with safety may lose influence and resources after that time. As an example, one of the chapters in the *Challenger* accident report, titled "The Silent Safety Program," lamented:

> Following the successful completion of the orbital flight test phase of the Shuttle program, the system was declared to be operational. Subsequently, several safety, reliability, and quality assurance organizations found themselves with reduced and/or reorganized functional capabilities. . . . The apparent reason for such actions was a perception that less safety, reliability, and quality assurance activity would be required during "routine" Shuttle operations. This reasoning was faulty.

While safety-guided design eliminates some hazards and creates controls for others, hazards and losses may still occur in operations due to:

- Inadequate attempts to eliminate or control the hazards in the system design, perhaps due to inappropriate assumptions about operations.
- Inadequate implementation of the controls that designers assumed would exist during operations.
- Changes that occur over time, including violation of the assumptions underlying the design.
- Unidentified hazards, sometimes new ones that arise over time and were not anticipated during design and development.

Treating operational safety as a control problem requires facing and mitigating these potential reasons for losses.

A complete system safety program spans the entire life of the system and, in some ways, the safety program during operations is even more important than during development. System safety does not stop after development; it is just getting started. The focus now, however, shifts to the operations safety control structure.

This chapter describes the implications of STAMP on operations. Some topics that are relevant here are left to the next chapter on management: organizational design, safety culture and leadership, assignment of appropriate responsibilities throughout the safety control structure, the safety information system, and corporate safety policies. These topics span both development and operations and many of the same principles apply to each, so they have been put into a separate chapter. A final section of this chapter considers the application of STAMP and systems thinking principles to occupational safety.

12.1 Operations Based on STAMP

Applying the basic principles of STAMP to operations means that, like development, the goal during operations is enforcement of the safety constraints, this time on the operating system rather than in its design. Specific responsibilities and control actions required during operations are outlined in chapter 13.

Figure 12.1 shows the interactions between development and operations. At the end of the development process, the safety constraints, the results of the hazard analyses, as well as documentation of the safety-related design features and design rationale, should be passed on to those responsible for the maintenance and evolution of the system. This information forms the baseline for safe operations. For example, the identification of safety-critical items in the hazard analysis should be used as input to the maintenance process for prioritization of effort.

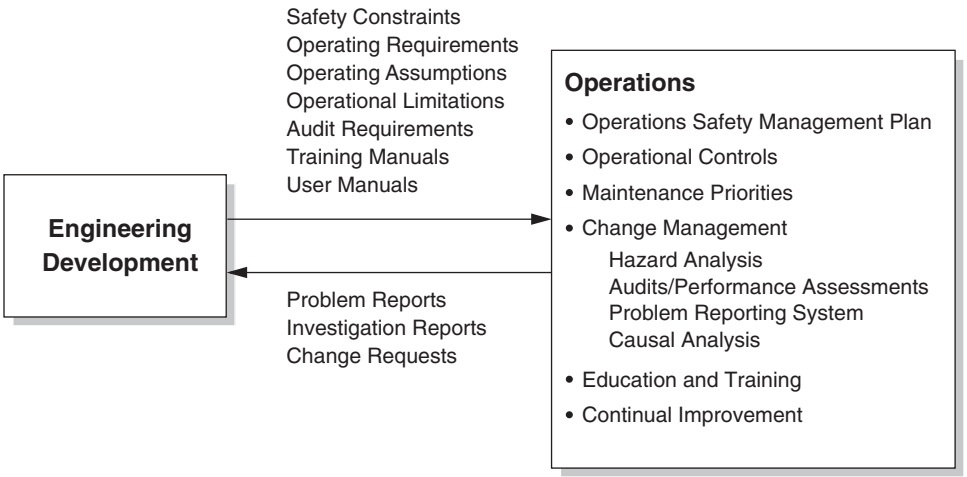


Figure 12.1
The relationship between development and operations.

At the same time, the accuracy and efficacy of the hazard analyses performed during development and the safety constraints identified need to be evaluated using the operational data and experience. Operational feedback on trends, incidents, and accidents should trigger reanalysis when appropriate. Linking the assumptions throughout the system specification with the parts of the hazard analysis based on that assumption will assist in performing safety maintenance activities. During field testing and operations, the links and recorded assumptions and design rationale can be used in safety change analysis, incident and accident analysis, periodic audits and performance monitoring as required to ensure that the operational system is and remains safe.

For example, consider the TCAS requirement that TCAS provide collision avoidance protection for any two aircraft closing horizontally at any rate up to 1,200 knots and vertically up to 10,000 feet per minute. As noted in the rationale, this requirement is based on aircraft performance limits at the time TCAS was created. It is also based on minimum horizontal and vertical separation requirements. The safety analysis originally performed on TCAS is based on these assumptions. If aircraft performance limits change or if there are proposed changes in airspace management, as is now occurring in new Reduced Vertical Separation Minimums (RVSM), hazard analysis to determine the safety of such changes will require the design rationale and the tracing from safety constraints to specific system design features as recorded in intent specifications. Without such documentation, the cost of reanalysis could be enormous and in some cases even impractical. In addition, the links between design and operations and user manuals in level 6 will ease updating when design changes are made.

In a traditional System Safety program, much of this information is found in or can be derived from the hazard log, but it needs to be pulled out and provided in a form that makes it easy to locate and use in operations. Recording design rationale and assumptions in intent specifications allows using that information both as the criteria under which enforcement of the safety constraints is predicated and in the inevitable upgrades and changes that will need to be made during operations. Chapter 10 shows how to identify and record the necessary information.

The design of the operational safety controls are based on assumptions about the conditions during operations. Examples include assumptions about how the operators will operate the system and the environment (both social and physical) in which the system will operate. These conditions may change. Therefore, not only must the assumptions and design rationale be conveyed to those who will operate the system, but there also need to be safeguards against changes over time that violate those assumptions.

The changes may be in the behavior of the system itself:

- Physical changes: the equipment may degrade or not be maintained properly.
- Human changes: human behavior and priorities usually change over time.
- Organizational changes: change is a constant in most organizations, including changes in the safety control structure itself, or in the physical and social environment within which the system operates or with which it interacts.

Controls need to be established to reduce the risk associated with all these types of changes.

The safeguards may be in the design of the system itself or in the design of the operational safety control structure. Because operational safety depends on the accuracy of the assumptions and models underlying the design and hazard analysis processes, the operational system should be monitored to ensure that:

1. The system is constructed, operated, and maintained in the manner assumed by the designers.
2. The models and assumptions used during initial decision making and design are correct.
3. The models and assumptions are not violated by changes in the system, such as workarounds or unauthorized changes in procedures, or by changes in the environment.

Designing the operations safety control structure requires establishing controls and feedback loops to (1) identify and handle flaws in the original hazard analysis and system design and (2) to detect unsafe changes in the system during operations before the changes lead to losses. Changes may be intentional or they may be unintended and simply normal changes in system component behavior or the environment over time. Whether intended or unintended, system changes that violate the safety constraints must be controlled.

12.2 Detecting Development Process Flaws during Operations

Losses can occur due to flaws in the original assumptions and rationale underlying the system design. Errors may also have been made in the hazard analysis process used during system design. During operations, three goals and processes to achieve these goals need to be established:

1. Detect safety-related flaws in the system design and in the safety control structure, hopefully before major losses, and fix them.

2. Determine what was wrong in the development process that allowed the flaws to exist and improve that process to prevent the same thing from happening in the future.
3. Determine whether the identified flaws in the process might have led to other vulnerabilities in the operational system.

If losses are to be reduced over time and companies are not going to simply engage in constant firefighting, then mechanisms to implement learning and continual improvement are required. Identified flaws must not only be fixed (symptom removal), but the larger operational and development safety control structures must be improved, as well as the process that allowed the flaws to be introduced in the first place. The overall goal is to change the culture from *a fixing orientation*—identifying and eliminating deviations or symptoms of deeper problems—to a *learning orientation* where systemic causes are included in the search for the source of safety problems [33].

To accomplish these goals, a feedback control loop is needed to regularly track and assess the effectiveness of the development safety control structure and its controls. Were hazards overlooked or incorrectly assessed as unlikely or not serious? Were some potential failures or design errors not included in the hazard analysis? Were identified hazards inappropriately accepted rather than being fixed? Were the designed controls ineffective? If so, why?

When numerical risk assessment techniques are used, operational experience can provide insight into the accuracy of the models and probabilities used. In various studies of the DC-10 by McDonnell Douglas, the chance of engine power loss with resulting slat damage during takeoff was estimated to be less than one in a billion flights. However, this highly improbable event occurred four times in DC-10s in the first few years of operation without raising alarm bells before it led to an accident and changes were made. Even one event should have warned someone that the models used might be incorrect. Surprisingly little scientific evaluation of probabilistic risk assessment techniques has ever been conducted [115], yet these techniques are regularly taught to most engineering students and widely used in industry. Feedback loops to evaluate the assumptions underlying the models and the assessments produced are an obvious way to detect problems.

Most companies have an accident/incident analysis process that identifies the proximal failures that led to an incident, for example, a flawed design of the pressure relief valve in a tank. Typical follow-up would include replacement of that valve with an improved design. On top of fixing the immediate problem, companies should have procedures to evaluate and potentially replace all the uses of that pressure relief valve design in tanks throughout the plant or company. Even better would be to reevaluate pressure relief valve design for all uses in the plant, not just in tanks.

But for long-term improvement, a causal analysis—CAST or something similar—needs to be performed on the process that created the flawed design and that process improved. If the development process was flawed, perhaps in the hazard analysis or design and verification, then fixing that process can prevent a large number of incidents and accidents in the future.

Responsibility for this goal has to be assigned to an appropriate component in the safety control structure and feedback-control loops established. Feedback may come from accident and incident reports as well as detected and reported design and behavioral anomalies. To identify flaws before losses occur, which is clearly desirable, audits and performance assessments can be used to collect data for validating and informing the safety design and analysis process without waiting for a crisis. There must also be feedback channels to the development safety control structure so that appropriate information can be gathered and used to implement improvements. The design of these control loops is discussed in the rest of this chapter. Potential challenges in establishing such control loops are discussed in the next chapter on management.

12.3 Managing or Controlling Change

Systems are not static but instead are dynamic processes that are continually adapting to achieve their ends and to react to changes in themselves and their environment. In STAMP, adaptation or change is assumed to be an inherent part of any system, particularly those that include humans and organizational components: Humans and organizations optimize and change their behavior, adapting to the changes in the world and environment in which the system operates.

To avoid losses, not only must the original design enforce the safety constraints on system behavior, but the safety control structure must continue to enforce them as changes to the designed system, including the safety control structure itself, occur over time.

While engineers usually try to anticipate potential changes and to design for changeability, the bulk of the effort in dealing with change must necessarily occur during operations. Controls are needed both to prevent unsafe changes and to detect them if they occur.

In the friendly fire example in chapter 5, the AWACS controllers stopped handing off helicopters as they entered and left the no-fly zone. They also stopped using the Delta Point system to describe flight plans, although the helicopter pilots assumed the coded destination names were still being used and continued to provide them. Communication between the helicopters and the AWACS controllers was seriously degraded although nobody realized it. The basic safety constraint that all aircraft in the no-fly zone and their locations would be known to the AWACS controllers

became over time untrue as the AWACS controllers optimized their procedures. This type of change is normal; it needs to be identified by checking that the assumptions upon which safety is predicated remain true over time.

The deviation from assumed behavior during operations was not, in the friendly fire example, detected until after an accident. Obviously, finding the deviations at this time is less desirable than using audits, and other types of feedback mechanisms to detect hazardous changes, that is, those that violate the safety constraints, before losses occur. Then something needs to be done to ensure that the safety constraints are enforced in the future.

Controls are required for both intentional (planned) and unintentional changes.

12.3.1 Planned Changes

Intentional system changes are a common factor in accidents, including physical, process, and safety control structure changes [115]. The Flixborough explosion provides an example of a temporary physical change resulting in a major loss: Without first performing a proper hazard analysis, a temporary pipe was used to replace a reactor that had been removed to repair a crack. The crack itself was the result of a previous process modification [54]. The Walkerton water contamination loss in appendix C provides an example of a control structure change when the government water testing lab was privatized without considering how that would affect feedback to the Ministry of the Environment.

Before any planned changes are made, including organizational and safety control structure changes, their impact on safety must be evaluated. Whether this process is expensive depends on how the original hazard analysis was performed and particularly how it was documented. Part of the rationale behind the design of intent specifications was to make it possible to retrieve the information needed.

While implementing change controls limits flexibility and adaptability, at least in terms of the time it takes to make changes, the high accident rate associated with intentional changes attests to the importance of controlling them and the high level of risk being assumed by not doing so. Decision makers need to understand these risks before they waive the change controls.

Most systems and industries do include such controls, usually called Management of Change (MOC) procedures. But the large number of accidents occurring after system changes without evaluating their safety implies widespread nonenforcement of these controls. Responsibility needs to be assigned for ensuring compliance with the MOC procedures so that change analyses are conducted and the results are not ignored. One way to do this is to reward people for safe behavior when they choose safety over other system goals and to hold them accountable when they choose to ignore the MOC procedures, even when no accident results. Achieving this goal, in

turn, requires management commitment to safety (see chapter 13), as does just about every aspect of building and operating a safe system.

12.3.2 Unplanned Changes

While dealing with planned changes is relatively straightforward (even if difficult to enforce), unplanned changes that move systems toward states of higher risk are less straightforward. There need to be procedures established to prevent or detect changes that impact the ability of the operations safety control structure and the designed controls to enforce the safety constraints.

As noted earlier, people will tend to optimize their performance over time to meet a variety of goals. If an unsafe change is detected, it is important to respond quickly. People incorrectly reevaluate their perception of risk after a period of success. One way to interrupt this risk-reevaluation process is to intervene quickly to stop it before it leads to a further reduction in safety margins or a loss occurs. But that requires an alerting function to provide feedback to someone who is responsible for ensuring that the safety constraints are satisfied.

At the same time, change is a normal part of any system. Successful systems are continually changing and adapting to current conditions. Change should be allowed as long as it does not violate the basic constraints on safe behavior and therefore increase risk to unacceptable levels. While in the short term relaxing the safety constraints may allow other system goals to be achieved to a greater degree, in the longer term accidents and losses can cost a great deal more than the short-term gains.

The key is to allow flexibility in how safety goals are achieved, but not flexibility in violating them, and to provide the information that creates accurate risk perception by decision makers.

Detecting migration toward riskier behavior starts with identifying baseline requirements. The requirements follow from the hazard analysis. These requirements may be general ("Equipment will not be operated above the identified safety-critical limits" or "Safety-critical equipment must be operational when the system is operating") or specifically tied to the hazard analysis ("AWACS operators must always hand off aircraft when they enter and leave the no-fly zone" or "Pilots must always follow the TCAS alerts and continue to do so until they are canceled").

The next step is to assign responsibility to appropriate places in the safety control structure to ensure the baseline requirements are not violated, while allowing changes that do not raise risk. If the baseline requirements make it impossible for the system to achieve its goals, then instead of waiving them, the entire safety control structure should be reconsidered and redesigned. For example, consider the foam shedding problems on the Space Shuttle. Foam had been coming off the external tank for most of the operational life of the Shuttle. During development, a hazard had been identified and documented related to the foam damaging the thermal

control surfaces of the spacecraft. Attempts had been made to eliminate foam shedding, but none of the proposed fixes worked. The response was to simply waive the requirement before each flight. In fact, at the time of the *Columbia* loss, more than three thousand potentially critical failure modes were regularly waived on the pretext that nothing could be done about them and the Shuttle had to fly [74]. More than a third of these waivers had not been reviewed in the ten years before the accident.

After the *Columbia* loss, controls and mitigation measures for foam shedding were identified and implemented, such as changing the fabrication procedures and adding cameras and inspection and repair capabilities and other contingency actions. The same measures could, theoretically, have been implemented before the loss of *Columbia*. Most of the other waived hazards were also resolved in the aftermath of the accident. While the operational controls to deal with foam shedding raise the risk associated with a Shuttle accident above actually fixing the problem, the risk is lower than simply ignoring and waiting for the hazards to occur. Understanding and explicitly accepting risk is better than simply denying and ignoring it.

The NASA safety program and safety control structure had seriously degraded before both the *Challenger* and *Columbia* losses [117]. Waiving requirements interminably represents an abdication of the responsibility to redesign the system, including the controls during operations, after the current design is determined to be unsafe.

Is such a hard line approach impractical? SUBSAFE, the U.S. nuclear submarine safety program established after the *Thresher* loss, described in chapter 14, has not allowed waiving the SUBSAFE safety requirements for more than forty-five years, with one exception. In 1967, four years after SUBSAFE was established, SUBSAFE requirements for one submarine were waived in order to satisfy pressing Navy performance goals. That submarine and its crew were lost less than a year later. The same mistake has not been made again.

If there is absolutely no way to redesign the system to be safe and at the same time to satisfy the system requirements that justify its existence, then the existence of the system itself should be rethought and a major replacement or new design considered. After the first accident, much more stringent and perhaps unacceptable controls will be forced on operations. While the decision to live with risk is usually accorded to management, those who will suffer the losses should have a right to participate in that decision. Luckily, the choice is usually not so stark if flexibility is allowed in the way the safety constraints are maintained and long-term rather than short-term thinking prevails.

Like any set of controls, unplanned change controls involve designing appropriate control loops. In general, the process involves identifying the responsibility of the controller(s); collecting data (feedback); turning the feedback into useful

information (analysis) and updating the process models; generating any necessary control actions and appropriate communication to other controllers; and measuring how effective the whole process is (feedback again).

12.4 Feedback Channels

Feedback is a basic part of STAMP and of treating safety as a control problem. Information flow is key in maintaining safety.

There is often a belief—or perhaps hope—that a small number of "leading indicators" can identify increasing risk of accidents, or, in STAMP terms, migration toward states of increased risk. It is unlikely that general leading indicators applicable to large industry segments exist or will be useful. The identification of system safety constraints does, however, provide the possibility of identifying leading indicators applicable to a specific system.

The desire to predict the future often leads to collecting a large amount of information based on the hope that something useful will be obtained and noticed. The NASA Space Shuttle program was collecting six hundred metrics a month before the loss of *Columbia*. Companies often collect data on occupational safety, such as days without a lost time accident, and they assume that these data reflect on system safety [17], which of course it does not. Not only is this misuse of data potentially misleading, but collecting information that may not be indicative of real risk diverts limited resources and attention from more effective risk-reduction efforts.

Poorly defined feedback can lead to a decrease in safety. As an incentive to reduce the number of accidents in the California construction industry, for example, workers with the best safety records—as measured by fewest reported incidents—were rewarded [126]. The reward created an incentive to withhold information about small accidents and near misses, and they could not therefore be investigated and the causes eliminated. Under-reporting of incidents created the illusion that the system was becoming safer, when instead risk had merely been muted. The inaccurate risk perception by management led to not taking the necessary control actions to reduce risk. Instead, the reporting of accidents should have been rewarded.

Feedback requirements should be determined with respect to the design of the organization's safety control structure, the safety constraints (derived from the system hazards) that must be enforced on system operation, and the assumptions and rationale underlying the system design for safety. They will be similar for different organizations only to the extent that the hazards, safety constraints, and system design are similar.

The hazards and safety constraints, as well as the causal information derived by the use of STPA, form the foundation for determining what feedback is necessary to provide the controllers with the information they need to satisfy their safety

responsibilities. In addition, there must be mechanisms to ensure that feedback channels are operating effectively.

The feedback is used to update the controller's process models and understanding of the risks in the processes they are controlling, to update their control algorithms, and to execute appropriate control actions.

Sometimes, cultural problems interfere with feedback about the state of the controlled process. If the culture does not encourage sharing information and if there is a perception that the information can be used in a way that is detrimental to those providing it, then cultural changes will be necessary. Such changes require leadership and freedom from blame (see "Just Culture" in chapter 13). Effective feedback collection requires that those making the reports are convinced that the information will be used for constructive improvements in safety and not as a basis for criticism or disciplinary action. Resistance to airing dirty laundry is understandable, but this quickly transitions into an organizational culture where only good news is passed on for fear of retribution. Everyone's past experience includes individual mistakes, and avoiding repeating the same mistakes requires a culture that encourages sharing.

Three general types of feedback are commonly used: audits and performance assessments; reporting systems; and anomaly, incident, and accident investigation.

12.4.1 Audits and Performance Assessments

Once again, audits and performance assessments should start from the safety constraints and design assumptions and rationale. The goal should be to determine whether the safety constraints are being enforced in the operation of the system and whether the assumptions underlying the safety design and rationale are still true. Audits and performance assessments provide a chance to detect whether the behavior of the system and the system components still satisfies the safety constraints and whether the way the controllers think the system is working—as reflected in their process models—is accurate.

The entire safety control structure must be audited, not just the lower-level processes. Auditing the upper levels of the organization will require buy-in and commitment from management and an independent group at a high enough level to control audits as well as explicit rules for conducting them.

Audits are often less effective than they might be. When auditing is performed through contracts with independent companies, there may be subtle pressures on the audit team to be unduly positive or less than thorough in order to maintain their customer base. In addition, behavior or conditions may be changed in anticipation of an audit and then revert back to their normal state immediately afterward.

Overcoming these limitations requires changes in organizational culture and in the use of the audit results. Safety controllers (managers) must feel personal

responsibility for safety. One way to encourage this view is to trust them and expect them to be part of the solution and to care about safety. "Safety is everyone's responsibility" must be more than an empty slogan, and instead a part of the organizational culture.

A *participatory audit* philosophy can have an important impact on these cultural goals. Some features of such a philosophy are:

- Audits should not be punitive. Audits need to be viewed as a chance to improve safety and to evaluate the process rather than a way to evaluate employees.
- To increase buy-in and commitment, those controlling the processes being audited should participate in creating the rules and procedures and understand the reasons for the audit and how the results will be used. Everyone should have a chance to learn from the audit without it having negative consequences—it should be viewed as an opportunity to learn how to improve.
- People from the process being audited should participate on the audit team. In order to get an outside but educated view, using process experts from other parts of the organization not directly being audited is a better approach than using outside audit companies. Various stakeholders in safety may be included such as unions. The goal should be to inculcate the attitude that this is *our* audit and a chance to improve *our* practices. Audits should be treated as a learning experience for everyone involved—including the auditors.
- Immediate feedback should be provided and solutions discussed. Often audit results are not available until after the audit and are presented in a written report. Feedback and discussion with the audit team during the audit are discouraged. One of the best times to discuss problems found and how to design solutions, however, is when the team is together and on the spot. Doing this will also reinforce the understanding that the goal is to improve the process, not to punish or evaluate those involved.
- All levels of the safety control structure should be audited, along with the physical process and its immediate operators. Accepting being audited and implementing improvements as a result—that is, leading by example—is a powerful way for leaders to convey their commitment to safety and to its improvement.
- A part of the audit should be to determine the level of safety knowledge and training that actually exists, not what managers believe exists or what exists in the training programs and user manuals. These results can be fed back into the training materials and education programs. Under no circumstances, of course, should such assessments be used in a negative way or one that is viewed as punitive by those being assessed.

Because these rules for audits are so far from common practice, they may be viewed as unrealistic. But this type of audit is carried out today with great success. See chapter 14 for an example. The underlying philosophy behind these practices is that most people do not want to harm others and have innate belief in safety as a goal. The problems arise when other goals are rewarded or emphasized over safety. When safety is highly valued in an organizational culture, obtaining buy-in is usually not difficult. The critical step lies in conveying that commitment.

12.4.2 Anomaly, Incident, and Accident Investigation

Anomaly, incident, and accident investigations often focus on a single "root" cause and look for contributory causes near the events. The belief that there is a root cause, sometimes called *root cause seduction* [32], is powerful because it provides an illusion of control. If the root cause can simply be eliminated and if that cause is low in the safety control structure, then changes can easily be made that will eliminate accidents without implicating management or requiring changes that are costly or disruptive to the organization. The result is that physical design characteristics or low-level operators are usually identified as the root cause.

Causality is, however, much more complex than this simple but very entrenched belief, as has been argued throughout this book. To effect high-leverage policies and changes that are able to prevent large classes of future losses, the weaknesses in the entire safety control structure related to the loss need to be identified and the control structure redesigned to be more effective.

In general, effective learning from experience requires a change from a fixing orientation to a continual learning and improvement culture. To create such a culture requires high-level leadership by management, and sometimes organizational changes.

Chapter 11 describes a way to perform better analyses of anomalies, incidents, and accidents. But having a process is not enough; the process must be embedded in an organizational structure that allows the successful exploitation of that process. Two important organizational factors will impact the successful use of CAST: training and follow-up.

Applying systems thinking to accident analysis requires training and experience. Large organizations may be able to train a group of investigators or teams to perform CAST analyses. This group should be managerially and financially independent. Some managers prefer to have accident/incident analysis reports focus on the low-level system operators and physical processes and the reports never go beyond those factors. In other cases, those involved in accident analysis, while well-meaning, have too limited a view to provide the perspective required to perform an adequate causal analysis. Even when intentions are good and local skills and knowledge are available, budgets may be so tight and pressures to maintain performance

schedules so high that it is difficult to find the time and resources to do a thorough causal analysis using local personnel. Trained teams with independent budgets can overcome some of these obstacles. But while the leaders of investigations and causal analysis can be independent, participation by those with local knowledge is also important.

A second requirement is *follow-up*. Often the process stops after recommendations are made and accepted. No follow-up is provided to ensure that the recommendations are implemented or that the implementations were effective. Deadlines and assignment of responsibility for making recommendations, as well as responsibility for ensuring that they are made, are required. The findings in the causal analysis should be an input to future audits and performance assessments. If the same or similar causes recur, then that itself requires an analysis of why the problem was not fixed when it first was detected. Was the fix unsuccessful? Did the system migrate back to the same high-risk state because the underlying causal factors were never successfully controlled? Were factors missed in the original causal analysis? Trend analysis is important to ensure that progress is being made in controlling safety.

12.4.3 Reporting Systems

Accident reports very often note that before a loss, someone detected an anomaly but never reported it using the official reporting system. The response in accident investigation reports is often to recommend that the requirement to use reporting systems be emphasized to personnel or to provide additional training in using them. This response may be effective for a short time, but eventually people revert back to their prior behavior. A basic assumption about human behavior in this book (and in systems approaches to human factors) is that human behavior can usually be explained by looking at the system in which the human is operating. The reason in the system design for the behavior must be determined and changed: Simply trying to force people to behave in ways that are unnatural for them will usually be unsuccessful.

So the first question to ask is why people do not use reporting systems and to fix those factors. One obvious reason is that they may be designed poorly. They may require extra, time-consuming steps, such as logging into a web-based system, that are not part of their normal operating procedures or environment. Once they get to the website, they may be faced with a poorly designed form that requires them to provide a lot of extraneous information or does not allow the flexibility necessary to enter the information they want to provide.

A second reason people do not report is that the information they provided in the past appeared to go into a black hole, with nobody responding to it. There is little incentive to continue to provide information under these conditions, particularly when the reporting system is time-consuming and awkward to use.

A final reason for lack of reporting is a fear that the information provided may be used against them or there are other negative repercussions such as a necessity to spend time filling out additional reports.

Once the reason for failing to use reporting systems is understood, the solutions usually become obvious. For example, the system may need to be redesigned so it is easy to use and integrated into normal work procedures. As an example, email is becoming a primary means of communication at work. The first natural response in finding a problem is to contact those who can fix it, not to report it to some database where there is no assurance it will be processed quickly or get to the right people. A successful solution to this problem used on one large air traffic control system was to require only that the reporter add an extra "cc:" on their emails in order to get it reported officially to safety engineering and those responsible for problem reports [94].

In addition, the receipt of a problem report should result in both an acknowledgment of receipt and a thank-you. Later, when a resolution is identified, information should be provided to the reporter of the problem about what was done about it. If there is no resolution within a reasonable amount of time, that too should be acknowledged. There is little incentive to use reporting systems if the reporters do not think the information will be acted upon.

Most important, an effective reporting system requires that those making the reports are convinced the information will be used for constructive improvements in safety and not as a basis for criticism or disciplinary action. If reporting is considered to have negative consequences for the reporter, then anonymity may be necessary and a written policy provided for the use of such reporting systems, including the rights of the reporters and how the reported information will be used. Much has been written about this aspect of reporting systems (e.g., see Dekker [51]). One warning is that trust is hard to gain and easy to lose. Once it is lost, regaining it is even harder than getting buy-in at the beginning.

When reporting involves an outside regulatory agency or industry group, protection of safety information and proprietary data from disclosure and use for purposes other than improving safety must be provided.

Designing effective reporting systems is very difficult. Examining two successful efforts, in nuclear power and in commercial aviation, along with the challenges they face is instructive.

Nuclear Power

Operators of nuclear power plants in the United States are required to file a Licensee Event Report (LER) with the Nuclear Regulatory Commission (NRC) whenever an irregular event occurs during plant operation. While the NRC collected an enormous amount of information on the operating experience of plants in this

way, the data were not consistently analyzed until after the Three Mile Island (TMI) accident. The General Accounting Office (GAO) had earlier criticized the NRC for this failure, but no corrective action was taken until after the events at TMI [98].

The system also had a lack of closure: important safety issues were raised and studied to some degree, but were not carried through to resolution [115]. Many of the conditions involved in the TMI accident had occurred previously at other plants but nothing had been done about correcting them. Babcock and Wilcox, the engineering firm for TMI, had no formal procedures to analyze ongoing problems at plants they had built or to review the LERs on their plants filed with the NRC.

The TMI accident sequence started when a pilot-operated relief valve stuck open. In the nine years before the TMI incident, eleven of those valves had stuck open at other plants, and only a year before, a sequence of events similar to those at TMI had occurred at another U.S. plant.

The information needed to prevent TMI was available, including the prior incidents at other plants, recurrent problems with the same equipment at TMI, and engineers' critiques that operators had been taught to do the wrong thing in specific circumstances, yet nothing had been done to incorporate this information into operating practices.

In reflecting on TMI, the utility's president, Herman Dieckamp, said:

> To me that is probably one of the most significant learnings of the whole accident [TMI] the degree to which the inadequacies of that experience feedback loop . . . significantly contributed to making us and the plant vulnerable to this accident [98].

As a result of this wake-up call, the nuclear industry initiated better evaluation and follow-up procedures on LERs. It also created the Institute for Nuclear Power Operations (INPO) to promote safety and reliability through external reviews of performance and processes, training and accreditation programs, events analysis, sharing of operating information and best practices, and special assistance to member utilities. The IAEA (International Atomic Energy Agency) and World Association of Nuclear Operators (WANO) share these goals and serve similar functions worldwide.

The reporting system now provides a way for operators of each nuclear power plant to reflect on their own operating experience in order to identify problems, interpret the reasons for these problems, and select corrective actions to ameliorate the problems and their causes. Incident reviews serve as important vehicles for self-analysis, knowledge sharing across boundaries inside and outside specific plants, and development of problem-resolution efforts. Both INPO and the NRC issue various letters and reports to make the industry aware of incidents as part of operating experience feedback, as does IAEA's Incident Reporting System.

The nuclear engineering experience is not perfect, of course, but real strides have been made since the TMI wakeup call, which luckily occurred without major human losses. To their credit, an improvement and learning effort was initiated and has continued. High-profile incidents like TMI are rare, but smaller scale self-analyses and problem-solving efforts follow detection of small defects, near misses, and precursors and negative trends. Occasionally the NRC has stepped in and required changes. For example, in 1996 the NRC ordered the Millstone nuclear power plant in Connecticut to remain closed until management could demonstrate a "safety conscious work environment" after identified problems were allowed to continue without remedial action [34].

Commercial Aviation

The highly regarded ASRS (Aviation Safety Reporting System) has been copied by many individual airline information systems. Although much information is now collected, there still exist problems in evaluating and learning from it. The breadth and type of information acquired is much greater than the NRC reporting system described above. The sheer number of ASRS reports and the free form entry of the information make evaluation very difficult. There are few ways implemented to determine whether the report was accurate or evaluated the problem correctly. Subjective causal attribution and inconsistency in terminology and information included in the reports makes comparative analysis and categorization difficult and sometimes impossible.

Existing categorization schemes have also become inadequate as technology has changed, for example, with increased use of digital technology and computers in aircraft and ground operations. New categorizations are being implemented, but that creates problems when comparing data that used older categorization schemes.

Another problem arising from the goal to encourage use of the system is in the accuracy of the data. By filing an ASRS report, a limited form of indemnity against punishment is assured. Many of the reports are biased by personal protection considerations, as evidenced by the large percentage of the filings that report FAA regulation violations. For example, in a NASA Langley study of reported helicopter incidents in the ASRS over a nine-year period, nonadherence to FARs (Federal Aviation Regulations) was by far the largest category of reports. The predominance of FAR violations in the incident data may reflect the motivation of the ASRS reporters to obtain immunity from perceived or real violations of FARs and not necessarily the true percentages.

But with all these problems and limitations, most agree that the ASRS and similar industry reporting systems have been very successful and the information obtained extremely useful in enhancing safety. For example, reported unsafe airport

conditions have been corrected quickly and improvements in air traffic control and other types of procedures made on the basis of ASRS reports.

The success of the ASRS has led to the creation of other reporting systems in this industry. The Aviation Safety Action Program (ASAP) in the United States, for example, encourages air carrier and repair station personnel to voluntarily report safety information to be used to develop corrective actions for identified safety concerns. An ASAP involves a partnership between the FAA and the certified organization (called the *certificate holder*) and may also include a third party, such as the employees' labor organization. It provides a vehicle for employees of the ASAP participants to identify and report safety issues to management and to the FAA without fear that the FAA will use the reports accepted under the program to take legal enforcement action against them or the company or that companies will use the information to take disciplinary action against the employee.

Certificate holders may develop ASAP programs and submit them to the FAA for review and acceptance. Ordinarily, programs are developed for specific employee groups, such as members of the flightcrew, flight attendants, mechanics, or dispatchers. The FAA may also suggest, but not require, that a certificate holder develop an ASAP to resolve an identified safety problem.

When ASAP reports are submitted, an event review committee (ERC) reviews and analyzes them. The ERC usually includes a management representative from the certificate holder, a representative from the employee labor association (if applicable), and a specially trained FAA inspector. The ERC considers each ASAP report for acceptance or denial, and if accepted, analyzes the report to determine the necessary controls to put in place to respond to the identified problem.

Single ASAP reports can generate corrective actions and, in addition, analysis of aggregate ASAP data can also reveal trends that require action. Under an ASAP, safety issues are resolved through corrective action rather than through punishment or discipline.

To prevent abuse of the immunity provided by ASAP programs, reports are accepted only for inadvertent regulatory violations that do not appear to involve an intentional disregard for safety and events that do not appear to involve criminal activity, substance abuse, or intentional falsification.

Additional reporting programs provide for sharing data that is collected by airlines for their internal use. FOQA (Flight Operational Quality Assurance) is an example. Air carriers often instrument their aircraft with extensive flight data recording systems or use pilot generated checklists and reports for gathering information internally to improve operations and safety. FOQA provides a voluntary means for the airlines to share this information with other airlines and with the FAA

so that national trends can be monitored and the FAA can target its resources to address the most important operational risk issues.[1]

In contrast with the ASAP voluntary reporting of single events, FOQA programs allow the accumulation of accurate operational performance information covering all flights by multiple aircraft types such that single events or overall patterns of aircraft performance data can be identified and analyzed. Such aggregate data can determine trends specific to aircraft types, local flight path conditions, and overall flight performance trends for the commercial aircraft industry. FOQA data has been used to identify the need for changing air carrier operating procedures for specific aircraft fleets and for changing air traffic control practices at certain airports with unique traffic pattern limitations.

FOQA and other such voluntary reporting programs allow early identification of trends and changes in behavior (i.e., migration of systems toward states of increasing risk) before they lead to accidents. Follow-up is provided to ensure that unsafe conditions are effectively remediated by corrective actions.

A cornerstone of FOQA programs, once again, is the understanding that aggregate data provided to the FAA will be kept confidential and the identity of reporting personnel or airlines will remain anonymous. Data that could be used to identify flight crews are removed from the electronic record as part of the initial processing of the collected data. Air carrier FOQA programs, however, typically provide a gatekeeper who can securely retrieve identifying information for a limited amount of time, in order to enable follow-up requests for additional information from the specific flight crew associated with a FOQA event. The gatekeeper is typically a line captain designated by the air carrier's pilot association. FOQA programs usually involve agreements between pilot organizations and the carriers that define how the collected information can be used.

12.5 Using the Feedback

Once feedback is obtained, it needs to be used to update the controllers' process models and perhaps control algorithms. The feedback and its analysis may be passed to others in the control structure who need it.

Information must be provided in a form that people can learn from, apply to their daily jobs, and use throughout the system life cycle.

Various types of analysis may be performed by the controller on the feedback, such as trend analysis. If flaws in the system design or unsafe changes are detected, obviously actions are required to remedy the problems.

1. FOQA is voluntary in the United States but required in some countries.

In major accidents, precursors and warnings are almost always present but ignored or mishandled. While what appear to be warnings are sometimes simply a matter of hindsight, sometimes clear evidence does exist. In 1982, two years before the Bhopal accident, for example, an audit was performed that identified many of the deficiencies involved in the loss. The audit report noted such factors related to the later tragedy such as filter-cleaning operations without using slip blinds, leaking valves, and bad pressure gauges. The report recommended raising the capability of the water curtain and pointed out that the alarm at the flare tower was nonoperational and thus any leakage could go unnoticed for a long time. The report also noted that a number of hazardous conditions were known and allowed to persist for considerable amounts of time or inadequate precautions were taken against them. In addition, there was no follow-up to ensure that deficiencies were corrected. According to the Bhopal manager, all improvements called for in the report had been implemented, but obviously that was either untrue or the fixes were ineffective.

As with accidents and incidents, warning signs or anomalies also need to be analyzed using CAST. Because practice will naturally deviate from procedures, often for very good reasons, the gap between procedures and practice needs to be monitored and understood [50].

12.6 Education and Training

Everyone in the safety control structure, not just the lower-level controllers of the physical systems, must understand their roles and responsibilities with respect to safety and why the system—including the organizational aspects of the safety control structure—was designed the way it was.

People, both managers and operators, need to understand the risks they are taking in the decisions they make. Often bad decisions are made because the decision makers have an incorrect assessment of the risks being assumed, which has implications for training. Controllers must know exactly what to look for, not just be told to look for "weak signals," a common suggestion in the HRO literature. Before a bad outcome occurs, weak signals are simply noise; they take on the appearance of signals only in hindsight, when their relevance becomes obvious. Telling managers and operators to "be mindful of weak signals" simply creates a pretext for blame after a loss event occurs. Instead, the people involved need to be knowledgeable about the hazards associated with the operation of the system if we expect them to recognize the precursors to an accident. Knowledge turns unidentifiable weak signals into identifiable strong signals. People need to know what to look for.

Decision makers at all levels of the safety control structure also need to understand the risks they are taking in the decisions they make: Training should include

not just *what* but *why*. For good decision making about operational safety, decision makers must understand the system hazards and their responsibilities with respect to avoiding them. Understanding the safety rationale, that is, the "why," behind the system design will also have an impact on combating complacency and unintended changes leading to hazardous states. This rationale includes understanding why previous accidents occurred. The Columbia Accident Investigation Board was surprised at the number of NASA engineers in the Space Shuttle program who had never read the official *Challenger* accident report [74]. In contrast, everyone in the U.S. nuclear Navy has training about the *Thresher* loss every year.

Training should not be a one-time event for employees but should be continual throughout their employment, if only as a reminder of their responsibilities and the system hazards. Learning about recent events and trends can be a focus of this training.

Finally, assessing for training effectiveness, perhaps during regular audits, can assist in establishing an effective improvement and learning process.

With highly automated systems, an assumption is often made that less training is required. In fact, training requirements go up (not down) in automated systems, and they change their nature. Training needs to be more extensive and deeper when using automation. One of the reasons for this requirement is that human operators of highly automated systems not only need a model of the current process state and how it can change state but also a model of the automation and its operation, as discussed in chapter 8.

To control complex and highly automated systems safely, operators (controllers) need to learn more than just the procedures to follow: If we expect them to control and monitor the automation, they must also have an in-depth understanding of the controlled physical process and the logic used in any automated controllers they may be supervising. System controllers—at all levels—need to know:

- The system hazards and the reason behind safety-critical procedures and operational rules.
- The potential result of removing or overriding controls, changing prescribed procedures, and inattention to safety-critical features and operations: Past accidents and their causes should be reviewed and understood.
- How to interpret feedback: Training needs to include different combinations of alerts and sequences of events, not just single events.
- How to think flexibly when solving problems: Controllers need to be provided with the opportunity to practice problem solving.
- General strategies rather than specific responses: Controllers need to develop skills for dealing with unanticipated events.

- How to test hypotheses in an appropriate way: To update mental models, human controllers often use hypothesis testing to understand the system state better and update their process models. Such hypothesis testing is common with computers and automated systems where documentation is usually so poor and hard to use that experimentation is often the only way to understand the automation behavior and design. Such testing can, however, lead to losses. Designers need to provide operators with the ability to test hypotheses safely and controllers must be educated on how to do so.

Finally, as with any system, emergency procedures must be overlearned and continually practiced. Controllers must be provided with operating limits and specific actions to take in case they are exceeded. Requiring operators to make decisions under stress and without full information is simply another way to ensure that they will be blamed for the inevitable loss event, usually based on hindsight bias. Critical limits must be established and provided to the operators, and emergency procedures must be stated explicitly.

12.7 Creating an Operations Safety Management Plan

The operations safety management plan is used to guide operational control of safety. The plan describes the objectives of the operations safety program and how they will be achieved. It provides a baseline to evaluate compliance and progress. Like every other part of safety program, the plan will need buy-in and oversight.

The organization should have a template and documented expectations for operations safety management plans, but this template may need to be tailored for particular project requirements.

The information need not all be contained in one document, but there should be a central reference with pointers to where the information can be found. As is true for every other part of the safety control structure, the plan should include review procedures for the plan itself as well as how the plan will be updated and improved through feedback from experience.

Some things that might be included in the plan:

- General Considerations
  - Scope and objectives
  - Applicable standards (company, industry)
  - Documentation and reports
  - Review of plan and progress reporting procedures
- Safety Organization (safety control structure)
  - Personnel qualifications and duties

  - Staffing and manpower
  - Communication channels
  - Responsibility, authority, accountability (functional organization, organizational structure)
  - Information requirements (feedback requirements, process model, updating requirements)
  - Subcontractor responsibilities
  - Coordination
  - Working groups
  - System safety interfaces with other groups, such as maintenance and test, occupational safety, quality assurance, and so on.
- Procedures
  - Problem reporting (processes, follow-up)
  - Incident and accident investigation
    - Procedures
    - Staffing (participants)
    - Follow-up (tracing to hazard and risk analyses, communication)
  - Testing and audit program
    - Procedures
    - Scheduling
    - Review and follow-up
    - Metrics and trend analysis
    - Operational assumptions from hazard and risk analyses
  - Emergency and contingency planning and procedures
  - Management of change procedures
  - Training
  - Decision making, conflict resolution
- Schedule
  - Critical checkpoints and milestones
  - Start and completion dates for tasks, reports, reviews
  - Review procedures and participants
- Safety Information System
  - Hazard and risk analyses, hazard logs (controls, review and feedback procedures)

  - Hazard tracking and reporting system
  - Lessons learned
  - Safety data library (documentation and files)
  - Records retention policies
- Operations hazard analysis
  - Identified hazards
  - Mitigations for hazards
- Evaluation and planned use of feedback to keep the plan up-to-date and improve it over time

12.8 Applying STAMP to Occupational Safety

Occupational safety has, traditionally, not taken a systems approach but instead has focused on individuals and changing their behavior. In applying systems theory to occupational safety, more emphasis would be placed on understanding the impact of system design on behavior and would focus on changing the system rather than people. For example, vehicles used in large plants could be equipped with speed regulators rather than depending on humans to follow speed limits and then punishing them when they do not. The same design for safety principles presented in chapter 9 for human controllers apply to designing for occupational safety.

With the increasing complexity and automation of our plants, the line between occupational safety and engineering safety is blurring. By designing the system to be safe despite normal human error or judgment errors under competing work pressures, workers will be better protected against injury while fulfilling their job responsibilities.

For information about special quantity discounts, please email special_sales@mitpress.mit.edu

This book was set in Syntax and Times Roman by Toppan Best-set Premedia Limited. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Leveson, Nancy.
Engineering a safer world : systems thinking applied to safety / Nancy G. Leveson.
p. cm.—(Engineering systems)
Includes bibliographical references and index.
ISBN 978-0-262-01662-9 (hardcover : alk. paper)
1. Industrial safety. 2. System safety. I. Title.
T55.L466 2012
620.8′6—dc23
2011014046

10 9 8 7 6 5 4 3 2 1