WORKPLACE FATALITIES in 2003 dropped to 4,500—a 3.2 death rate per 100,000 workers, down from 3.4 in 2003, representing a 5.8 percent reduction in one year (Hoskin). This continued a trend that started in 1994. Does this mean industry did a better job of preventing injuries or was it luck? The question probably cannot be answered by simply looking at one metric: fatality rates.

For example, looking back at the period 1973 to 1999 and considering other metrics for this period:

- in total cases, there was a 45-percent reduction (or 1.6 percent per year);
- in cases without lost workdays, there was a 57-percent reduction (or 2.1 percent per year);
- in lost workday cases, there was an 11.4-percent reduction (or 0.4 percent per year).

While a 5.8-percent reduction in a year sounds good, 0.4 percent per year is 93-percent less good. Clearly, the answer depends on the selected measure (Hoskin).

Success—and failure—in occupational safety and health can be measured in many ways. Before OSHA was created, ANSI Z16.1 contained three measures of injury experience:

1) Disabling injury frequency rate (the most popular expression of industrial safety performance). Its formula:

F = \frac{\text{(number of disabling injuries}}{10^6})/\text{employee hours of exposure}.

Example: Assuming an establishment experienced 12 employee disabling injuries during a one-month measurement period, and recorded a total employee hours of exposure amounting to 2,189,243, then, in the frequency formula:

F = \frac{(12 \times 10^6)}{2,189,243} = 5.48 disabling injuries per million employee hours of exposure.

2) Disabling injury severity rate (essentially a weighted frequency rate). This measure expresses the days actually lost due to temporary total disabilities and the days charged (arbitrarily by an ANSI schedule) for the fatal and permanently disabling cases. The formula, where total days charged equals temporary total days lost plus schedule charges for permanent disabilities, was:

S = \frac{(\text{total days charged} \times 10^6)}{\text{employee hours of exposure}}.

Example: Assuming in the first example that the establishment experienced 12 temporary total cases that collectively totaled the days lost due to disability and one death for which the schedule charge is 6,000 days, plus one loss of an eye for which the schedule charge is 1,800 days. The combined total days lost and charged would equal 7,872 days.

S = \frac{(7,872 \times 10^6)}{2,189,243} = 3,596 days lost and charged per million employee hours of exposure.

3) Average days charged per disabling injury. This measure is the ratio of severity to frequency rates. It may also be calculated as the ratio of the total days lost and charged to the total of disabling injuries.

S/F = \frac{\text{total days charged}}{\text{number of disabling injuries}}.

Example: Using the data for the assumed establishment in the first two examples and employing the given ratios: Average days charged per disability injury = 3,596/5.48 or 7,872/12 = 656 (ANSI).

Since 1972, SH&E practitioners have been using OSHA measures. These OSHA measures—all per 100 full-time
Levels of Severity
Consider these four levels of severity:
1) injuries without lost time;
2) injuries with lost time;
3) fatalities;
4) catastrophes.
The first three are often described with these labels under workers’ compensation:
1) temporary total injuries (with or without lost time);
2) permanent partial injuries (permanent disabilities);
3) permanent total injuries (fatalities).

Levels of injury severity can be measured in different ways, with different agencies adopting different systems. The Bureau of Labor Statistics in the U.S. adopted the Bureau of Labor Statistics’s Census of Fatal Occupational Injuries figure as the authoritative count of work-related deaths beginning with the 1992 data.

Of these five measures, only the first, total cases, seems to be universally used to measure a company’s or a manager’s safety performance. The OSHA incident rate was not intended to be used for this purpose. According to Steve Newell, a consultant with Organization Resources Counselors (ORC) in Washington, DC, the current system for OSHA reporting was developed in 1970 to give the U.S. government a snapshot of occupational health and injuries in the workplace.

The primary ANSI measure used was the disabling injury frequency rate, which is comparable to the (seldom used) OSHA total lost workday measure. Comparing these two shows a 0.4 percent per 100,000 reduction for the 25-year period (or 0.4 percent per year mentioned earlier, the worst improvement of all measures).

Why the discrepancy in results as measured by different measures of different levels of injury severity? Could it be that different causal factors are involved for different levels of severity? Would that suggest different controls needed for each?

Accident Types & Severity

<table>
<thead>
<tr>
<th>Type of Accident</th>
<th>Temp. Total %</th>
<th>Perm. Partial %</th>
<th>Perm. Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Handling materials</td>
<td>24.3</td>
<td>20.9</td>
<td>5.6</td>
</tr>
<tr>
<td>Falls</td>
<td>18.1</td>
<td>16.2</td>
<td>15.9</td>
</tr>
<tr>
<td>Falling objects</td>
<td>10.4</td>
<td>8.4</td>
<td>18.1</td>
</tr>
<tr>
<td>Machines</td>
<td>11.9</td>
<td>25.0</td>
<td>9.1</td>
</tr>
<tr>
<td>Vehicles</td>
<td>8.5</td>
<td>8.4</td>
<td>23.0</td>
</tr>
<tr>
<td>Hand tools</td>
<td>8.1</td>
<td>7.8</td>
<td>1.1</td>
</tr>
<tr>
<td>Electricity</td>
<td>3.5</td>
<td>2.5</td>
<td>13.4</td>
</tr>
<tr>
<td>Other</td>
<td>15.2</td>
<td>10.8</td>
<td>13.8</td>
</tr>
</tbody>
</table>

Source: Petersen.
accounts for 25 percent of all temporary total disabilities and 21 percent of all permanent partials. But only six percent of all permanent total injuries and fatalities (Hoskin). Electricity accounts for 13 percent of all permanent totals and fatalities, but accounts for a negligible percentage of temporary totals and permanent partials.

These percentages would not differ if the causes of frequency and severity were the same; therefore, they are not the same. Different sets of circumstance surround severity. Thus, to control serious injuries, one should try to predict where they will occur—and now it is possible to do just that. For example, it appears that severe injuries are fairly predictable in certain situations, including:

- **Unusual, nonroutine work.** This is the job that occurs only occasionally or perhaps only once. These situations may arise in production or nonproduction departments. The normal controls that apply to routine work have little effect in the nonroutine situation.

- **Nonproduction activities.** Much of safety’s efforts have focused on production work. However, tremendous potential exposure to loss exists in nonproduction activities such as maintenance, and research and development. In such activities, most work tends to be nonroutine. However, since no production is involved, these tasks rarely require safety attention, and often lack formal procedures. Severity is predictable here.

- **Sources of high energy.** In most cases, high energy can be associated with severity. Electricity, steam, compressed gases and flammable liquids are examples.

- **Construction situations.** Examples include high-rise erection, tunneling and working over water. In fact, construction severity is often an amalgam of the previously described high-severity situations.

These are just a beginning point. One could develop a long list to more extensively identify those areas where severity is predictable.

Under the old ANSI system, this was not a problem because the main measure—“frequency rate”—included some relatively serious injuries with perhaps considerable lost time, but excluded the very minor injuries with no lost time (e.g., cuts, bruises, paper cuts, bee stings).

Under the OSHA measures, these minor injuries are often the larger percentage of what is included. To appear to have good control of the accident problem, many companies concentrate on the minor injuries and at times go to great lengths to reduce the amount or type of prescription drugs used in an individual case so that the injuries will not “count.” If the safety goal is to reduce these minor injuries, are the causes of the serious ones being overlooked? Is this why in cases without lost workdays, the national record is 93-percent better than the success in reducing lost workday cases?

Years ago, John Grimaldi stated that “qualitative judgments of safety performance, reached exclusively in terms of the frequency, are apt to be grossly incorrect” (Grimaldi). He further suggested that severity rates correlate much better with the costs of accidents than with frequency rates, making them more meaningful to upper management.

He further suggested nonaccident measures, such as behavior sampling, critical incident techniques, inspections and audits, as more meaningful in judging safety system effectiveness. Other such measures might include interviews and perception surveys. In fact, Newell suggests SH&E professionals identify a subset of causes for measurement and accountability that are reasonably serious and connected to the workplace instead of OSHA incident rates (Newell).

**Suggested Modifications**

Experience with standards for reporting and measuring accident experience has led to widespread dissatisfaction on the part of many responsible for safety management activities and performance because of the limitations described. One consequence has been the development and application of various reporting modifications, generally proposed for application concurrently with standard reporting. The following discussion highlights three of these modifications.

**Serious Injury Index (1965)**

This involves an attempt to enlarge the sample of occurrences beyond those that are defined as disabling injury accidents. In effect, it includes certain types of non-disabling injuries along with those disabling injuries that meet the lost-time criterion. These include specified kinds of non-disabling eye injuries, fractures, lacerations and other injuries for which work restrictions are prescribed (Gilmore).

**High Potential Accident Index (1967)**

This involves enlarging the universe of events beyond disabling injury accidents alone while also generating a safety performance measure that remains sensitive to serious rather than minor hazards (Allison). In practice, a reporting system is established to encourage the reporting of all injury-producing accidents, whether disabling or not. These events are then studied in order to identify those few that investigators believe had the potential for serious human or property damage (estimated at 4/10 of one percent of all accidents versus 7/1,000 of one percent for lost-time accidents). These high-potential accidents represent the measurement base of interest. In support of this proposal, it has been argued that indexes based merely on raw minor injury data tend to reflect reporting rather than incidence; place pressure on reporting and minor injury reduction; and do not reflect the overall level of accident cost performance (Allison).

**Property Damage Accident Index (1966)**

This measurement system includes property damage accidents (Bird). The ratio of property to disabling injury accidents is at least 500 to one. The measurement index most commonly mentioned is that of total realized dollar cost for replacement, repair, lost production time and related factors. The plan as it has been applied in practice requires a workable scheme for enforced reporting of all property damage accidents (Bird).

**Catastrophes**

The fourth level of severity—the catastrophe—is somewhat different. It usually is most infrequent within a single organization, and constitutes several serious injuries, much property damage, fines, lost production and similar losses. It does not lend itself to easy measurement. Because of this, it is rarely discussed in the safety literature, as well as in corporate planning and systems. The “Incidents & Costs” sidebar on pg. 46 provides a partial listing from the 1940s to the 1990s (sources included CNN, AP and safety reports).

Perhaps the only effective measure is to look at the research in this area and develop a picture of what is behind a catastrophe—what are the causal factors? Often, safety systems, regulations and recordkeeping focus on controlling all injuries with equal emphasis, regardless of their potential severity and financial loss to the organization.

In the areas of regulation and required recordkeeping, a bee sting and a fatality are counted equally. At times, a paper cut suffered by an office worker is counted as a recordable injury, whereas an explosion that killed members of the public but no employees is not recordable. SH&E pro-
Incidents & Costs
This list of sample incidents and costs was generated from accessible CNN, AP, safety and press reports, and is only a small sample of what has occurred.

Texas City Explosion
Date: April 16, 1947
Incident type: French ship exploded while docked.
Details: Ammonia nitrate blew up.
Costs: 576 fatalities; 5,000 injured.

Mexico City Gas Storage Explosion
Date: Nov. 19, 1983
Incident type: Four spherical 420,000-gallon tanks ignited from propane truck at loading dock.
Details: Homes were allowed to be built near the facility.
Costs: 30 acres of homes destroyed, 30 acres damaged; 540 fatalities; 2,200+ injured; 10,000+ homeless.

Soviet Nuclear Incident at Chelyabinsk-65
Date: Sept. 29, 1957
Incident type: Tank of radioactive waste exploded.
Details: Probably contaminated 357 square miles. 10,000 evacuated.
Costs: Possibly several hundred fatalities; 200 million rubles.

Farmer’s Export Grain Elevator Explosion
Date: Dec. 27, 1977
Details: Grain dust ignited by spark.
Costs: 18 fatalities.

Vila Soco Pipeline Fire
Date: Feb. 25, 1984
Incident type: Pipeline gasoline blaze exploded, burned at over 1,000ºC through Brazilian village.
Details: The wrong pipeline was opened the day preceding the fire.
Costs: 500+ fatalities (child casualties under age five had to be estimated since they were totally incinerated).
Proposed/reported fines/penalties: Petrobras paid hospital costs and damages.

San Juanico Pemex Gas Explosion
Date: Nov. 19, 1984
Incident type: Series of liquefied gas storage explosions in San Juanico, Mexico.
Details: Fireball flashed through suburban area at 5:43 am.
Costs: 503 fatalities; 4,000+ injuries.
Proposed/reported fines/penalties: Pemex held liable by federal attorney general. By 1986, Pemex had $5 million in claims.

Chernobyl
Date: April 26, 1986
Incident type: Nuclear plant meltdown and radiation release. Poor engineering and operation combined.
Details: Released 50 million curies of radiation into surrounding area.
Costs: 250+ deaths; $26 billion planned to move 200,000 additional residents; $2 billion planned to rebury the plant.

Alaskan Oil Spill
Date: March 24, 1989
Incident type: 987-foot tanker smashed into Bligh Reef and spilled 11 million gallons of oil into Prince William Sound.
Details: Captain left the bridge during maneuvers. He and crew have been blamed by government officials and others.
Costs: Eventually labeled as a human fatigue incident by NTSB investigators.
Proposed/reported fines/penalties: Ongoing. U.S. Congress passed a bill allowing states to adopt stricter spill liability laws than the federal government requires.

Phillips Petroleum Pasadena Explosion
Date: Oct. 23, 1989
Incident type: Gas release led to explosion that destroyed portion of a polyethylene plant.
Details: Phillips Petroleum said the company’s own investigation showed the explosion “was the result of a departure from established routine procedures.”
Costs: 23 fatalities; 314 injuries; Phillips experienced $431 million decrease in net income that year.
Proposed/reported fines/penalties: OSHA first proposed $6.4 million in fines. Later reduced them in exchange for promise to institute process safety management procedures at Pasadena and three other plants.

Channelview Texas Chemical Plant Explosion
Date: July 5, 1990
Incident type: Houston Arco Channelview plant suffered explosion that burned city-block-sized area. Fire lasted more than four hours.
Details: Inadequate training and excessive overtime work have been mentioned as possible causes of the accident.
Costs: 17 fatalities.
Proposed/reported fines/penalties: $3.48 million in fines.

As noted, some research has focused on the causes of catastrophes, particularly those caused by humans. One of the best investigations was that of Edwin Zebrowski of Aptech Engineering Services, which focused on several major human-caused catastrophes to determine whether any common negative characteristics were present before the incidents occurred. Several of his conclusions are enlightening (Petersen).

The most fundamental human factor is obviously management—the capabilities, organization and degree of involvement in proactive safety and reliability practices. Sometimes it takes a great catastrophe to bring the needed capabilities and involvement into play. Some recent examples—Three Mile Island (TMI), Chernobyl, Bhopal, the Challenger shuttle and the Piper Alpha oil platform—are important to study and to understand because they illustrate the risky attitude “it can’t happen here” and highlight the fact that “good practices” are not cliches. The basic lesson is that the absence (or weakening) of just a few good practices can lead to a catastrophe. It is not difficult to make a list of the practices whose absence can make a catastrophe not only probable, but essentially certain—that is, only a matter of time.

When one examines common factors in large man-made catastrophes, one always finds that many relevant and ultimately
crucial factors were ignored. In most cases, nobody wanted to look at these factors or the protective and/or remedial actions needed. One benefit of structured decision analysis is to make key factors in a decision explicit rather than implicit, and to get in view any “sacred cows” that can put blinders on how decisions are made or delayed.

Eleven negative attributes found have had medium to large degrees of commonality in the basis for TMI, Chernobyl, Challenger and Bhopal events. They are:

1) Diffused responsibilities with rigid communication channels and large organization distances from decision makers to the plant;
2) Mindset that success is routine with neglect of severe risks which are present;
3) Rule compliance and the belief that this is enough to ensure safety;
4) Team player emphasis with dissent not allowed even for evident risk;
5) Experience from other facilities not processed systematically for application locally;
6) Lessons learned from the past or from others disregarded;
7) Safety analysis and responses subordinate to other performance goals in operating priorities;
8) Emergency procedures, plans, training and regular drills for severe events lacking;
9) Design and operational features allowed to persist even though recognized elsewhere as hazardous;
10) Project and risk management techniques available but not used;
11) Organization with undefined responsibilities and authorities for recognizing and integrating safety matters.

The matrix in Table 2 lists these 11 attributes and identifies the catastrophe to which each was a major contributor. Note that all of these systems shared at least 10 of the attributes. Chernobyl was surely not an outlier, and the fact that disasters have resulted from multiple failures means that in the future, people at the scene of an impending failure may have the opportunity to break the chain of events.

Table 2
Matrix of Common Attributes: Four Severe Accidents

<table>
<thead>
<tr>
<th></th>
<th>Bhopal</th>
<th>Challenger</th>
<th>Chernobyl</th>
<th>TMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responsibilities</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>“Mindset”</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>True compliance</td>
<td>(x)</td>
<td>x</td>
<td>0</td>
<td>x</td>
</tr>
<tr>
<td>Team agreement</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Prior events</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Narrow experience</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Output vs. safety</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>(x)</td>
</tr>
<tr>
<td>Severe accidents</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Known hazards</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Risk techniques</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Safety integration</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Environmental conditions were several levels removed from detailed technical knowledge of actual hazards present.

Certain clear patterns repeat throughout 200 years of man/machine interfaces, and these are at the heart of James Chiles’s Inviting Disaster: Lessons from the Edge of Technology. This book distills lessons out of 64 technological catastrophes and close calls.

The repeating patterns of error make up the book’s chapters and include undocumented field changes, failure to test new products thoroughly, poor communication and maintenance lapses. A lack of personal accountability helped transform small failures into big problems. Most of the disasters involved no particularly outrageous conduct; instead, heedlessness—a failure to take seriously the full consequences of error—was to blame.

The idea of a fracture extending through points of weakness helps explain why the initial event leading to a technological disaster can be exceedingly small. Consider the rapid-fire set of events that caused the crash of an Air France Concorde in July 2000. A 16-inch strip of titanium lying on the Paris runway blows one of the plane’s tires; the tire flings off fragments; a 10-pound chunk of rubber spins off to thump against a wing tank; a shock wave through the fuel blows a hole in the tank wall from inside; a torrent of kerosene spills onto the left-side engine intakes and imbalances the thrust during climbout. This string of small events adds up to a fiery crash into a hotel, killing all on board and four others on the ground.

One must also note that the Concorde fleet had 32 tire blowouts over the preceding years—six of which had caused fuel-tank leaks. The fact that disasters spring from multiple failures means that the future, people at the scene of an impending failure may have the opportunity to break the chain of events.

Managerial Climate, Groupthink & Culture

Most of the negative attributes highlighted by Zebrowski are management
characters. Further assessment of the causes of the Challenger and Columbia disasters are provided in Apollo, Challenger, Columbia: The Decline of the Space Program by Phillip Tomkins. In one section, he discusses the “ideal managerial climate” as espoused by Charles Redding. It consists of five factors:
1. supportiveness;
2. participative decision making;
3. trust, confidence and credibility;
4. openness and candor;
5. emphasis on high-performance goals.
Tomkins also pursues the idea that NASA’s “curse” was groupthink.
At NASA, it really is rocket science, and the decision makers really are rocket scientists. But a body of research that is getting more and more attention points to the ways that smart people working collectively can be dumber than the sum of their brains.
The culture of NASA could lead to group dynamics as a problem, and the concept of groupthink, a term coined by Irving Janis, then a professor of social psychology at Yale University. Janis’s definition of groupthink was “a mode of thinking that people engage in when they are deeply involved in a cohesive in-group, when the members’ strivings for unanimity override their motivation to realistically appraise alternative courses of action.” Janis found this phenomenon in the Kennedy administration’s ill-fated decision to invade Cuba’s Bay of Pigs and in the escalation of the Vietnam War. He also believed it applied to the explosion of the Challenger.
To avoid groupthink, leadership must ask penetrating questions and listen hard to what seem to be deviant messages. Indeed, leaders should encourage members to play devil’s advocate and members should accept that role without fear of being punished for advocating a different theory.

The Criteria to Judge
Tomkins proposed eight criteria to assess what stance might be appropriate to assess causal factors:
• Causal force. Did the action or inaction of the individual have causal force, or did it amount merely to being one act in a long chain of events?
• Hierarchy. What was the person’s formal degree of authority and responsibility?
• Values of the culture. Was the individual’s act consistent with the ultimate and avowed values of the culture?

• Consequences. Were the consequences of the individual’s act or acts trivial or significant?
• Justice. Does the punishment fit the crime? Would the effect of blame be commensurate with the act?
• Defense. Did the person truthfully deny or accept responsibility for the act?
• Actor agency. Did the person have autonomy or control over the act or attribute that offends?
• Future actions. Might the person(s) accountable make the same mistake again (Tomkins)?

NASA’s Safety Culture
In assessing NASA’s culture, Tomkins examined the agency’s “safety culture” separately and observed that it is “straining to hold together the vestiges of a once robust system safety program.” A short history lesson made it clear that the safety culture had been weakened by government decisions over the years to reduce the workforce and make NASA more dependent on contractors for technical and safety support. This eroded NASA’s in-house engineering depth, making it a slimmed-down agency largely run by contractors.

Need for a New Approach
Given this, with respect to the control of severity, is SH&E progressing or regressing? The major emphasis remains on frequency with little emphasis on severity, much less on catastrophic losses. The focus continues to be on frequency reduction (if the frequency comes down, performance is better). Meanwhile, severity stays roughly the same. Granted, fatalities apparently are reduced, but serious (lost-time) statistics remain stagnant. Is this acceptable? This is up to SH&E professionals. And how can one know? By which numbers are results judged? These are questions today’s professionals must address.

Perhaps the answer is to use more than one metric for both goal setting and measurement. Many companies do this today, typically combining results measures with one or more activity measures, thus using both upstream and downstream—or leading and lagging—indicators to assess progress. Rarely, however, is a severity indicator included in the mix.

Perhaps the ideal strategy would include a frequency measure, a severity measure, and one or more activity measures; or a measure of activities (e.g., audits), a measure of results (noninjury sampling), or a valid measure of people’s perception of the success of the safety effort (perception survey). Together, these might provide a more-accurate assessment of the system’s value.

Considerable new and innovative thinking is taking place in many organizations. As with safety system content, there is no one right way when it comes to safety metrics. Each organization must determine its own “right way.” In addition, after deciding what components to include, one must decide how each component should be weighted, making it possible to come up with a single metric if so desired. At some point, SH&E professionals may have to do this. The sooner we start down this road, the better.

References

Dan Petersen, Ph.D., PE, is a consultant specializing in safety management and organizational behavior. He holds a B.S. in Industrial Engineering from Iowa State University, an M.S. in Industrial Psychology from University of Nebraska, and a Ph.D. in Organizational Behavior and Management. Petersen is an ASSE Fellow, a professional member of ASSE’s Arizona Chapter and a member of the Society’s Management Practice Specialty.

Professional Safety encourages readers to write letters in response to specific articles, editorials, letters, columns and news reports published in PS. The editors reserve the right to select which letters will appear and to edit letters for brevity and clarity. Send Reader Feedback comments to professionalsafety@asse.org.