Chapter 14

In Search of Safety

Introduction

Safety is a term defined more by its absence than its presence. This last chapter seeks to redress the balance by presenting two models that highlight different aspects of the positive and more hidden face of safety. The safety space model deals with goal-setting, intrinsic resistance to operational hazards, the relationship between proactive process measures and reactive negative outcome data, and the importance of both cultural drivers and navigational aids in achieving the maximum practical state of intrinsic resilience. As a corollary to this model, I will show how the three principal cultural drivers – commitment, cognisance and competence – map on to the four 'Ps' of management – principles, policies, procedures and practices – to provide a broad description of what a resilient and safe organisation might look like.

The second model exploits the mechanical properties of a knotted rubber band to enlarge upon the notion of safety as a dynamic non-event. Together the two models provide complimentary views of the nature of safety. The safety space model addresses the more strategic aspects of safety, while the rubber band model focuses on the tactical issues of local control.

What Does the Term 'Safety' Mean?

Like 'health,' the word 'safety' suffers from an imbalance of understanding. Far more is known about its momentary absences than about its longer-lasting presence. We are much better at describing, comprehending and quantifying the occasional deviations from this state, expressed very concretely as accidents,

injuries, losses, incidents and close calls than we are at explaining what it means to be safe.

Dictionaries take us no further since they also treat safety as the absence of something. The *Concise Oxford Dictionary*, for example defines safety as 'freedom from danger or risks.' The *Shorter Oxford English Dictionary* gives its meaning as 'exemption from hurt or injury, freedom from dangerousness . . . the quality of being unlikely to cause hurt or injury.'

These everyday uses are of little help to those engaged in the safety sciences or in the management of risk. They neither capture the reality of such activities as aviation, health care and nuclear power generation where the hazards – gravity, terrain, weather, error, hospital-acquired infections, radioactive materials and the like – are ever-present, nor do they tell us much about the nature of the goals that those working within hazardous systems must strive to attain. These people, naturally enough, see their target as the reduction and elimination of harm and losses. But this is only partially within their control. Moreover, in most modern, well-defended technologies such unhappy outcomes are so infrequent as to provide little or no guidance on how to restrict or prevent bad events. Of course, this is not necessarily true for the more 'close encounter' industries, such as mining, transport, health care and construction, but the main focus of this chapter is upon those high-technology activities in which the dangers are potentially great and far-reaching, but where the frequency of adverse events is generally low.

Compared to the natural sciences, where worth is gauged by how much empirical activity a particular theory generates, safety scientists face an additional challenge. As well as provoking interest, a safety-related contribution must also have practical utility. But it can only achieve this if it is readily communicable to those people engaged in the day-to-day business of managing the safety of hazardous operations. Here, as in the behavioural sciences, models, images, metaphors and analogies have an essential part to play. Not only can they convey complex ideas in a concise and digestible fashion (and few enterprises are more complex than the pursuit of safety), they also make it easier for safety specialists, working within dangerous systems, to disseminate these ideas with in their respective organisations.

Such models do not have to be 'true' in the literal sense, nor do they have to be consistent one with another; rather each should tell a story (or draw a picture) that captures some important aspect of an otherwise elusive and mysterious phenomenon. The most useful models also have an internal logic or explanatory 'engine' that highlights the significance of some hitherto unrevealed, or at least unremarked, safety process. The ultimate criterion, though, is a very practical one. Do the ideas communicated by the model lead to measures that enhance a system's resistance to its operational hazards? In short, does it improve safety?

The Two Faces of Safety

Safety has a negative and a positive aspect, though it is mainly the former that claims attention. They are summarised below:

- The negative face is revealed by reactive outcome measures: accidents, fatalities, injuries, loss of assets, environmental damage, patient safety incidents and adverse events of all kinds. Then there are also close calls, near misses and 'free lessons'. All of these are readily quantified and hence much preferred by number-hungry technical managers. These numbers may be convenient and easy to manipulate, but beyond a certain point, they have very dubious validity, as we shall see later.
- The positive face of safety relates to the system's intrinsic resistance to its operational hazards. It is assessed by proactive process measures – indices that reflect an organisation's 'health' both in regard to production and safety. I will say more about these indices later.

The main purpose of the 'safety space' model is to elucidate what exactly is meant by the positive face of safety. But before describing it, let me sneak in yet another metaphor that illustrates the notions of vulnerability and resilience. Figure 14.1 shows a ball-bearing (representing the system) sitting on top of variously shaped metal blocks. Both the ball-bearing and the block are subject to continual jigglings or perturbations that seek to topple the ball-bearing off the block – equivalent to an accident.

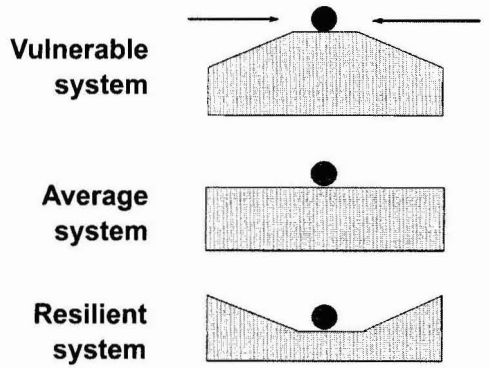

Figure 14.1 Illustrating vulnerability and resistance. The arrows at the top of the figure represent perturbing forces

It is clear that the top ball-and-block are the most vulnerable (easily toppled), while the bottom set are the most resistant. Note, however, that even in this bottom configuration, it is still possible to dislodge the ball. If you prefer a more homely example, think of a tray with an egg and a Pyrex bowl on it. In the vulnerable arrangement, the bowl is inverted and the egg is on top. In the resistant arrangement, the egg is inside the bowl. Perturbations come from tiltings of the tray.

The 'Safety Space' Model

The first model to be described here embodies a navigational metaphor. It presents the notion of a 'safety space' within which comparable organisations can be distributed according to their relative vulnerability or resistance to the dangers that beset their particular activities. They are also free to move to and fro within this space. An important feature of this model is that it seeks to specify an attainable safety goal for real world systems. This is not zero accidents, since safety is not an absolute state; rather it is the achievement and maintenance of the maximum intrinsic resistance to operational hazards.

The model had it's origins in analyses of individual differences in the numbers of accidents experienced by groups of people

exposed to comparable hazards over the same time period (we discussed this at some length in Chapter 6). These variations in liability are normally expressed in relation to the predictions of some chance theoretical distribution – the Poisson exponential series. A Poisson distribution looks roughly like the right hand half of a bell-shaped (normal or Gaussian) distribution. But the accident liability distribution is of necessity one-sided; it can only assess degrees of liability. Our concern is with the opposite and previously neglected end of the distribution, most especially with the fact that more than half of the groups assessed in this way have zero accidents. Was this simply due to chance? Were these people simply lucky? It is probable that some of them were. But it is also likely that others possessed characteristics that rendered them less susceptible to accidental harm.

In other words, this unidirectional account of accident liability – discriminating, as it does, degrees of 'unsafety' within a given time period – might actually conceal a bi-directional distribution reflecting variations in personal safety ranging from a high degree of intrinsic resistance to considerable vulnerability. It is a short step from this notional bi-directional distribution of individual accident liability to the 'safety space' shown in Figure 14.2.

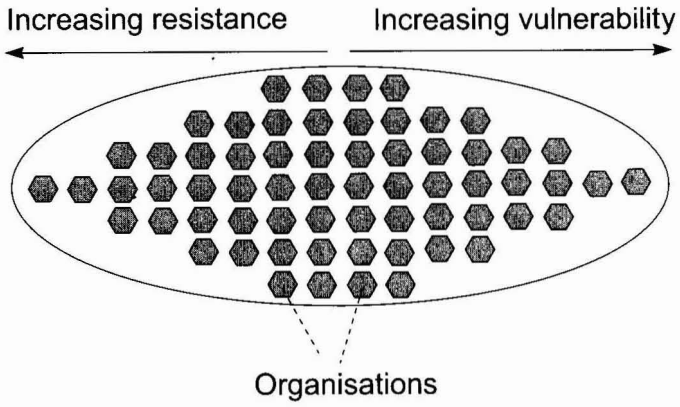

Figure 14.2 Showing a number of hypothetical organisations within the same hazardous domain distributed throughout the safety space

The horizontal axis of the space runs from an extreme of maximum attainable resistance to operational hazards (and still stay in business) on the left to a maximum of survivable vulnerability on the right. Rather than individuals, however, we have plotted the position of a number of hypothetical organisations operating within the same hazardous conditions along this resistance-vulnerability dimension. The space's cigar shape acknowledges that most organisations will occupy an approximately central position with very few located at either extreme.

There will probably be some relationship between an organisation's position along the resistance-vulnerability dimension and the number of bad events it suffers during a given accounting period, but it is likely to be a very tenuous one. If, and only if, the system managers had complete control over all the accident-producing conditions within their organisations would we expect their accident and incident rates to bear a direct relationship to the quality of their efforts. But this is not the case. Chance also plays a large part in accident causation. So long as operational hazards, local variations and human fallibility continue to exist, chance can combine with them in ways that breach the system's defences.[1] Thus, even the most resistant organisations can still have bad accidents. By the same token, even the most vulnerable organisations can evade disaster, at least for a time. Luck works both ways: it can afflict the deserving and protect the unworthy.

The imperfect correlation between an organisation's position along the resistance–vulnerability continuum and the number of adverse events it sustains in a given accounting period has a further implication. When the accident rates within a particular sphere of activity fall to very low levels, as they have in aviation and nuclear power, the occurrence or not of negative outcomes reveals very little about an organisation's position within the safety space. This means that organisations with comparably low levels of accidents could occupy quite different locations along the resistance–vulnerability continuum, and not know it. So how can an organisation establish its own position within the space? In short, what navigational aids are available?

1 Reason, J. (1997) *Managing the Risks of Organizational Accidents*. Aldershot: Ashgate Publishing.

Each commercial organisation has two imperatives: to keep its risks as low as possible and still stay in business. It is clear that for any organisation continuing to operate profitably in dangerous conditions, the state of maximum resistance will not confer total immunity from harm. Maximum resistance is only the best that an organisation can reasonably achieve within the limits of its finite resources and current technology. Given these constraints, there are two ways by which it can locate its position within the safety space: from reactive and proactive indices.

Where major accidents are few and far between, the reactive measures will be derived mainly from near miss and incident reporting systems, or 'free lessons.' Such safety information systems have been considered at length elsewhere[2] and will not be discussed further here. We can, however, summarise their likely benefits:

- If the right lessons are learned from these retrospective data, they can act like vaccines to mobilise the organisation's defences against some more serious occurrence in the future. And, like vaccines, they can do this without lasting harm to the system.
- These data can also inform us about which safeguards and barriers remained effective, thus thwarting a more damaging event.
- Near misses, close calls and 'free lessons' provide qualitative insights into how small defensive failures could combine to cause major accidents.
- Such data can also yield the large numbers required for more far-reaching quantitative analyses. The analysis of several domain-related incidents can reveal patterns of cause and effect that are rarely evident in single-case investigations.
- More importantly, the understanding and dissemination of these data serve to slow down the inevitable process of forgetting to be afraid of the (rarely experienced) operational dangers, particularly in systems, such as nuclear power plants, where the operators are physically remote from both the processes they control and their associated hazards.

2 Van der Schraaf, T.W., Lucas, D.A., and Hale, A.R. (1991) *Near Miss Reporting as a Safety Tool*. Oxford: Butterworth-Heinemann.

Proactive measures identify in advance those factors likely to contribute to some future event. Used appropriately, they help to make visible to those who operate and manage the system the latent conditions and 'resident pathogens' that are an inevitable part of any hazardous technology (see Chapter 7). Their great advantage is that they do not have to wait upon an accident or incident; they can be applied now and at any time. Proactive measures involve making regular checks upon the organisation's defences and upon its various essential processes: designing, building, forecasting, scheduling, budgeting, specifying, maintaining, training, selecting, creating procedures, and the like. There is no single comprehensive measure of an organisation's 'safety health'.[3] Just as in medicine, establishing fitness means sampling a subset of a much larger collection of leading indicators, each reflecting the various systemic vital signs

Effective safety management requires the use of both reactive and proactive measures. In combination, they provide essential information about the state of the defences and about the systemic and workplace factors known to contribute to bad outcomes. The main elements of their integrated employment are summarised in Table 14.1.

Table 14.1 Summarising the interactions between reactive and proactive measures

	Type of navigational aid	
	Reactive measures	Proactive measures
Local and organisational conditions	*Analysis of many incidents can reveal recurrent patterns of cause and effect.*	*Identify those conditions most needing correction, leading to steady gains in resistance or 'fitness'.*
Defences barriers and safeguards	*Each event shows a partial or complete trajectory through the defences.*	*Regular checks reveal where holes exist now and where they are most likely to appear next.*

3 Reason (1997).

Navigational aids are necessary but insufficient. Without some internal driving force, organisations would be subject to the 'tides and currents' present within the safety space. These external forces run in opposite directions, getting stronger the nearer an organisation comes to either end.

The closer an organisation approaches the high-vulnerability end of the space, the more likely it is to suffer bad events – though, as mentioned earlier, this is by no means inevitable. Few things alert top management to the perils of their business more than losses or a frightening near miss. Together with regulatory and public pressures, these events provide a powerful impetus for creating enhanced safety measures which, in turn, drive the organisation towards the high-resistance end of the space. However, such improvements are often short-lived. Managers forget to be afraid and start to redirect their limited resources back to serving productive rather than protective ends. Organisations become accustomed to their apparently safer state and allow themselves to drift back into regions of greater vulnerability. Without an 'engine,' organisations will behave like flotsam, subject only to the external forces acting within the space.

Consideration of the 'safety engine' brings us to the cultural core of an organisation. Three factors, in particular, are needed to fuel the 'engine, all of them lying within the province of what Mintzberg called the 'strategic apex' of the system.[4] These driving forces are commitment, competence and cognisance.

Commitment has two components: motivation and resources. The motivational issue hinges on whether an organisation strives to be a domain model for good safety practices, or whether it is content merely to keep one step ahead of regulatory sanctions (see Chapter 5 for a discussion of the differences between 'generative' and 'pathological' organisations). The resource issue is not just a question of money, though that is important. It also concerns the calibre and status of those people assigned to direct the management of system safety. Does such a task put an individual in the career fast lane, or is it a long-term parking area for under-powered or burned out executives?

4 Mintzberg, H. (1989) *Mintzberg on Management: Inside Our Strange World of Organizations*. New York: The Free Press.

Commitment by itself is not enough. An organisation must also possess the technical competence necessary to achieve enhanced safety. Have the hazards and safety-critical activities been identified? How many crises have been prepared for? Are crisis plans closely linked to business-recovery plans? Do the defences, barriers and safeguards possess adequate diversity and redundancy? Is the structure of the organisation sufficiently flexible and adaptive? Is the right kind of safety-related information being collected and analysed appropriately? Does this information get disseminated? Does it get acted upon? An effective safety information system is a prerequisite for a resilient system.[5]

Neither commitment nor competence will suffice unless the organisation is adequately cognisant of the dangers that threaten its activities. Cognisant organisations understand the true nature of the struggle for enhanced resilience. For them, a lengthy period without adverse events does not signal 'safe enough'. They see it correctly as a period of heightened danger and so review and strengthen their defences accordingly. In short, cognisant organisations maintain a state of intelligent wariness even in the absence of bad outcomes. This is the very essence of a safe culture.

Figure 14.3 summarises the argument so far. It also identifies the primary goal of safety management: to reach that region of the space associated with the maximally attainable level of intrinsic resistance – and then staying there. Simply moving in the right direction is relatively easy. But sustaining this goal state is very difficult. Maintaining such a position against the strong countervailing currents requires both a skilful use of navigational aids – the reactive and proactive measures – and a powerful cultural 'engine' that continues to exert its driving force regardless of the inclinations of the current leadership team. A good safety culture has to be CEO-proof. CEOs are, by nature, birds of passage: changing jobs frequently is how they got to

5 Kjellen, U. (1983) 'An evaluation of safety information systems of six medium-sized and large firms.' *Journal of Occupational Accidents*, 3: 273–288. Smith, M.J., Cohen, H., Cohen, A., and Cleveland, R.J. (1988) 'Characteristics of successful safety programs.' *Journal of Safety Research*, 10: 5–14

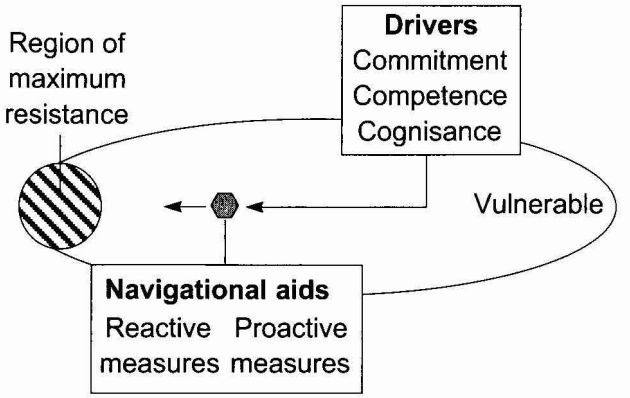

Figure 14.3 Summarising the driving forces and navigational aids necessary to propel an organisation towards the region of maximum resistance

where they are today – and there is no reason to suppose that they are going to behave any differently in the future.

Achieving this practicable safety goal depends very largely upon managing the manageable. Many organisations treat safety management as something akin to a negative production process. They set as targets the achievement of some reduced level of negative outcomes. But unplanned events, by their nature, are not directly controllable. So much of their variance lies outside the organisation's sphere of influence. The safety space model suggests an alternative approach: the long-term fitness programme. Rather than struggling vainly to reduce an already low and perhaps asymptotic level of adverse events, the organisation should regularly assess and improve those basic processes – design, hardware, maintenance, planning, procedures, scheduling, budgeting, communicating – that are known to influence the likelihood of bad events. These are the manageable factors determining a system's intrinsic resistance to its operational hazards. And they, in any case, are the things that managers are hired to manage. In this way, safety management becomes an essential part of the organisation's core business, and not just an add-on.

What Does a Resilient System Look Like?

Mapping the 3Cs on to the 4Ps

Earl Wiener, the eminent American human factors expert, devised the 4Ps (philosophy, policies, procedures and practices) framework[6] to differentiate the various aspects of management activity. I have borrowed the 4Ps here to present one axis of a 3 x 4 table (see Table 14.2). The other axis is made up of the 3Cs (commitment, cognisance and competence.

In each of the 12 cells, we are asking the question: how would each of the cultural drivers (the 3Cs) manifest itself in each of the 4Ps of organisational management? In Cell 1, for example, we are interested in how top-level commitment would reveal itself in the organisation's basic philosophy. In each cell, there are a set of indicators for the influence of the 3Cs upon the three Ps. Collectively, the indicators in the matrix provide a snapshot of what a resilient organisation might look like. The numbers below correspond to the cells in Table 14.2. Some of the contents of the

Table 14.2 Combining the 3Cs and the 4Ps to produce 12 sets of indicators

	Commitment	Cognisance	Competence
Principles (Philosophy			
Policies			
Procedures			
Practices			

6 Degani, A. and Wiener, E.L. (1994) 'The four "P"s of flight deck operation.' In N. Johnston, N. McDonald and R. Fuller (eds) *Aviation Psychology in Practice*. Aldershot: Avebury Technical.

cells have a health-care flavour, but these are readily generalised to other hazardous domains:

1. Principles and commitment:
   - Safety is recognised as being everyone's responsibility, not just that of the risk management team.
   - The organisation's mission statement makes safety a primary goal, and this is continually endorsed by the leadership's words, presence, actions and the allocation of resources.
   - Top management accepts errors, setbacks and nasty surprises as inevitable. It repeatedly reminds staff to be wary and vigilant.
   - Safety-related issues are considered at high-level meeting on a regular basis, not just after a bad event
2. Principles and cognisance:
   - Past events are thoroughly reviewed at high-level meetings and the lessons learned are implemented as global reforms rather than local repairs.
   - After some mishap, the primary aim of top management is to identify the failed system defences and improve them, rather than seeking to pin blame on specific individuals at the 'sharp end'.
   - It is understood that effective risk management depends critically upon the collection, analysis and dissemination of relevant safety-related information.
3. Principles and competence:
   - Top management adopts a proactive stance towards safety
     - strives to seek out and remove recurrent error traps,
     - eliminates error-provoking factors in the system,
     - brainstorms new scenarios of failure,
     - conducts regular 'health' checks on organisational 'vital signs'.
   - Top management recognises that error-provoking systemic factors are easier to correct than fleeting psychological states.
4. Policies and commitment:
   - Safety-related information has direct access to the top.
   - Safety management is fast-track not a long-term 'parking lot.'

- Meetings relating to safety are attended by staff from a wide variety of levels and departments.
- Schedulers and planners seek to ensure that teams remain intact when they are known to be effective and where conditions permit.

5. Policies and cognisance:
   - The organisation prioritises clinical goals over non-clinical demands on health-care staff wherever that is possible.
   - Policies are in place to reduce potential sources of non-clinical distraction in clinics, wards and operating theatres.
   - Policies ensure that senior staff are available and present throughout high-risk procedures.
6. Policies and competence:
   - Reporting system policies
     - qualified indemnity against sanctions,
     - confidentiality and/or de-identification,
     - separation of data collection from disciplinary procedures.
   - Disciplinary system policies
     - agreed distinction between acceptable and unacceptable behaviour,
     - peers involved in disciplinary proceedings.
7. Procedures and commitment:
   - The training of junior staff goes beyond the conventional apprenticeship system and procedures are in place to ensure that trainees reach pre-established competency criteria and receive adequate mentoring and supervision.
   - Procedures are in place within the system to facilitate the retraining and continuing professional development of senior staff, particularly with regard to new drugs and techniques.
8. Procedures and cognisance
   - Protocols backed by training in the recognition and recovery of errors.
   - Staff informed by feedback on recurrent error patterns.
   - Shift handovers are proceduralised to ensure adequate communication regarding local conditions.
   - Comparable procedures are in place to ensure safe transitions from the ward or operating theatre to the intensive care unit.

9. Procedures and competence:
   - Clinical supervisors train their charges in the mental as well as technical skills necessary to achieve safe and effective performance.
   - Clinical teams are briefed at the outset of complex or unusual procedures. And, where necessary, they are also debriefed afterwards.
   - The knowledge required to do a job should be shared between procedures, reminders and forcing functions.
10. Practices and commitment:
    - Safety-related issues are discussed by all staff whenever the need arises.
    - Nurses (in particular) should be discouraged from doing 'workarounds' to overcome (often chronic) systemic deficiencies.
    - Rather, they should be rewarded for bringing these problems to the attention of their line management.
11. Practices and cognisance:
    - Frontline personnel (nurses and junior doctors) should be provided with the tools and mental skills necessary to recognise high-risk situations.
    - Junior staff should be empowered to step back from situations for which they have been inadequately trained, where there is no local supervision, and where the conditions are highly error-provoking.
12. Practices and competence:
    - There should be rapid, useful and intelligible feedback on lessons learned and actions needed.
    - Bottom-up information should be listened to and acted upon where necessary.
    - Patient partnering and openness should be encouraged.
    - And, when mishaps occur ...
      - acknowledge responsibility,
      - apologise,
      - convince victims and their relatives that the lessons learned will reduce the chance of a recurrence.

The Knotted Rubber Band Model

The Model Applied to Some Continuous Control Process

What follows is an attempt to elucidate the phrase 'reliability is a dynamic non-event' using the mechanical properties of a rubber band as a model. We are concerned here with the actions of someone on the frontline of the system who has control over some process or piece of equipment.

Imagine a rubber band knotted in the middle. The knot represents the system-to-be-controlled and its spatial position is determined by the horizontal forces exerted on both ends of the band. Three configurations of the knotted rubber band are shown in Figure 14.4.

The stippled area in the centre of the diagram is the safe operating zone. The task of the controller is to keep the knot in this region by countering dangerous perturbations with appropriate compensatory corrections to the other end of the band. The top illustration in Figure 14.4 is a relatively stable state in which moderate and equal tensions on both ends of the band maintain the knot within the safety zone. The middle picture shows an unstable – or unsafe – condition in which an unequal force has been applied to one side of the band, pulling the knot out the safety zone. The bottom configuration depicts a corrected state in which the previous perturbation has been compensated for

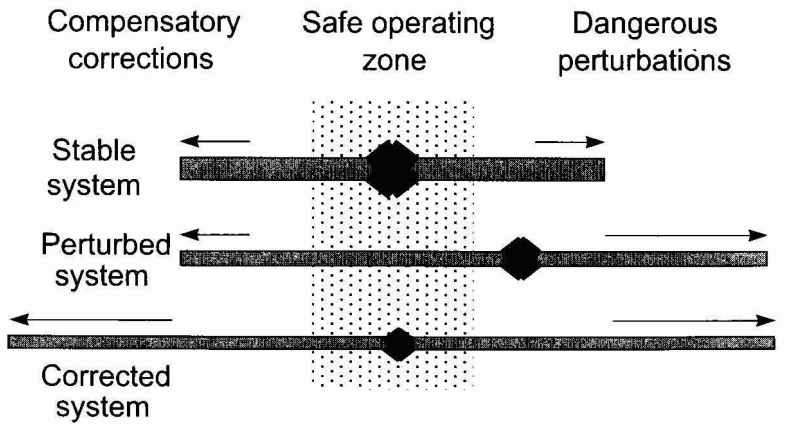

Figure 14.4 Three states of the knotted rubber band

by an equal pull in the opposite direction. There are, of course, many other states, but these are best appreciated by actually manipulating a knotted rubber band yourself.

The rubber band has a further important property. In order to maintain the position of the knot relative to the safety zone, it is necessary to apply an equal, opposite and *simultaneous* correction to any perturbation. Any delay in making this correction will take the knot outside of the safety zone, at least for a short while. I call this the *simultaneity principle.*

In applying this model to complex, highly automated technologies, such as nuclear power plants, chemical process plants and modern commercial aircraft, we should recognise that most of the foreseeable perturbations have already been anticipated by the designers and compensated for by the provision of engineered safety devices. These come into play automatically when the system parameters deviate from acceptable operational limits. This means that the large majority of the residual perturbations – those not anticipated by the designers – are likely to be due either to unexpected variations in local conditions, or to unforeseen actions on the part of the system's human elements – controllers, pilots, maintainers and the like. The latter are likely to include both errors and violations of safe operating procedures (see Chapters 3 and 4).

What are the consequences of the simultaneity principle for the human controllers of complex technologies, taking into account the nature of the residual perturbations just discussed? The first implication is that the timely application of appropriate corrections requires the ability to anticipate their occurrence. This, in turn, demands considerable understanding of what causes these perturbations. That is, it will depend upon the knowledge and experience of the human system controllers regarding, among other things, the roots of their own fallibility. As Weick has argued,[7] these qualities are more likely to be present in systems subject to fairly frequent perturbations (or in which periods of likely perturbation can be anticipated) than in stable systems in which the operating parameters remain constant for long periods of time. Clearly, there will be limits to this generalisation. Just

7 Weick, K.E. (1987) 'Organizational culture as a source of high reliability.' *California Management Review,* 19: 112–127.

as the inverted-U curve (the Yerkes-Dodson law) predicts that optimal human performance will lie between states of low and high arousal, we would similarly expect optimal system performance to lie between the extremes of virtual constancy and unmanageable perturbation.

Support for this view comes from field study observations of nuclear power generation, aircraft carrier flight deck operations and air traffic control.[8] In order to anticipate the conditions likely to provoke error, system operators need to experience them directly, learning from their own and other people's mistakes, as well during simulated training sessions. Error detection and error recovery are acquires skills and must be practised. This need to keep performance skills finely honed has been offered as an explanation for why ship-handlers manoeuvre closer to other vessels than is necessary in the prevailing seaway conditions.[9] Watchkeepers, it was suggested, gain important avoidance skills from such deliberately contrived close encounters.

The Model Applied to the Tension Between Productive and Protective Resources

Figure 14.5 shows the knotted rubber band in a different setting in order to demonstrate its resource implications. Every organisation needs to keep an optimal balance between production and protection (touched upon in Chapter 7 and discussed at length elsewhere).[10] The stippled region is now called the optimal operating zone and on either side there are protective and productive resources, represented as rectangles. The rubber band is a limited resource system. The more it is stretched, the less potential it has for controlling the rubber band – except, of course, by releasing the tension on one or other side.

Three configurations are shown in Figure 14.5. The top one is a balanced state in which the knot is centrally located with considerable potential for corrective action. Configuration A

8 Ibid.

9 Habberley, J.S., Shaddick, C.A., and Taylor, D.H. (1986) *A Behavioural Study of the Collision Avoidance Task in Bridge Watchkeeping*. Southampton: The College of Marine Studies.

10 Reason (1997), Chapter 1.

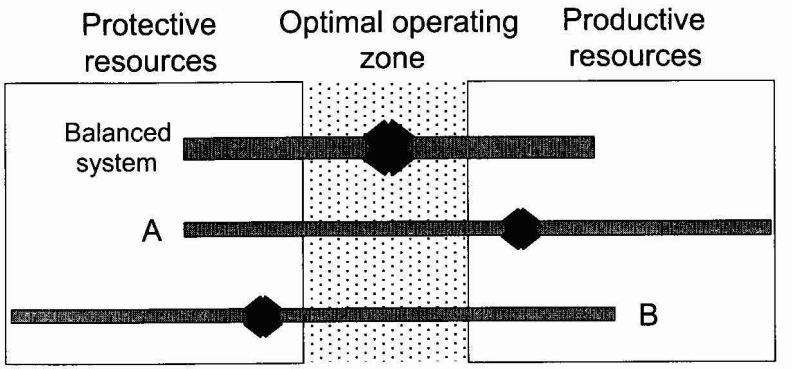

Figure 14.5 Showing the resource implications of the knotted rubber band model

shows an unbalanced state in which the pursuit of productive goals has pulled the knot out of the optimal zone. Configuration B is similarly out of balance, but in the opposite direction. Both configuration A and B have undesirable resource implications. Configuration A provides little or no possibility of compensating for some additional pull in the direction of productive goals, and is potentially dangerous. Configuration B, on the other hand, involves the unnecessary consumption of protective resources and so constitutes a serious economic drain upon the system. The risk in the former case is the unavailability of additional protective resources in the event of an increase in operational hazards; the risk in the latter case is, at the extreme, bankruptcy.

The Model Applied to the Diminution of Coping Abilities

The knotted rubber band has a further application derived from its capacity to become over-stretched and thus lose its potential for correcting the position of the knot. You will recall that in our discussion of the arterial switch operation (Chapter 9), it was noted that the ability of surgeons to compensate for adverse events was inversely related to the total number of events, both major and minor, that were encountered during the procedure. The implication was clear: coping resources are finite. They are used up by repeated stressors.

In the previous consideration of this phenomenon, I used Cheddar cheese to represent the limited coping resources, and a mouse that nibbled it away as representing the cumulative effects of the adverse events. But it is also possible to apply the knotted rubber band model. Let us assume that compensating for these events involves stretching the rubber band to neutralise each perturbation. Given enough of these events, the band becomes over-stretched and is unable to cope with these disturbances until the tension is released equally on both ends.

Defining the Nature of Positive Safety

The purpose of this concluding section is to combine the two models into a single view of safety that does not rely exclusively upon infrequent episodes of 'unsafety'. We will begin by summarising the main features of each model in turn.

Summarising the Properties of the Safety Space Model

- Both people and organisations differ not only in the frequency with which they suffer adverse events, but also in their intrinsic resistance to the hazards of their particular operations.
- It was argued that resistance was a determined rather than a random property. Unlike accidents, that have a large chance component in their causation, the factors contributing to the degree of intrinsic resistance are – to a much greater extent – under the control of those who manage and operate the system. These properties include such generic processes as forecasting, designing, specifying, planning, operating, maintaining, budgeting, communicating, proceduralising, managing, training and the like.
- Because of the chance element, even highly resistant systems can still experience negative outcomes. Safety is never absolute. There is no total freedom from danger. Conversely, even vulnerable systems can escape accidents for lengthy periods. Thus, the relationship between a system's resistance or vulnerability and its accident record, while generally positive over the long run, can be quite tenuous within any specific accounting period.

- Contrary to the spirit of most definitions of safety, it was argued that negative outcome data are imperfect, even misleading, indices of a system's state of intrinsic resistance. This is especially the case when the accident rate is very low or asymptotic – as it is in many contemporary industries.
- It was proposed that an organisation's current level of safety could be represented by its location within a cigar-shaped space, bounded at either end by high degrees of resistance and vulnerability to operational dangers.
- Any organisation is subject to external forces that act to push them away from both end of the space. If they were subject only to these 'tides and currents,' organisations would simply drift to and fro, moving from relative vulnerability to relative resistance, and then back again.
- It was argued that, for any organisation, the only attainable safety goal is not zero accidents, but to strive to reach the zone of maximum practicable resistance and then remain there for as long as possible. For this, each organisation requires both reliable navigational aids and some internal means of propulsion.
- The navigational aids comprise both reactive and proactive data: an effective safety information system that collects, analyses and disseminates information regarding accidents, incidents and near misses that is used in conjunction with regular diagnostic checks upon the system's 'vital signs' and the continuous improvement of those processes most in need of attention at any one time.
- An organisation's engine is essentially cultural. An ideal culture is one that continues to drive an organisation towards the resistant end of the space regardless of the commercial concerns of the current leadership. Three factors are seen to lie at the core of a safe culture: commitment, competence and cognisance (the 3Cs).
- Our consideration of the safety space model concluded with a 12-cell matrix in which indications that each of the 3Cs was influencing each of the 4Ps (principles, policies, procedures and practices) were listed. The matrix as a whole provided a snapshot summary of what a safe and resilient organisation might look like.

ıarising the Main Features of the Knotted Rubber Band Model

Echoing Weick, the model emphasised the dynamic character of the control actions required to keep a system in a reliable and stable state.

The mechanical properties of the rubber band model also highlighted the equal, opposite and simultaneous corrections necessary to keep the knot (the system) within the safe operating zone.

It was argued that most of the foreseeable perturbations (in complex, well-defended systems) will have been anticipated by the designers and corrected for automatically by engineered control devices. Those that remain are likely to arise from local variations and/or from unsafe acts. These human contributions can be both long-standing latent conditions, generated within the upper echelons of the organisation, and active failures (errors and violations) committed by those at the human–system interface.

In order to achieve the timely correction of these residual perturbations, the system operators must be able to recognise the conditions that foretell their occurrence. To do this, they need to have experienced them (in reality or simulation) and have developed the requisite error detection and correction skills.

- It follows from this assertion that operators of systems subject to relatively frequent disturbances are more likely to possess these skills than those who supervise comparatively stable systems. Systems in which these off-normal conditions are known through direct experience are likely to be safer than those in which this opportunity is largely denied.
- Like the systems it seeks to model, the knotted rubber band is resource-limited. Its corrective potential – beyond a certain point of necessary tension – is inversely related to the degree that it is stretched. When the band is extended near to its breaking point, the only way the knot can be moved is by reducing the tension on one or both sides.
- If it is assumed that the force exerted on one side of the band is primarily protective, while that acting on the other side is essentially productive, two unbalanced system states can be modelled. One is where productive forces hold the knot outside the optimal operating zone, and the other is the reverse situation in which excessive protective forces have been applied. Both states are potentially dangerous. In the first case, there is the risk of an uncorrected perturbation leading to a bad outcome. In the second, the risk is economic ruin due to an investment in protection that goes beyond that required to counter the operational hazards.
- A third application of the model was in regard to the limited capacity of the coping resources (see the surgeons in Chapter 9). If each perturbation requires a compensatory extension of the rubber band, it eventually becomes over-stretched and so lacks the capacity to counter further disturbances.

How can we integrate these two sets of features into a single coherent account of safety? The task is made easier by the fact that the two models address complementary but somewhat different levels of description. The emphasis of the safety space model is upon the broader strategic aspects of safety, while the rubber band model deals with the more tactical, moment-to-moment, control issues.

The safety space model defines the goal of safety management: the attainment and preservation of a state of maximum practicable resistance to operational hazards. It also indicates, in general terms, to achieve it: that is, the use of reactive and proactive navigational aids and the necessity of a cultural motive force. The dynamics of the knotted rubber band model, on the other hand, are more attuned to the local details of system control, most particularly with the need for anticipating error-provoking conditions, the timing of corrections and a suitable balance between the deployment of protective and productive resources.

Final Words

As I come to write the final words of this book, I am conscious of how ragged and inconclusive it is. I have provided you with little in the way of formulae or prescriptions for safer operation. But, at least, I hope you would be suspicious of anything I (or any consultant) might have offered in that regard. If someone tells you that they have a safe culture you will know to be deeply

suspicious, just as you might be if someone told you that they had achieved a state of grace. These are goals that have to be constantly striven for rather than achieved. It's the journey rather than the arrival that matters. Safety is a guerrilla war that you will probably lose (since entropy gets us all in the end), but you can still do the best you can.

I have greatly enjoyed writing about the heroic recoverers. Unfortunately, I have not given you much that could be 'bottled' and passed on to your workforce. Whatever it takes resides largely within very special people. Let's hope you have some such person (or people) when the occasion arises.

Index