7 Fundamentals

All the parts of the process described in the following chapters start from the same fundamental system engineering activities. These include defining, for the system involved, accidents or losses, hazards, safety requirements and constraints, and the safety control structure.

7.1 Defining Accidents and Unacceptable Losses

The first step in any safety effort involves agreeing on the types of accidents or losses to be considered.

In general, the definition of an accident comes from the customer and occasionally from the government for systems that are regulated by government agencies. Other sources might be user groups, insurance companies, professional societies, industry standards, and other stakeholders. If the company or group developing the system is free to build whatever they want, then considerations of liability and the cost of accidents will come into play.

Definitions of basic terms differ greatly among industries and engineering disciplines. A set of basic definitions is used in this book (see appendix A) that reflect common usage in System Safety. An *accident* is defined as:

Accident: An undesired or unplanned event that results in a loss, including loss of human life or human injury, property damage, environmental pollution, mission loss, etc.

An accident need not involve loss of life, but it does result in some loss that is unacceptable to the stakeholders. System Safety has always considered non-human losses, but for some reason, many other approaches to safety engineering have limited the definition of a loss to human death or injury. As an example of an inclusive definition, a spacecraft accident might include loss of the astronauts (if the spacecraft is manned), death or injury to support personnel or the public, non-accomplishment of the mission, major equipment damage (such as damage to launch

facilities), environmental pollution of planets, and so on. An accident definition used in the design of an explorer spacecraft to characterize the icy moon of a planet in the Earth's solar system, for example, was [151]:

A1. Humans or human assets on earth are killed or damaged.

A2. Humans or human assets off of the earth are killed or damaged.

A3. Organisms on any of the moons of the outer planet (if they exist) are killed or mutated by biological agents of Earth origin.

Rationale: Contamination of an icy outer planet moon with biological agents of Earth origin could have catastrophically adverse effects on any biological agents indigenous to the icy outer planet moon.

A4. The scientific data corresponding to the mission goals is not collected.

A5. The scientific data corresponding to the mission goals is rendered unusable (i.e., deleted or corrupted) before it can be fully investigated.

A6. Organisms of Earth origin are mistaken for organisms indigenous to any of the moons of the outer planet in future missions to study the outer planet's moon.

Rationale: Contamination of a moon of an outer planet with biological agents of Earth origin could lead to a situation in which a future mission discovers the biological agents and falsely concludes that they are indigenous to the moon of the outer planet.

A7. An incident during this mission directly causes another mission to fail to collect, return, or use the scientific data corresponding to its mission goals.

Rationale: It is possible for this mission to interfere with the completion of other missions through denying the other missions access to the space exploration infrastructure (for example, overuse of limited Deep Space Network[1] (DSN) resources, causing another mission to miss its launch window because of damage to the launch pad during this mission, etc.)

Prioritizing or assigning a level of severity to the identified losses may be useful when tradeoffs among goals are required in the design process. As an example, consider an industrial robot to service the thermal tiles on the Space Shuttle, which

1. The Deep Space Network is an international network of large antennas and communication facilities that supports interplanetary spacecraft missions and radio and radar astronomy observations for the exploration of the solar system and the universe. The network also supports some Earth-orbiting missions.

is used as an example in chapter 9. The goals for the robot are (1) to inspect the thermal tiles for damage caused during launch, reentry, and transport of a Space Shuttle and (2) to apply waterproofing chemicals to the thermal tiles.

Level 1:

Al-1: Loss of the orbiter and crew (e.g., inadequate thermal protection)

Al-2: Loss of life or serious injury in the processing facility

Level 2:

A2–1: Damage to the orbiter or to objects in the processing facility that results in the delay of a launch or in a loss of greater than x dollars

A2–2: Injury to humans requiring hospitalization or medical attention and leading to long-term or permanent physical effects

Level 3:

A3–1: Minor human injury (does not require medical attention or requires only minimal intervention and does not lead to long-term or permanent physical effects)

A3–2: Damage to orbiter that does not delay launch and results in a loss of less than x dollars

A3–3: Damage to objects in the processing facility (both on the floor or suspended) that does not result in delay of a launch or a loss of greater than x dollars

A3–4: Damage to the mobile robot

Assumption: It is assumed that there is a backup plan in place for servicing the orbiter thermal tiles in case the tile processing robot has a mechanical failure and that the same backup measures can be used in the event the robot is out of commission due to other reasons.

The customer may also have a safety policy that must be followed by the contractor or those designing the thermal tile servicing robot. As an example, the following is similar to a typical NASA safety policy:

General Safety Policy: All hazards related to human injury or damage to the orbiter must be eliminated or mitigated by the system design. A reasonable effort must be made to eliminate or mitigate hazards resulting at most in damage to the robot or objects in the work area. For any hazards that cannot be eliminated, the hazard analysis as well as the design features and development procedures, including any tradeoff studies, must be documented and presented to the customer for acceptance.

7.2 System Hazards

The term *hazard* has been used in different ways. For example, in aviation, a hazard is often used to denote something in the environment of the system, for example a mountain, that is in the path of the aircraft. In contrast, in System Safety, a hazard is defined as within the system being designed (or its relationship to an environmental object) and not just in its environment. For example, an aircraft flying too close to a mountain would be a hazard.

Hazard: A system state or set of conditions that, together with a particular set of worst-case environmental conditions, will lead to an accident (loss).

This definition requires some explanation. First, hazards may be defined in terms of conditions, as here, or in terms of events as long as one of these choices is used consistently. While there have been arguments about whether hazards are events or conditions, the distinction is irrelevant and either can be used. Figure 2.6 depicts the relationship between events and conditions: conditions lead to events which lead to conditions which lead to events. . . . The hazard for a chemical plant could be stated as the release of chemicals (an event) or chemicals in the atmosphere (a condition). The only difference is that events are limited in time while the conditions caused by the event persist over time until another event occurs that changes the prevailing conditions. For different purposes, one choice might be advantageous over the other.

Second, note that the word *failure* does not appear anywhere. Hazards are not identical to failures—failures can occur without resulting in a hazard and a hazard may occur without any precipitating failures. C. O. Miller, one of the founders of System Safety, cautioned that "distinguishing hazards from failures is implicit in understanding the difference between safety and reliability" [138].

Sometimes, hazards are defined as something that "has the potential to do harm" or that "can lead to an accident." The problem with this definition is that most every system state has the potential to do harm or can lead to an accident. An airplane that is in the air is in a hazardous state according to this definition, but there is little that the designer of an air traffic control system or an air transportation system, for example, can do about designing a system where the planes never leave the ground. For practical reasons, the definition should preclude states that the system must normally be in to accomplish the mission. By limiting the definition of hazard to states that the system should never be in (that is, closer to the accident or loss event), the designer has greater freedom and ability to design hazards out of the system. For air traffic control, the hazard would not be two planes in the air but two planes that violate minimum separation standards.

An accident is defined with respect to the environment of the system or component:

Hazard + Environmental Conditions ⇒ Accident (Loss)

As an example, a release of toxic chemicals or explosive energy will cause a loss only if there are people or structures in the vicinity. Weather conditions may affect whether a loss occurs in the case of a toxic release. If the appropriate environmental conditions do not exist, then there is no loss and, by definition, no accident. This type of non-loss event is commonly called an *incident.* When a hazard is defined as an event, then hazards and incidents are identical.

7.2.1 Drawing the System Boundaries

What constitutes a hazard, using the preceding definition, depends on where the boundaries of the system are drawn. A system is an abstraction, and the boundaries of the system can be drawn anywhere the person defining the system wants. Where the boundaries are drawn will determine which conditions are considered part of the hazard and which are considered part of the environment. Because this choice is arbitrary, the most useful way to define the boundaries, and thus the hazard, is to draw them to include the conditions related to the accident over which the system designer has some control. That is, if we expect designers to create systems that eliminate or control hazards and thus prevent accidents, then those hazards must be in their design space. This control requirement is the reason for distinguishing between hazards and accidents—accidents may involve aspects of the environment over which the system designer or operator has no control.

In addition, because of the recursive nature of the definition of a system—that is, a system at one level may be viewed as a subsystem of a larger system—higher-level systems will have control over the larger hazards. But once boundaries are drawn, system designers can be held responsible only for controlling the accident factors that they have the ability to control, including those that have been passed to them from system designers above them as component safety requirements to ensure the encompassing system hazards are eliminated or controlled.

Consider the chemical plant example. While the hazard could be defined as death or injury of residents around the plant (the loss event), there may be many factors involved in such a loss that are beyond the control of the plant designers and operators. One example is the atmospheric conditions at the time of the release, such as velocity and direction of the wind. Other factors in a potential accident or loss are the location of humans around the plant and community emergency preparedness, both of which may be under the control of the local or state government. The designers of the chemical plant have a responsibility to provide the information necessary for the design and operation of appropriate emergency preparedness equipment and procedures, but their primary design responsibility is the part of a potential

accident that is under their design control, namely the design of the plant to prevent release of toxic chemicals.

In fact, the environmental conditions contributing to a loss event may change over time: potentially dangerous plants may be originally located far from population centers, for example, but over time human populations tend to encroach on such plants in order to live close to their jobs or because land may be cheaper in remote areas or near smelly plants. The chemical plant designer usually has no design control over these conditions so it is most convenient to draw the system boundaries around the plant and define the hazard as uncontrolled release of chemicals from the plant. If the larger sociotechnical system is being designed or analyzed for safety, which it should be, the number of potential hazards and actions to prevent them increases. Examples include controlling the location of plants or land use near them through local zoning laws, and providing for emergency evacuation and medical treatment.

Each component of the sociotechnical system may have different aspects of an accident under its control and is responsible for different parts of the accident process, that is, different hazards and safety constraints. In addition, several components may have responsibilities related to the same hazards. The designers of the chemical plant and relevant government regulatory agencies, for example, may both be concerned with plant design features potentially leading to inadvertent toxic chemical release. The government role, however, may be restricted to design and construction approvals and inspection processes, while the plant designers have basic design creation responsibilities.

As another example of the relationship between hazards and system boundaries, consider the air traffic control system. If an accident is defined as a collision between aircraft, then the appropriate hazard is the violation of minimum separation between aircraft. The designer of an airborne collision avoidance system or a more general air traffic control system theoretically has control over the separation between aircraft, but may not have control over other factors that determine whether two aircraft that get close together actually collide, such as visibility and weather conditions or the state of mind or attentiveness of the pilots. These are under the control of other system components such as air traffic control in directing aircraft away from poor weather conditions or the control of other air transportation system components in the selection and training of pilots, design of aircraft, and so on.

Although individual designers and system components are responsible for controlling only the hazards in their design space, a larger system safety engineering effort preceding component design will increase overall system safety while decreasing the effort, cost, and tradeoffs involved in component safety engineering. By considering the larger sociotechnical system and not just the individual technical components, the most cost-effective way to eliminate or control hazards can be

identified. If only part of the larger system is considered, the compromises required to eliminate or control the system hazard in one piece of the overall system design may be much greater than would be necessary if other parts of the overall system were considered. For example, a particular hazard associated with launching a spacecraft might be controllable by the spacecraft design, by the physical launch infrastructure, by launch procedures, by the launch control system, or by a combination of these. If only the spacecraft design is considered in the drawing of system boundaries and the hazard identification process, hazard control may require more tradeoffs than if the hazard is partially or completely eliminated or controlled by design features in other parts of the system.

All that is being suggested here is that top-down system engineering is critical for engineering safety into complex systems. In addition, when a new component is introduced into an existing system, such as the introduction of a collision avoidance system in the aircraft, the impact of the addition on the safety of the aircraft itself as well as the safety of air traffic control and the larger air transportation system safety needs to be considered.

Another case is when a set of systems that already exist are combined to create a new system.[2] While the individual systems may have been designed to be safe within the system for which they were originally created, the safety constraints enforced in the components may not adequately control hazards in the combined system or may not control hazards that involve interactions among new and old system components.

The reason for this discussion is to explain why the definition of the hazards associated with a system is an arbitrary but important step in assuring system safety and why a system engineering effort that considers the larger sociotechnical system is necessary. One of the first steps in designing a system, after the definition of an accident or loss and the drawing of boundaries around the subsystems, is to identify the hazards that need to be eliminated or controlled by the designers of that system or subsystem.

7.2.2 Identifying the High-Level System Hazards

For practical reasons, a small set of high-level system hazards should be identified first. Starting with too large a list at the beginning, usually caused by including refinements and causes of the high-level hazards in the list, often leads to a disorganized and incomplete hazard identification and analysis process. Even the most complex system seldom has more than a dozen high-level hazards, and usually less than this.

2. Sometimes called a *system* of *systems*, although all systems are subsystems of larger systems.

Hazards are identified using the definition of an accident or loss along with additional safety criteria that may be imposed by regulatory or industry associations and practices. For example, the hazards associated with the outer planets explorer accident definition in section 7.1 might be defined as [151]:

H1. Inability of the mission to collect data [A4]

H2. Inability of the mission to return collected data [A5]

H3. Inability of the mission scientific investigators to use the returned data [A5]

H4. Contamination of the outer planet moon with biological agents of Earth origin on mission hardware [A6]

H5. Exposure of Earth life or human assets on Earth to toxic, radioactive, or energetic elements of the mission hardware [A1]

H6. Exposure of Earth life or human assets off Earth to toxic, radioactive, or energetic elements of the mission hardware [A2]

H7. Inability of other space exploration missions to use the shared space exploration infrastructure to collect, return, or use data [A7]

The numbers in the square brackets identify the accidents related to each of these hazards.

The high-level system hazards that might be derived from the accidents defined for the NASA thermal tile processing robot in section 7.1 might be:

H1. Violation of minimum separation between mobile base and objects (including orbiter and humans)

H2. Unstable robot base

H3. Movement of the robot base or manipulator arm causing injury to humans or damage to the orbiter

H4. Damage to the robot

H5. Fire or explosion

H6. Contact of human with DMES waterproofing chemical

H7. Inadequate thermal protection

During the design process, these high-level hazards will be refined as the design alternatives are considered. Chapter 9 provides more information about the refinement process and an example.

Aircraft collision control provides a more complex example. As noted earlier, the relevant accident is a collision between two airborne aircraft and the overall system hazard to be avoided is violation of minimum physical separation (distance) between aircraft.

One (but only one) of the controls used to avoid this type of accident is an airborne collision avoidance system like TCAS (Traffic alert and Collision Avoidance System), which is now required on most commercial aircraft. While the goal of TCAS is increased safety, TCAS itself introduces new hazards associated with its use. Some hazards that were considered during the design of TCAS are:

H1. TCAS causes or contributes to a near midair collision (NMAC), defined as a pair of controlled aircraft violating minimum separation standards.

H2. TCAS causes or contributes to a controlled maneuver into the ground.

H3. TCAS causes or contributes to the pilot losing control over the aircraft.

H4. TCAS interferes with other safety-related aircraft systems.

H5. TCAS interferes with the ground-based Air Traffic Control system (e.g., transponder transmissions to the ground or radar or radio services).

H6. TCAS interferes with an ATC advisory that is safety-related (e.g., avoiding a restricted area or adverse weather conditions).

Ground-based air traffic control also plays an important role in collision avoidance, although it has responsibility for a larger and different set of hazards:

H1. Controlled aircraft violate minimum separation standards (NMAC).

H2. An airborne controlled aircraft enters an unsafe atmospheric region.

H3. A controlled airborne aircraft enters restricted airspace without authorization.

H4. A controlled airborne aircraft gets too close to a fixed obstacle other than a safe point of touchdown on assigned runway (known as controlled flight into terrain or CFIT).

H5. A controlled airborne aircraft and an intruder in controlled airspace violate minimum separation.

H6. Loss of controlled flight or loss of airframe integrity.

H7. An aircraft on the ground comes too close to moving objects or collides with stationary objects or leaves the paved area.

H8. An aircraft enters a runway for which it does not have a clearance (called runway incursion).

Unsafe behavior (hazards) at the system level can be mapped into hazardous behaviors at the component or subsystem level. Note, however, that the reverse (bottom-up) process is not possible, that is, it is not possible to identify the system-level hazards by looking only at individual component behavior. Safety is a system property, not a component property. Consider an automated door system. One

reasonable hazard when considering the door alone is the door closing on someone. The associated safety constraint is that the door must not close on anyone in the doorway. This hazard is relevant if the door system is used in any environment. If the door is in a building, another important hazard is not being able to get out of a dangerous environment, for example, if the building is on fire. Therefore, a reasonable design constraint would be that the door opens whenever a door open request is received. But if the door is used on a moving train, an additional hazard must be considered, namely, the door opening while the train is moving and between stations. In a moving train, different safety design constraints would apply compared to an automated door system in a building. Hazard identification is a top-down process that must consider the encompassing system and its hazards and potential accidents.

Let's assume that the automated door system is part of a train control system. The system-level train hazards related to train doors include a person being hit by closing doors, someone falling from a moving train or from a stationary train that is not properly aligned with a station platform, and passengers and staff being unable to escape from a dangerous environment in the train compartment. Tracing these system hazards into the related hazardous behavior of the automated door component of the train results in the following hazards:

1. Door is open when the train starts.
2. Door opens while train is in motion.
3. Door opens while not properly aligned with station platform.
4. Door closes while someone is in the doorway.
5. Door that closes on an obstruction does not reopen or reopened door does not reclose.
6. Doors cannot be opened for emergency evacuation between stations.

The designers of the train door controller would design to control these hazards. Note that constraints 3 and 6 are conflicting, and the designers will have to reconcile such conflicts. In general, attempts should first be made to eliminate hazards at the system level. If they cannot be eliminated or adequately controlled at the system level, then they must be refined into hazards to be handled by the system components.

Unfortunately, no tools exist for identifying hazards. It takes domain expertise and depends on subjective evaluation by those constructing the system. Chapter 13 in *Safeware* provides some common heuristics that may be helpful in the process. The good news is that identifying hazards is usually not a difficult process. The later steps in the hazard analysis process are where most of the mistakes and effort occurs.

There is also no right or wrong set of hazards, only a set that the system stakeholders agree is important to avoid. Some government agencies have mandated the hazards they want considered for the systems they regulate or certify. For example, the U.S. Department of Defense requires that producers of nuclear weapons consider four hazards:

1. Weapons involved in accident or incidents, or jettisoned weapons, produce a nuclear yield.
2. Nuclear weapons are deliberately prearmed, armed, launched, fired, or released without execution of emergency war orders or without being directed to do so by a competent authority.
3. Nuclear weapons are inadvertently prearmed, armed, launched, fired, or released.
4. Inadequate security is applied to nuclear weapons.

Sometimes user or professional associations define the hazards for the systems they use and that they want developers to eliminate or control. In most systems, however, the hazards to be considered are up to the developer and their customer(s).

7.3 System Safety Requirements and Constraints

After the system and component hazards have been identified, the next major goal is to specify the system-level safety requirements and design constraints necessary to prevent the hazards from occurring. These constraints will be used to guide the system design and tradeoff analyses.

The system-level constraints are refined and allocated to each component during the system engineering decomposition process. The process then iterates over the individual components as they are refined (and perhaps further decomposed) and as design decisions are made.

Figure 7.1 shows an example of the design constraints that might be generated from the automated train door hazards. Again, note that the third constraint potentially conflicts with the last one and the resolution of this conflict will be an important part of the system design process. Identifying these types of conflicts early in the design process will lead to better solutions. Choices may be more limited later on when it may not be possible or practical to change the early decisions.

As the design process progresses and design decisions are made, the safety requirements and constraints are further refined and expanded. For example, a safety constraint on TCAS is that it must not interfere with the ground-based air traffic control system. Later in the process, this constraint will be refined into more detailed constraints on the ways this interference might occur. Examples include

HAZARD		SAFETY DESIGN CONSTRAINT
1	Train starts with door open	Train must not be capable of moving with any door open
2	Door opens while train is in motion	Doors must remain closed while train is in motion
3	Door opens while improperly aligned with station platform	Door must be capable of opening only after train is stopped and properly aligned with platform unless emergency exists (see hazard 6 below)
4	Door closes while someone is in the doorway	Door areas must be clear before door closing begins
5	Door that closes on an obstruction does not reopen or reopened door does not reclose	An obstructed door must reopen to permit removal of obstruction and then automatically reclose
6	Doors cannot be opened for emergency evacuation	Means must be provided to open doors anywhere when the train is stopped for emergency evacuation

Figure 7.1
Design constraints for train door hazards.

constraints on TCAS design to limit interference with ground-based surveillance radar, with distance-measuring equipment channels, and with radio services. Additional constraints include how TCAS can process and transmit information (see chapter 10).

Figure 7.2 shows the high-level requirements and constraints for some of the air traffic control hazards identified above. Comparing the ATC high-level constraints with the TCAS high-level constraints (figure 7.3) is instructive. Ground-based air traffic control has additional requirements and constraints related to aspects of the collision problem that TCAS cannot handle alone, as well as other hazards and potential aircraft accidents that it must control.

Some constraints on the two system components (ATC and TCAS) are closely related, such as the requirement to provide advisories that maintain safe separation between aircraft. This example of overlapping control raises important concerns about potential conflicts and coordination problems that need to be resolved. As noted in section 4.5, accidents often occur in the boundary areas between controllers and when multiple controllers control the same process. The inadequate resolution of the conflict between multiple controller responsibilities for aircraft separation contributed to the collision of two aircraft over the town of Überlingen (Germany)

HAZARD		SAFETY DESIGN CONSTRAINTS
1	A pair of controlled aircraft violate minimum separation standards	a. ATC must provide advisories that maintain safe separation between aircraft b. ATC must provide conflict alerts
2	A controlled aircraft enters an unsafe atmospheric region (icing conditions, windsher areas, thunderstorm cells)	a. ATC must not issue advisories that direct aircraft into areas with unsafe atmospheric conditions b. ATC must provide weather advisories and alerts to flight crews c. ATC must warn aircraft that enter an unsafe atmospheric region
3	A controlled aircraft enters restricted airspace without authorization	a. ATC must not issue advisories that direct an aircraft into restricted airspace unless avoiding a greater hazard b. ATC must provide timely warnings to aircraft to prevent their incursion into restricted airspace
4	A controlled aircraft gets too close to a fixed obstacle or terrain other than a safe point of touchdown on its assigned runway	ATC must provide advisories that maintain safe separation between aircraft and terrain or physical obstacles
5	A controlled aircraft and an intruder in controlled airspace violate minimum separation standards	ATC must provide alerts and advisories to avoid intruders if at all possible
6	Loss of controlled flight or loss of airframe integrity	a. ATC must not issue advisories outside the the safe performance envelope of the aircraft b. ATC advisories must not distract or disrupt the crew from maintaining safety of flight c. ATC must not issue advisories that the pilot or aircraft cannot fly or that degrade the continued safe flight of the aircraft d. ATC must not provide advisories that cause an aircraft to fall below the standard glidepath or intersect it at the wrong place

Figure 7.2
High-level requirements and design constraints for air traffic control.

	HAZARD	SAFETY DESIGN CONSTRAINT
1	TCAS causes or contributes to an NMAC (near midair collision)	a. TCAS must provide effective warnings and appropriate collision avoidace guidance on potentially dangerous threats and must provide them within an appropriate time limit b. TCAS must not cause or contribute to an NMAC that would not have occurred had the aircraft not carried TCAS)
2	TCAS causes or contributes to a controlled maneuver into the ground	TCAS must not cause or contribute to controlled flight into terrain
3	TCAS causes or contributes to a pilot losing control over the aircraft	a. TCAS must not disrupt the pilot and ATC operations during critical phases of flight nor disrupt aircraft operation b. TCAS must operate with an acceptably low level of unwanted or nuisance alarms. The unwanted alarm rate must be sufficiently low to pose no safety of flight hazard nor adversely affect the workload in the cockpit c. TCAS must not issue advisories outside the safe performance envelope of the aircraft and degrade the continued safe flight of the aircraft (e.g., reduce stall margins or result in stall warnings)
4	TCAS interferes with other safety-related aircraft systems	TCAS must not interfere with other safety-related aircraft systems or contribute to non-separation-related hazards
5	TCAS interferes with the ground ATC system (e.g., transponder, radar. or radar transmissions)	TCAS must not interfere with the ground ATC system or other aircraft transmissions to the ground ATC system
6	TCAS interferes with a safety-related ATC advisory (e.g., avoiding a restricted area or adverse weather conditions)	TCAS must generate advisories that require as little deviation as possible from ATC clearances

Figure 7.3
High-level design constraints for TCAS.

in July 2002 when TCAS and the ground air traffic controller provided conflicting advisories to the pilots. Potentially conflicting responsibilities must be carefully handled in system design and operations and identifying such conflicts are part of the new hazard analysis technique described in chapter 8.

Hazards related to the interaction among components, for example the interaction between attempts by air traffic control and by TCAS to prevent collisions, need to be handled in the safety control structure design, perhaps by mandating how the pilot is to select between conflicting advisories. There may be considerations in handling these hazards in the subsystem design that will impact the behavior of multiple subsystems and therefore must be resolved at a higher level and passed to them as constraints on their behavior.

7.4 The Safety Control Structure

The safety requirements and constraints on the physical system design shown in section 7.3 act as input to the standard system engineering process and must be incorporated into the physical system design and safety control structure. An example of how they are used is provided in chapter 10.

Additional system safety requirements and constraints, including those on operations and maintenance or upgrades will be used in the design of the safety control structure at the organizational and social system levels above the physical system. There is no one correct safety control structure: what is practical and effective will depend greatly on cultural and other factors. Some general principles that apply to all safety control structures are described in chapter 13. These principles need to be combined with specific system safety requirements and constraints for the particular system involved to design the control structure.

The process for engineering social systems is very similar to the regular system engineering process and starts, like any system engineering project, with identifying system requirements and constraints. The responsibility for implementing each requirement needs to be assigned to the components of the control structure, along with requisite authority and accountability, as in any management system; controls must be designed to ensure that the responsibilities can be carried out; and feedback loops created to assist the controller in maintaining accurate process models.

7.4.1 The Safety Control Structure for a Technical System

An example from the world of space exploration is used in this section, but many of the same requirements and constraints could easily be adapted for other types of technical system development and operations.

The requirements in this example were generated to perform a programmatic risk assessment of a new NASA management structure called Independent

Technical Authority (ITA) recommended in the report of the Columbia Accident Investigation Board. The risk analysis itself is described in the chapter on the new hazard analysis technique called STPA (chapter 8). But the first step in the safety or risk analysis is the same as for technical systems: to identify the system hazards to be avoided, to generate a set of requirements for the new management structure, and to design the control structure.

The new safety control structure for the NASA manned space program was introduced to improve the flawed engineering and management decision making leading to the Columbia loss. The hazard to be eliminated or mitigated was:

System Hazard: Poor engineering and management decision making leading to a loss.

Four high-level system safety requirements and constraints for preventing the hazard were identified and then refined into more specific requirements and constraints.

1. Safety considerations must be first and foremost in technical decision making.
   a. State-of-the art safety standards and requirements for NASA missions must be established, implemented, enforced, and maintained that protect the astronauts, the workforce, and the public.
   b. Safety-related technical decision making must be independent from programmatic considerations, including cost and schedule.
   c. Safety-related decision making must be based on correct, complete, and up-to-date information.
   d. Overall (final) decision making must include transparent and explicit consideration of both safety and programmatic concerns.
   e. The Agency must provide for effective assessment and improvement in safety-related decision making.
2. Safety-related technical decision making must be done by eminently qualified experts, with broad participation of the full workforce.
   a. Technical decision making must be credible (executed using credible personnel, technical requirements, and decision-making tools) .
   b. Technical decision making must be clear and unambiguous with respect to authority, responsibility, and accountability.
   c. All safety-related technical decisions, before being implemented by the Program, must have the approval of the technical decision maker assigned responsibility for that class of decisions.

   d. Mechanisms and processes must be created that allow and encourage all employees and contractors to contribute to safety-related decision making.
3. Safety analyses must be available and used starting in the early acquisition, requirements development, and design processes and continuing through the system life cycle.
   a. High-quality system hazard analyses must be created.
   b. Personnel must have the capability to produce high-quality safety analyses.
   c. Engineers and managers must be trained to use the results of hazard analyses in their decision making.
   d. Adequate resources must be applied to the hazard analysis process.
   e. Hazard analysis results must be communicated in a timely manner to those who need them. A communication structure must be established that includes contractors and allows communication downward, upward, and sideways (e.g., among those building subsystems).
   f. Hazard analyses must be elaborated (refined and extended) and updated as the design evolves and test experience is acquired.
   g. During operations, hazard logs must be maintained and used as experience is acquired. All in-flight anomalies must be evaluated for their potential to contribute to hazards.
4. The Agency must provide avenues for the full expression of technical conscience (for safety-related technical concerns) and provide a process for full and adequate resolution of technical conflicts as well as conflicts between programmatic and technical concerns.
   a. Communication channels, resolution processes, adjudication procedures must be created to handle expressions of technical conscience.
   b. Appeals channels must be established to surface complaints and concerns about aspects of the safety-related decision making and technical conscience structures that are not functioning appropriately.

Where do these requirements and constraints come from? Many of them are based on fundamental safety-related development, operations and management principles identified in various chapters of this book, particularly chapters 12 and 13. Others are based on experience, such as the causal factors identified in the Columbia and Challenger accident reports or other critiques of the NASA safety culture and of NASA safety management. The requirements listed obviously reflect the advanced technology and engineering domain of NASA and the space program that was the focus of the ITA program along with some of the unique aspects of the NASA

culture. Other industries will have their own requirements. An example for the pharmaceutical industry is shown in the next section of this chapter.

There is unlikely to be a universal set of requirements that holds for every safety control structure beyond a small set of requirements too general to be very useful in a risk analysis. Each organization needs to determine what its particular safety goals are and the system requirements and constraints that are likely to ensure that it reaches them.

Clearly buy-in and approval of the safety goals and requirements by the stakeholders, such as management and the broader workforce as well as anyone overseeing the group being analyzed, such as a regulatory agency, is important when designing and analyzing a safety control structure.

Independent Technical Authority is a safety control structure used in the nuclear Navy SUBSAFE program described in chapter 14. In this structure, safety-related decision making is taken out of the hands of the program manager and assigned to a Technical Authority. In the original NASA implementation, the technical authority rested in the NASA Chief Engineer, but changes have since been made. The overall safety control structure for the original NASA ITA is shown in figure 7.4.[3]

For each component of the structure, information must be determined about its overall role, responsibilities, controls, process model requirements, coordination and communication requirements, contextual (environmental and behavior-shaping) factors that might bear on the component's ability to fulfill its responsibilities, and inputs and outputs to other components in the control structure. The responsibilities are shown in figure 7.5. A risk analysis on ITA and the safety control structure is described in chapter 8.

7.4.2 Safety Control Structures in Social Systems

Social system safety control structures often are not designed but evolve over time. They can, however, be analyzed for inherent risk and redesigned or "reengineered" to prevent accidents or to eliminate or control past causes of losses as determined in an accident analysis.

The reengineering process starts with the definition of the hazards to be eliminated or mitigated, system requirements and constraints necessary to increase safety, and the design of the current safety-control structure. Analysis can then be used to drive the redesign of the safety controls. But once again, just like every system that has been described so far in this chapter, the process starts by identifying the hazards

3. The control structure was later changed to have ITA under the control of the NASA center directors rather than the NASA chief engineer; therefore, this control structure does not reflect the actual implementation of ITA at NASA, but it was the design at the time of the hazard analysis described in chapter 8.

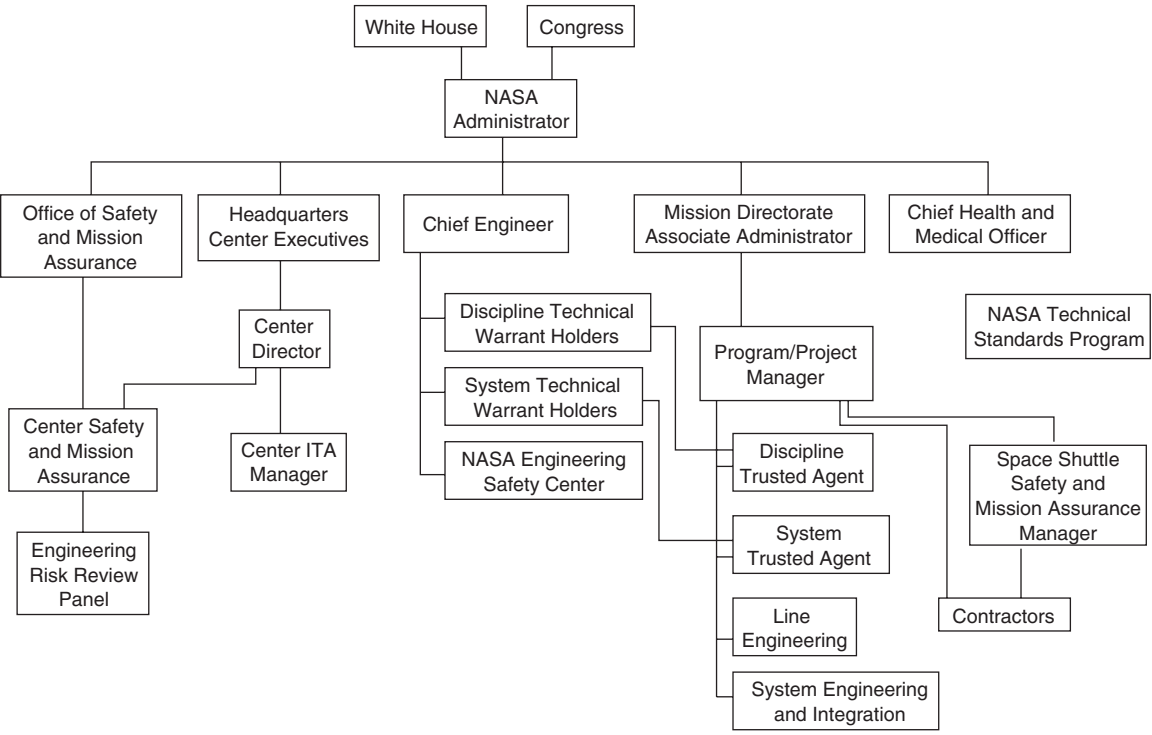

Figure 7.4
The NASA safety control structure under the original ITA design.

and safety requirements and constraints derived from them. The process is illustrated using drug safety.

Dozens of books have been written about the problems in the pharmaceutical industry. Everyone appears to have good intentions and are simply striving to optimize their performance within the existing incentive structure. The result is that the system has evolved to the point where each group's individual best interests do not necessarily add up to or are not aligned with the best interests of society as a whole. A safety control structure exists, but does not necessarily provide adequate satisfaction of the system-level goals, as opposed to the individual component goals.

This problem can be viewed as a classic system engineering problem: optimizing each component does not necessarily add up to a system optimum. Consider the air transportation system, as noted earlier. When each aircraft tries to optimize its path from its departure point to its destination, the overall system throughput may not be optimized when they all arrive in a popular hub at the same time. One goal of the air traffic control system is to control individual aircraft movement in order to

Executive Branch
- Appointment of NASA Administrator
- Setting of high-level goals and vision for NASA
- Creation of a draft budget appropriation for NASA

Congress
- Approval of NASA Administrator appointment
- NASA budget allocation
- Legislation affecting NASA operations

NASA Administrator
- Appointment of Chief Engineer (ITA) and head of Office of Safety and Mission Assurance
- Providing funding and authority to Chief Engineer to execute the Independent Technical Authority
- Demonstraion of commitment to safety over programmatic concerns through concrete actions
- Providing the directives and procedural requirements that define the ITA program
- Adjudication of differences between the Mission Directorate Associate Administrators and the Chief Engineer (ITA)

Chief Engineer
- Implementing ITA
- Effectiveness of the ITA program
- Communication channels with and among Warrant Holders
- Communication of decisions and lessons learned
- Establishment, monitoring, and approval of technical requirements, products, and policy and all changes, variances and waivers to the requirements
- Safety, risk, and trend analysis
- Independent assessment of flight (launch) readiness
- Conflict Resolution
- Developing a Technical Conscience throughout the engineering community

System Technical Warrant Holder
- Establishment and maintenance of technical policy, technical standards, requirements, and processes for a particular system or systems
- Technical product compliance with requirements, specifications, standards
- Primary interace between system and ITA (Chief Engineer)
- Assist Discipline Technical Warrant Holder in access to data, rationale, and other experts
- Production, quality, and use of FMEA/SIL, trending analysis, hazard and risk analyses
- Timely, day-to-day technical positions on issues pertaining to safe and reliable operations
- Establishing appropriate communication channels and networks
- Succession Planning
- Documentation of all methodologies, actions or closures, and decisions
- Sustaining the Agency knowledge base through communication of decisions and lessons learned
- Assessment of launch readiness from the standpoint of safe and reliable flight and operations
- Budget and resource requirements definition
- Maintaining competence
- Leading the technical conscience for the warranted system(s)

Figure 7.5
The responsibilities of the components in the NASA ITA safety control structure.

Discipline Technical Warrant Holder
- Interface to specialized knowledge within the Agency
- Assistance to System Technical Warrant Holders in carrying out their responsibilities
- Ownership of technical specifications and standards for warranted discipline (including system safety standards)
- Sustaining the Agency knowledge base in the warranted discipline
- Sustaining the general health of the warranted discipline throughout the Agency
- Succession Planning
- Leading the technical conscience for the warranted discipline
- Budget and resource requirements definition

Trusted Agents
- Screening: evaluate all changes and variances and perform all functions requested by Technical Warrant Holders
- Conducting daily business for System Technical Warrant Holder (represent on boards, meetings, committees)
- Providing information to Technical Warrant Holders about specific projects (e.g., safety analyses)

In-Line Engineers
- Provide unbiased technical positions to warrant holders, safety and mission assurance, trusted agents, and programs and projects
- Conduct system safety engineering (analyses and incorporation of results into design, development, and operations
- Evaluate contractor-produced analyses and incorporation of results into contractor products
- Act as the technical conscience of the Agency

Chief Safety and Mission Assurance Officer (OSMA)
- Leadership, policy direction, functional oversight, and coordination of assurance activities across the Agency
- Assurance of safety and reliability on programs and projects
- Incident and accident investigation

Center Safety and Mission Assurance (S&MA)
- Assure compliance with all requirements, standards, directives, policies, and procedures
- Perform quality (reliability and safety) assessments
- Participate in reviews
- Intervention in any activity to avoid an unnecessary safety risk
- Recommend a Safety, Reliability, and Quality Assurance plan for each project
- Chair Engineering Risk Review Panels at each Space Operations Center

Lead Engineering Risk Review Panel Manager and Panels
- Conduct formal safety reviews of accepted and controlled hazards
- Oversee and resolve integrated hazards
- Assure compliance with requirements, accuracy of all data and hazard analyses, and proper classification of hazards

Space Shuttle Program Safety and Mission Assurnace Manager
- Assure compliance with requirements in activities of prime contractors and technical support personnel from the NASA Centers
- Integrate and provide guidance for safety, reliability, and quality engineering activities performed by Space Operations Centers

Figure 7.5
(Continued)

Program/Project Managers
- Communication of ITA understanding through program or project team
- Prioritization safety over programmatic concerns among those reporting to him or her
- Support of Trusted Agents
- Provision of complete and timely data to Technical Warrant Holder
- Compliance with System Technical Warrant Holders decisions

System Engineering and Integration Office
- Integrated hazard analyses and anomaly investigation at system level
- Communication of system-level, safety-related requirements and constraints to and from contractors
- Update hazard analyses and maintain hazard logs during test and operations

Contractors
- Production and use of hazard analyses in their designs
- Communication of hazard information to NASA System Engineering and Integration

Center Director
- Practice of technical conscience at the Center
- Preservation of ITA financial and managerial independence at the Center
- Support of ITA activities at the Center
- Support of safety activities at the Center

Center ITA Manager
- Administrative support for Technical Warrant Holders

NASA Engineering and Safety Center (NESC)
- In-depth technical reviews, assessments, and analyses of high-risk projects
- Selected mishap investigations
- In-depth system engineering analyses

Headquarters Center Executives
- Alignment of Center's organization and processes to support and maintain independence of ITA
- Monitoring of ITA and expression and resolution of technical conscience at their Center
- Oversight of safety and mission assurance at their Center

Mission Directorate Associate Administrators
- Leadership and accountability for all engineering and technical work for their mission
- Alignment of financial, personnel, and engineering infrastructure with ITA
- Resolution of differences between warrant holders and program or project managers

NASA Technical Standards Program
- Coordination of standards activities with ITA

Figure 7.5
(Continued)

optimize overall system throughput while trying to allow as much flexibility as possible for the individual aircraft and airlines to achieve their goals. The air traffic control system and the rules of operation of the air transportation system resolve conflicting goals when public safety is at stake. Each airline might want its own aircraft to land as quickly as possible, but the air traffic controllers ensure adequate spacing between aircraft to preserve safety margins. These same principles can be applied to non-engineered systems.

The ultimate goal is to determine how to reengineer or redesign the overall pharmaceutical safety control structure in a way that aligns incentives for the greater good of society. A well-designed system would make it easier for all stakeholders to do the right thing, both scientifically and ethically, while achieving their own goals as much as possible. By providing the decision makers with information about ways to achieve the overall system objectives and the tradeoffs involved, better decision making can result.

While system engineering is applicable to pharmaceutical (and more generally medical) safety and risk management, there are important differences from the classic engineering problem that require changes to the traditional system safety approaches. In most technical systems, managing risk is simpler because not doing something (e.g., not inadvertently launching the missile) is usually safe and the problem revolves around preventing the hazardous event (inadvertent launch): a risk/no risk situation. The traditional engineering approach identifies and evaluates the costs and potential effectiveness of different ways to eliminate or control the hazards involved in the operational system. Tradeoffs require comparing the costs of various solutions, including costs that involve reduction in desirable system functions or system reliability.

The problem in pharmaceutical safety is different: there is risk in prescribing a potentially unsafe drug, but there is also risk in not prescribing the drug (the patient dies from their medical condition): a risk/risk situation. The risks and benefits conflict in ways that greatly increase the complexity of decision making and the information needed to make decisions. New, more powerful system engineering techniques are required to deal with risk/risk decisions.

Once again, the basic goals, hazards, and safety requirements must first be identified [43].

System Goal: *To provide safe and effective pharmaceuticals to enhance the long-term health of the population.*

Important loss events (accidents) we are trying to avoid are:

1. Patients get a drug treatment that negatively impacts their health.
2. Patients do not get the treatment they need.

Three system hazards can be identified that are related to these loss events:

H1: The public is exposed to an unsafe drug.
1. The drug is released with a label that does not correctly specify the conditions for its safe use.
2. An approved drug is found to be unsafe and appropriate responses are not taken (warnings, withdrawals from the market, etc.)
3. Patients are subjected to unacceptable risk during clinical trials.

H2: Drugs are taken unsafely.
1. The wrong drug is prescribed for the indication.
2. The pharmacist provides a different medication than was prescribed.
3. Drugs are taken in an unsafe combination.
4. Drugs are not taken according to directions (dosage, timing).

H3: Patients do not get an effective treatment they require.
1. Safe and effective drugs are not developed, are not approved for use, or are withdrawn from the market.
2. Safe and effective drugs are not affordable by those who need them.
3. Unnecessary delays are introduced into development and marketing.
4. Physicians do not prescribe needed drugs or patients have no access to those who could provide the drugs to them.
5. Patients stop taking a prescribed drug due to perceived ineffectiveness or intolerable side effects.

From these hazards, a set of system requirements can be derived to prevent them:

1. Pharmaceutical products are developed to enhance long-term health.
   a. Continuous appropriate incentives exist to develop and market needed drugs.
   b. The scientific knowledge and technology needed to develop new drugs and optimize their use is available.
2. Drugs on the market are adequately safe and effective.
   a. Drugs are subjected to effective and timely safety testing.
   b. New drugs are approved by the FDA based upon a validated and reproducible decision-making process.
   c. The labels attached to drugs provide correct information about safety and efficacy.

d. Drugs are manufactured according to good manufacturing practices.

e. Marketed drugs are monitored for adverse events, side effects, and potential negative interactions. Long-term studies after approval are conducted to detect long-term effects and effects on subpopulations not in the original study.

f. New information about potential safety risk is reviewed by an independent advisory board. Marketed drugs found to be unsafe after they are approved are removed, recalled, restricted, or appropriate risk/benefit information is provided.

3. Patients get and use the drugs they need for good health.
   a. Drug approval is not unnecessarily delayed.
   b. Drugs are obtainable by patients.
   c. Accurate information is available to support decision making about risks and benefits.
   d. Patients get the best intervention possible, practical, and reasonable for their health needs.
   e. Patients get drugs with the required dosage and purity.
4. Patients take the drugs in a safe and effective manner.
   a. Patients get correct instructions about dosage and follow them.
   b. Patients do not take unsafe combinations of drugs.
   c. Patients are properly monitored by a physician while they are being treated.
   d. Patients are not subjected to unacceptable risk during clinical trials.

In system engineering, the requirements may not be totally achievable in any practical design. For one thing, they may be conflicting among themselves (as was demonstrated in the train door example) or with other system (non-safety) requirements or constraints. The goal is to design a system or to evaluate and improve an existing system that satisfies the requirements as much as possible today and to continually improve the design over time using feedback and new scientific and engineering advances. Tradeoffs that must be made in the design process are carefully evaluated and considered and revisited when necessary.

Figure 7.6 shows the general pharmaceutical safety control structure in the United States. Each component's assigned responsibilities are those assumed in the design of the structure. In fact, at any time, they may not be living up to these responsibilities.

Congress provides guidance to the FDA by passing laws and providing directives, provides any necessary legislation to ensure drug safety, ensures that the FDA has

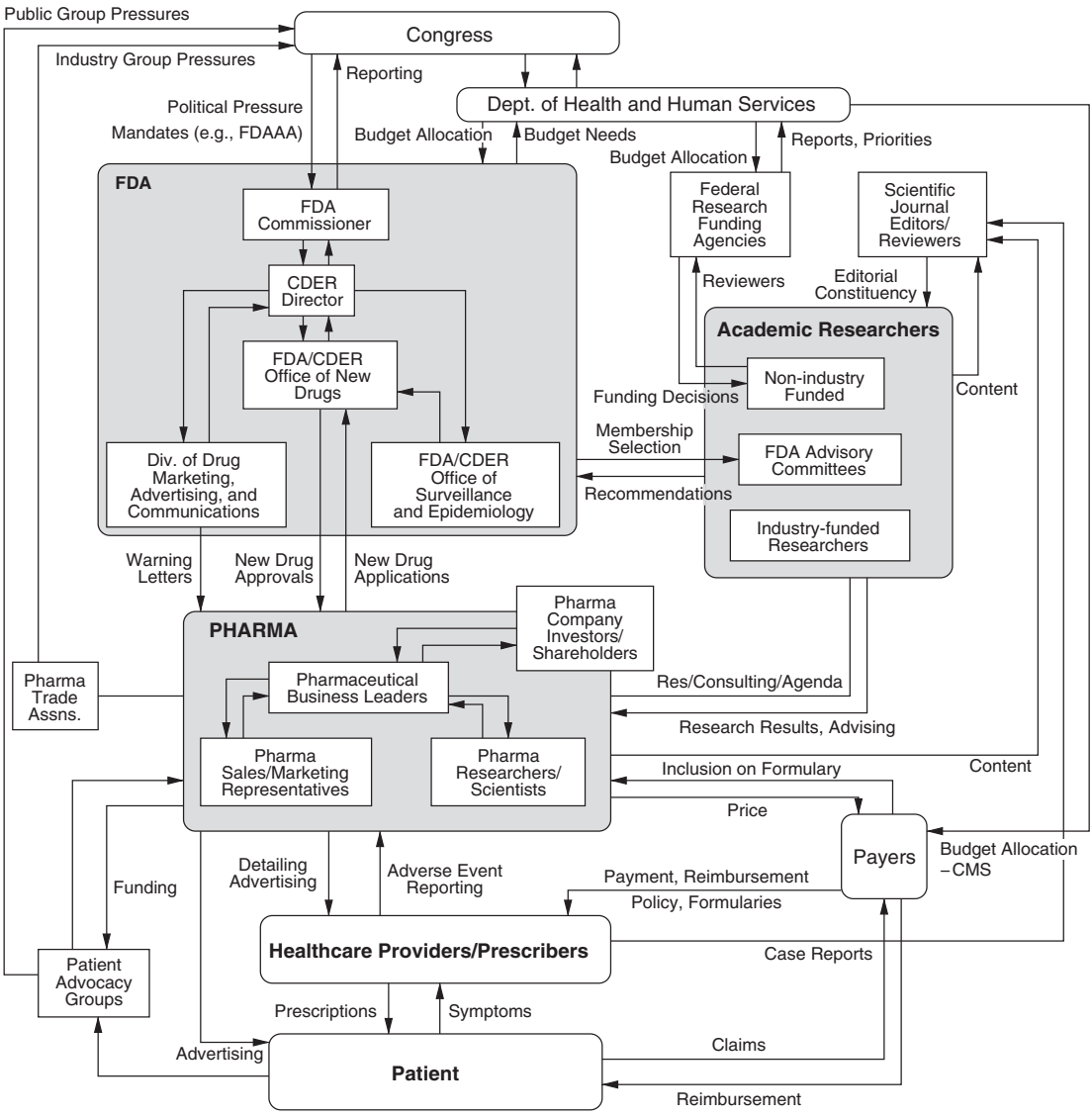

Figure 7.6
The U.S. pharmaceutical safety control structure.

enough funding to operate independently, provides legislative oversight on the effectiveness of FDA activities, and holds committee hearings and investigations of industry practices.

The FDA CDER (Center for Drug Evaluation and Research) ensures that the prescription, generic, and over-the-counter drug products are adequately available to the public and are safe and effective; monitors marketed drug products for unexpected health risks; and monitors and enforces the quality of marketed drug products. CDER staff members are responsible for selecting competent FDA advisory committee members, establishing and enforcing conflict of interest rules, and providing researchers with access to accurate and useful adverse event reports.

There are three major components within CDER. The Office of New Drugs (OND) is in charge of approving new drugs, setting drug labels and, when required, recalling drugs. More specifically, OND is responsible to:

- Oversee all U.S. human trials and development programs for investigational medical products to ensure safety of participants in clinical trials and provide oversight of the Institutional Review Boards (IRBs) that actually perform these functions for the FDA.
- Set the requirements and process for the approval of new drugs.
- Critically examine a sponsor's claim that a drug is safe for intended use (New Drug Application Safety Review). Impartially evaluate new drugs for safety and efficacy and approve them for sale if deemed appropriate.
- Upon approval, set the label for the drug.
- Not unnecessarily delay drugs that may have a beneficial effect.
- Require Phase IV (after-market) safety testing if there is a potential for long-term safety risk.
- Remove a drug from the market if new evidence shows that the risks outweigh the benefits.
- Update the label information when new information about drug safety is discovered.

The second office within the FDA CDER is the Division of Drug Marketing, Advertising, and Communications (DDMAC). This group provides oversight of the marketing and promotion of drugs. It reviews advertisements for accuracy and balance.

The third component of the FDA CDER is the Office of Surveillance and Epidemiology. This group is responsible for ongoing reviews of product safety, efficacy, and quality. It accomplishes this goal by performing statistical analysis of adverse event data it receives to determine whether there is a safety problem. This office reassesses risks based on new data learned after a drug is marketed and recommends

ways to manage risk. Its staff members may also serve as consultants to OND with regard to drug safety issues. While they can recommend that a drug be removed from the market if new evidence shows significant risks, only OND can actually require that it be removed.

The FDA performs its duties with input from FDA Advisory Boards. These boards are made up of academic researchers whose responsibility is to provide independent advice and recommendations that are in the best interest of the general public. They must disclose any conflicts of interest related to subjects on which advice is being given.

Research scientists and centers are responsible for providing independent and objective research on a drug's safety, efficacy, and new uses and give their unbiased expert opinion when it is requested by the FDA. They should disclose all their conflicts of interest when publishing and take credit only for papers on which they have significantly contributed.

Scientific journals are responsible for publishing articles of high scientific quality and provide accurate and balanced information to doctors.

Payers and insurers pay the medical costs for the people insured as needed and only reimburse for drugs that are safe and effective. They control the use of drugs by providing formularies or lists of approved drugs for which they will reimburse claims.

Pharmaceutical developers and manufacturers also have responsibilities within the drug safety control structure. They must ensure that patients are protected from avoidable risks by providing safe and effective drugs, testing drugs for effectiveness, properly labeling their drugs, protecting patients during clinical trials by properly monitoring the trial, not promoting unsafe use of their drugs, removing a drug from the market if it is no longer considered safe, and manufacturing their drugs according to good manufacturing practice. They are also responsible for monitoring drugs for safety by running long-term, post-approval studies as required by the FDA; running new trials to test for potential hazards; and providing, maintaining, and incentivizing adverse-event reporting channels.

Pharmaceutical companies must also give accurate and up-to-date information to doctors and the FDA about drug safety by educating doctors, providing all available information about the safety of the drug to the FDA, and informing the FDA of potential new safety issues in a timely manner. Pharmaceutical companies also sponsor research for the development of new drugs and treatments.

Last, but not least, are the physicians and patients. Physicians have the responsibility to:

- Make treatment decisions based on the best interests of their clients
- Weigh the risks of treatment and non-treatment

- Prescribe drugs according to the limitations on the label
- Maintain up-to-date knowledge of the risk/benefit profile of the drugs they are prescribing
- Monitor the symptoms of their patients under treatment for adverse events and negative interactions
- Report adverse events potentially linked to the use of the drugs they prescribe

Patients are taking increasing responsibility for their own health in today's world, limited by what is practical. Traditionally they have been responsible to follow their physician's instructions and take drugs as prescribed, accede to the doctor's superior knowledge when appropriate, and go through physicians or appropriate channels to get prescription drugs.

As designed, this safety control structure looks strong and potentially effective. Unfortunately, it has not always worked the way it was supposed to work and the individual components have not always satisfied their responsibilities. Chapter 8 describes the use of the new hazard analysis technique, STPA, as well as other basic STAMP concepts in analyzing the potential risks in this structure.