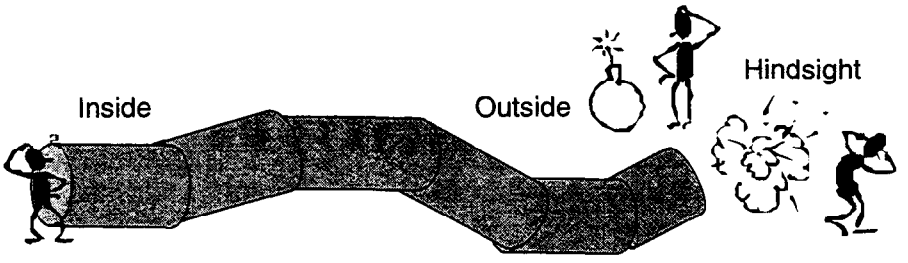
The Field Guide to Human Error

draft, August 2000



by Sidney Dekker

to be published by Cranfield University Press, early 2001
in collaboration with Ashgate Publishing Co.

Comments welcome to the author at: IKP, Linköping Institute of Technology, S-581 83 Sweden. sidde@ikp.liu.se. Fax +46-13-282579
If you want to refer to it: Dekker, S. W. A. (in press). The field guide to human error. Bedford, UK: Cranfield University Press.

Contents

Preface VII

PART I
Human error as a cause of mishaps

1. The Bad Apple Theory 3
2. Reacting To Failure 13
3. What Is The Cause? 29
4. Human Error By Any Other Name 35
5. Human Error—In The Head Or In The World? 51
6. Put Data Into Context 55

PART II
Human error as symptom of trouble deeper inside the system

7. Human Error—The New View 65
8. Reconstruct The Unfolding Mindset 71
9. Clues In The Rubble 87
10. Human Factors Data 99
11. Writing Recommendations 107
12. Learning From Failure 115

Subject Index 125

Preface

You are faced with an incident or accident that has a significant human contribution in it. What do you do? How do you make sense out of the controversial and puzzling assessments and actions by people? You basically have two options, and your choice determines the focus, questions, answers and ultimately the success of your probe, as well as the potential for progress on safety:

- You can see human error as the cause of trouble;
- You can see human error as the symptom of deeper trouble.

The first is called the old view of human error, while the second—itself already 25 years in the making—is the new view of human error.

The old view of human error	The new view of human error
Human error is a cause of accidents	Human error is a symptom of trouble deeper inside a system
To explain failure, you must seek failure. You must find people's: inaccurate assessments, wrong decisions, bad judgments.	To explain failure, do not try to find where people went wrong. Instead, find how people's assessments and actions made sense at the time, given the circumstances that surrounded them.

Table 0.1: Two views on human error

This Field Guide helps you reconstruct the human contribution to system failure according to the new view. In Part II, it presents a method for how to "reverse engineer" the evolving mindset of people who were caught up in a complex, unfolding situation. The Field Guide also wants to make you aware of the biases and difficulties in under-

standing past puzzling behavior—which is what Part I is about.

PART I OF THE FIELD GUIDE

The first six chapters of The Field Guide talk about the old view of human error—the problems it holds, the traps it represents, and the temptations that can make you fall in to them. These chapters help you understand:

- The bad apple theory: why throwing out a few bad apples does not get rid of the underlying human error problem;
- Reactions to failure: why the surprising nature of failure makes you revert easily to the bad apple theory;
- That there is no such thing as a root or primary cause: accidents are the result of multiple factors—each necessary and only jointly sufficient;
- That large psychological labels may give you the illusion of understanding human error but that they hide more than they explain;
- Why human error cannot be explained by going into the brain: you have to understand the situation in which behavior took place;
- Why human factors data need to be left in the context from which they came: cherry picking and micro-matching robs data of its original meaning.

PART II OF THE FIELD GUIDE

The last six chapters show you that human error is not necessarily something slippery or something hard to pin down. They show you how to concretely "reverse engineer" human error, just like any other component that needs to be put back together in a mishap investigation. It shows how to rebuild systematic connections between human behavior and features of the tasks and tools that people worked with, and of the operational and organizational environment in which they carried out their work. The Field Guide will encourage you to build a picture of:

- how a process and other circumstances unfolded around people;
- how people's assessments and actions evolved in parallel with their changing situation;
- how features of people's tools and tasks and organizational and operational environment influenced their assessments and actions.

The premise is that if you really understand the evolving situation in which people's behavior took place, you will understand the behavior that took place inside of it. Here is what the last six chapters talk about:

- Human error as a symptom of deeper trouble: connecting people's behavior with the circumstances surrounding them shows the sources of trouble and explains the behavior;
- A method for the reconstruction of people's unfolding mindset—this is the central part around which the rest of The Field Guide revolves;
- Where to look in the evidence for deeper clues about people's behavior: checking out the operational history, organizational environment and features of the technology they worked with;
- How and where to get human factors data: from historical sources, interviews and debriefings, and process recordings;
- Writing meaningful human factors recommendations;
- Learning from failure as ultimate goal of an investigation: failures represent opportunities for learning—opportunities that can fall by the wayside for a variety of reasons.

The Field Guide was born through participation in various incident and accident investigations. I want to thank those who alerted me to the need for this book and who inspired me to write it, in particular Air Safety Investigator Maurice Peters and Captain Örjan Goteman. It was written with support from the Swedish Flight Safety Directorate and Arne Axelsson, its director.

I am indebted to the following people for "the new view" on human error: David Woods, Erik Hollnagel, Edwin Hutchins, James Reason, John Flach, Gary Klein, Judith Orasanu, Diane Vaughan and Gene Rochlin. The ideas in The Field Guide are inspired by them and their ideas, although any misrepresentations or biases in this book are of course my responsibility.

S.D.
Linköping, Sweden
Summer 2000

1 The Bad Apple Theory

There are basically two ways of looking at human error. The first view could be called "the bad apple theory". It maintains that:

- Complex systems would be fine, were it not for the erratic behavior of some unreliable people (bad apples) in it;
- Human errors cause accidents: humans are the dominant contributor to more than two thirds of mishaps;
- Human error—by any other name (for example: loss of situation awareness, complacency, negligence)—explains system failures;
- Human errors come as an unpleasant surprises. They are unexpected and do not belong in the system. Errors are introduced to the system only through the inherent unreliability of people.

This chapter is about the first view, and the following five are about the problems and confusion that lie at its root.

A nation-wide debate about the death penalty is once again raging in the United States. Studies have found a system so fraught with vulnerabilities and error that some states are halting proceedings altogether, while others are scrambling to invest more in countermeasures against the executions of the innocent.

The debate is a window on people's beliefs about the sources of error. Says one protagonist: "The system of protecting the rights of accused is good. It's the people who are administring it who need improvement: The judges that make mistakes and don't permit evidence to be introduced. We also need improvement of the defense attorneys."[1] The system is basically safe, but it contains bad apples. Countermeasures against miscarriages of justice begin with them. Get rid of them, retrain them, discipline them.

But what is the practice of employing the least experienced, least skilled, least paid public defenders in many death penalty cases other than systemic? What are the rules for judges' permission of evidence other than systemic? What is the ambiguous nature of evidence other than inherent to a system that often relies on eyewitness accounts to make or break a case?

---

1 International Herald Tribune, 13 June 2000.

Each debate about error reveals two possibilities. Error is either the result of a bad apple, where disastrous outcomes could have been avoided if somebody had paid a bit more attention or made a little more effort. In this view, we wonder how we can cope with the unreliability of the human element in our systems.

Or errors are the inevitable by-product of people doing the best they can in systems that themselves contain multiple subtle vulnerabilities. These systems themselves are inherent contradictions between operational efficiency on the one hand and safety (for example: protecting the rights of the accused) on the other. In this view, errors are symptoms of trouble deeper inside a system.

Like debates about human error, investigations into human error mishaps face the choice. The choice between the bad apple theory in one of its many versions, or what has become known as the new view of human error.

Learning from failure

The ultimate goal of an investigation is to learn from failure. The road towards learning—the road taken by most investigations—is paved with intentions to follow the new view. Investigators intend to find the systemic vulnerabilities behind individual errors. They want to address the error-producing conditions that, if left in place, will repeat the same basic pattern of failure.

In practice, however, investigations often return disguised versions of the bad apple theory—in both findings and recommendations. They sort through the rubble of a mishap to:

- Find evidence for erratic, wrong or inappropriate behavior;
- Bring to light people's bad decisions, inaccurate assessments, deviations from written guidance;
- Single out particularly ill-performing practitioners.

Investigations often end up concluding how front-line operators failed to notice certain data, or did not adhere to procedures that appeared relevant after the fact. They recommend the demotion or retraining of particular individuals; the tightening of procedures or oversight. The reasons for regression into the bad apple theory are many. For example:

- Resource constraints on investigations. Findings may need to be produced in a few months time, and money is limited;
- Reactions to failure, which make it difficult not to be judgmental about seemingly bad performance;
- The hindsight bias, which confuses our reality with the one that surrounded the people we investigate;
- Political distaste of deeper probing into sources of failure, which may de facto limit access to certain data or discourage certain kinds of recommendations;

- Limited human factors knowledge on part of investigators. While wanting to probe the deeper sources behind human errors, investigators may not really know where or how to look.

In one way or another, The Field Guide will try to deal with these reasons. But it is foremost the lack of methodical guidance to reconstruct the human contribution to failure that allows investigations to relapse into the bad apple theory—a gap which The Field Guide intends to fill.

UNRELIABLE PEOPLE IN BASICALLY SAFE SYSTEMS

This chapter discusses the bad apple theory of human error. This way sees human error as a threat to systems that are basically safe. In this view on human error, progress on safety is driven by one unifying idea:

COMPLEX SYSTEMS ARE BASICALLY SAFE

THEY NEED TO BE PROTECTED FROM UNRELIABLE PEOPLE

Charges will be brought against the pilots who flew a VIP jet with a malfunction in its pitch control system (which makes the plane go up or down). Severe oscillations during descent killed seven of their un-strapped passengers in the back. Significant in the sequence of events was that the pilots "ignored" the relevant alert light in the cockpit as a false alarm, and that they had not switched on the fasten seatbelt sign from the top of descent, as recommended by jet's procedures. The pilot oversights were captured on video, shot by one of the passengers who died not much later. The pilots, wearing seatbelts, survived the upset.[1]

To protect safe systems from the vagaries of human behavior, recommendations typically propose to:

- Tighten procedures and close regulatory gaps. This reduces the bandwidth in which people operate, leaving less room for error;

---

1 FLIGHT International, 6-12 June 2000.

- Introduce more technology to monitor or replace human work. If machines do the work, then humans can no longer make errors doing it. And if machines monitor human work, they can snuff out any erratic human behavior;
- Make sure that defective practitioners (the bad apples) do not contribute to system breakdown again. Put them on "administrative leave"; demote them to a lower status; educate them to behave better next time; instill some fear in them and their peers by taking them to court or reprimanding them.

In this view of human error, investigations can safely conclude with the label "human error"—by whatever name (for example: ignoring a warning light, violating a procedure). Such a conclusion and its implications supposedly get to the causes of system failure.

AN ILLUSION OF PROGRESS ON SAFETY

The shortcomings of the bad apple theory are severe and deep. Pro-gress on safety based on this view is an illusion.

Throwing out the bad apples

For example, focusing on individual failures does not take away the underlying problem. Removing "defective" practitioners fails to remove the potential for the errors they made.

As it turns out, the VIP jet aircraft had been flying for a long time with a malfunctioning pitch feel system ('Oh that light? Yeah, that's been on for four months now'). These pilots inherited a systemic problem from the airline that operated the VIP jet, and the organization charged with its maintenance.

Adding more procedures

Adding or enforcing procedures does not guarantee compliance:

Seatbelt sign on from top of descent in a VIP jet? The layout of furni-ture in these machines and the way in which their passengers are

pressured to make good use of their time by meeting, planning, working, discussing, does every-thing to discourage people from strapping in any earlier than strictly neces-sary. Pilots can blink the light all they want, you could understand that over time it may become pointless to switch it on from 41,000 feet on down.

And who typically employs the pilot of a VIP jet? The person in the back. So guess who can tell who what to do. And why have the light on only from the top of descent? This is hypocritical—only in the VIP jet upset discussed here was that relevant because loss of control occurred during descent. But other incidents with in-flight deaths have occurred during cruise. Procedures are insensitive to this kind of natural variability.

New procedures can also get buried in masses of regulatory paperwork. Mismatches between procedures and practice grow not necessarily because of people's conscious non-adherence but because of the amount and increasingly tight constraints of procedures.

The vice president of a large airline commented recently how he had seen various of his senior colleagues retire over the past few years. Almost all had told him how they had gotten tired of updating their aircraft operating manuals with new procedures that came out—one after the other—often for no other reason than to close just the next gap that had been revealed in the latest little incident. Faced with a growing pile of paper in their mailboxes, they had just not bothered. Yet these captains all retired alive and probably flew very safely during their last few years.

Adding a bit more technology

More technology does not remove the potential for human error, but relocates or changes it.

A warning light does not solve a human error problem, it creates new ones. What is this light for? How do we respond to it? What do we do to make it go away? It lit up yesterday and meant nothing. Why listen to it today?

WHY IS THE BAD APPLE THEORY POPULAR?

Cheap and easy

So why would anyone adhere to the bad apple theory of human error? There are many reasons. One is that it is a relatively straightforward approach to dealing with safety. It is simple to understand and simple, and relatively cheap, to implement.

Saving face

In the aftermath of failure, pressure can exist to save public image. Taking out defective practitioners is always a good start to saving face. It tells people that the mishap is not a systemic problem, but just a local glitch in an otherwise smooth operation.

Two hard disks with classified information went missing from the Los Alamos nuclear laboratory, only to reappear under suspicious circumstances behind a photocopier a few months later. Under pressure to assure that the facility was secure and such lapses extremely uncommon, the Energy Secretary attributed the incident to "human error, a mistake". The hard drives were probably misplaced out of negligence or inattention to security procedures, officials said. The Deputy Energy Secretary added that "the vast majority are doing their jobs well at the facility, but it probably harbored "a few bad apples" who had compromised security out of negligence.[1]

Personal responsibility and the illusion of omnipotence

Another reason to adhere to the bad apple theory of human error is that practitioners in safety-critical domains typically assume great personal responsibility for the outcomes of their actions. Practitioners get trained and paid to carry this responsibility, and are proud of it.

---

1 International Herald Tribune, 19 June 2000.

But the other side of taking this responsibility is the assumption that one has the authority, the power, to match it; to live up to it. The assumption is that people can simply choose between making errors and not making them—independent of the world around them. This, however, is an illusion of omnipotence. It is commonly entertained by children in their pre-teens, and by the airline captain who said, "If I didn't do it, it didn't happen."

Investigators are often practitioners themselves or have been practitioners, which makes it easy to overestimate the freedom of choice allotted to fellow practitioners.

The pilot of an airliner accepted a different runway with a more direct approach to the airport. The crew got in a hurry and made a messy landing that resulted in some minor damage to the aircraft. Asked why they accepted the runway, the crew cited a late arrival time and many connecting passengers on board. The investigator's reply was that real pilots are of course immune to those kinds of pressures.

The reality is that people are not immune to those pressures, and the organizations that employ them would not want them to be. People do not operate in a vacuum, where they can decide and act all-powerfully. Instead, their work is subject to and constrained by factors more or less outside of their control. Individual responsibility is not always matched by individual authority. Authority is restricted by other people or parts in the system, by other pressures, other deficiencies.

In the VIP jet's case, it was found that there was no checklist that told pilots what to do in case of a pitch feel indication light. The procedure to avoid the oscillations would have been to reduce airspeed to less than 260 knots indicated. But the procedure was not in any manual. It was not available in the cockpit. And it's hardly the kind of thing you can think up on the fly.

WHAT IS NOT RIGHT WITH THIS PICTURE?

Something was not right with the picture of the VIP jet from the start. How, really, could anyone claim that pilots "ignored" a light for which there was no procedure available? You cannot "ignore" something if you do not know what to do with it. Factors from the outside seriously constrained what the pilots could have possibly done. Problems existed with this particular aircraft. No procedure was available to deal with the warning light.

Any picture of human error is probably not right or not complete if it contains a generous helping of negligence or complacency; a large measure of people not motivated to try hard enough.

In 1995 Srebrenica, a Muslim town in Bosnia, was captured by Bosnian Serbs during one of the Post-Yugoslavian wars. The town had been nominally guarded by a contingent of a few hundred Dutch peacekeepers. Once the town had fallen, the Bosnian Serbian army massacred thousands of Muslims with impunity. Media at the time were fond of portraying a complacent stereotype of the Dutch army—long haired, bearded, marijuana-smoking—as if this would explain the events at Srebrenica.

The actual story behind this failure revealed fundamental shortcomings pervading the entire peacekeeping operation and organization. United Nations mandates were ambiguous and limited, lines of command confused by distribution across multiple countries and services, and supplies of material and manpower woefully inadequate, leaving soldiers on the ground effectively with their hands tied behind their backs. Almost a decade on, the debate about problems and vulnerabilities in peacekeeping still reverberates at UN headquarters. Also, times in Srebrenica had been trying. Warring parties displayed no willingness to heed UN treaties—peacekeepers were blocked, abducted, robbed and murdered. Infighting between Muslim warlords had also blurred the traditional "good guy—bad guy" portrait and undermined any possible opposition against a highly determined Serbian Bosnian onslaught.

Whatever label is in fashion (complacency, negligence, ignorance), if a human error picture makes sense by relying on "bad apples" who lack the motivation to

perform better, it is probably missing the real story behind failure, or at least large parts of it.

Local rationality

The point is, people in safety-critical jobs are very likely doing the right thing under the circumstances. They are doing reasonable things given their point of view and focus of attention; their limited knowledge of the situation; their objectives.

Yes, they do want to be bothered. They do want to pick up and integrate the data that are critical. They do want to do things the right way. People in these jobs do not go out of their way to hurt or kill other people, or hurt or kill themselves.

But these people are also concerned with other objectives existent in their jobs—the pressures to produce; to not cost their organization unnecessary money; to be on time; to get results; to keep customers happy. Their sensitivity to these factors, and their ability to juggle them in parallel with demands for safety, is one reason why they were chosen for the jobs, and why they are allowed to keep them.

In the Los Alamos nuclear research facility, complacency was no longer a feature of a few individuals—if it ever had been. Under pressure to perform daily work in a highly cumbersome context of checking, double-checking and registering the use of sensitive materials, "complacency" (if one could still call it that) had become a feature of the entire laboratory. Scientists routinely moved classified material without witnesses or signing logs. Doing so was not a sign of malice. The practice had grown over time, bending to production pressures from which the laboratory owed its existence.[1]

The safe assumption to make is that people were doing reasonable things given the circumstances. They were doing their best given the complexities, dilemmas, trade-offs and uncertainty that surrounded them. Understanding critical features of the circumstances in which people worked, and had worked for a while, will help you understand the behavior inside those situations much better.

---

1 International Herald Tribune, 20 June 2000.

2. Reacting To Failure

Have you ever caught yourself asking, "How could they not have noticed?", or, "How could they not have known?"? Then you were reacting to failure.

TO UNDERSTAND FAILURE, WE MUST FIRST UNDERSTAND OUR REACTIONS TO FAILURE

We all react to failure. In fact, our reactions to failure often make us see human error as the cause of a mishap; they promote the bad apple theory. Failure, or people doing things with the potential for failure, is generally not something we expect to see. It surprises us; it does not fit our assumptions about the system we use or organization we work in. It goes against our beliefs and views. As a result, we try to reduce that surprise—we react failure. Reactions share the following reatures:

- Retrospective. Reactions arise from our ability to look back on a sequence of events, of which we know the outcome;
- Proximal. They focus on those people who were closest in time and space to causing or potentially preventing the mishap;
- Counterfactual. They lay out in detail what these people could have done to prevent the mishap;
- Judgmental. They say what people should have done, or failed to do, to prevent the mishap.

Reactions to failure interfere with our understanding of failure. Yet fin-dings and conclusions about human error are often driven by reactions to failure, and written in their language.

RETROSPECTIVE

Looking back on a sequence of events, knowing the outcome

INVESTIGATIONS AIM TO EXPLAIN A PART OF THE PAST

YET ARE CONDUCTED IN THE PRESENT, AND THUS INEVITABLY INFLUENCED BY IT

As investigator, you are likely to know:

- The outcome of a sequence of events you are investigating;
- Which cues and indications were critical in the light of the outcome—what were the signs of danger?
- Which assessments and actions would have prevented the outcome.

A highly automated airliner crashed on a golf course short of the runway at an airport in India. During the final approach, the aircraft's automation had been in "open descent mode", which manages airspeed by pitching the nose up or down, rather than through engine power. When they ended up too low on the approach, the crew could not recover in time. In hindsight, the manufacturer of the aircraft commented that "the crew should have known they were in open descent mode". Once outside observers learned its importance, the question became how the crew could have missed or miss-understood such a critical piece of information.

One of the safest bets you can make as an investigator or outside observer is that you know more about the incident or accident than the people who were caught up in it—thanks to hindsight:

- Hindsight means being able to look back, from the outside, on a sequence of events that led to an outcome you already know about;
- Hindsight gives you almost unlimited access to the true nature of the situation that surrounded people at the time (for example: where they actually were versus where they thought they were; what state their system was in versus what they thought it was in, etc.);
- Hindsight allows you to pinpoint what people missed and shouldn't have missed; what they didn't do but should have done.

Hindsight biases your investigation towards items that you *now* know were important ("open descent mode"). As a result, you may assess people's decisions and actions mainly in the light of their failure to pick up this critical piece of data. It artificially narrows your examination of the evidence and potentially misses alternative or wider explanations of people's behavior.

Inside the tunnel

Look at figure 2.1. You see an unfolding sequence of events there. It has the shape of a tunnel which is meandering its way to an outcome. The figure shows two different perspectives on the pathway to failure:

- **The perspective from the outside and hindsight** (typically your perspective). From here you can oversee the entire sequence of events—the triggering conditions, its various twists and turns, the outcome, and the true nature of circumstances surrounding the route to trouble.
- **The other perspective sees only the inside of the tunnel**. This is the point of view of the person in the unfolding situation. To that someone, the outcome was not known. That someone changed the direction of the sequence of events on the basis of what he or she saw on the *inside* of the unfolding situation. To understand human error, you need to attain this perspective.

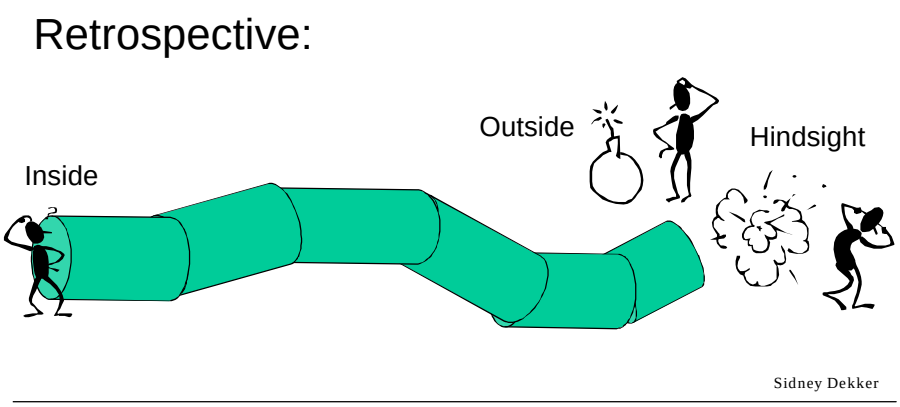


Fig. 2.1: Different perspectives on a sequence of events: Looking from the outside and hindsight you have knowledge of the outcome and dangers involved. From the inside, you may have neither.

The Field Guide invites you to go inside the tunnel of figure 2.1. It will help you understand the evolving situation from the point of view of the people inside of it, and to try to see why their assessments and actions would have been reasonable and seemed right at the time.

Hindsight is everywhere

Hindsight is baked deeply into the language of accident stories we tell one another. Take a common problem today—people losing track of what mode their automated systems are operating in. This happens in cockpits, operating rooms,

process control plants and many other workplaces. In hindsight, when you know how things developed and turned out, this problem is often called "losing mode awareness". Or, more broadly, "loss of situation awareness". What are we really saying? Look at figure 2.2. Loss of situation awareness is the difference between:

- what you *now* know the situation actually was like;
- what people understood it to be at the time.

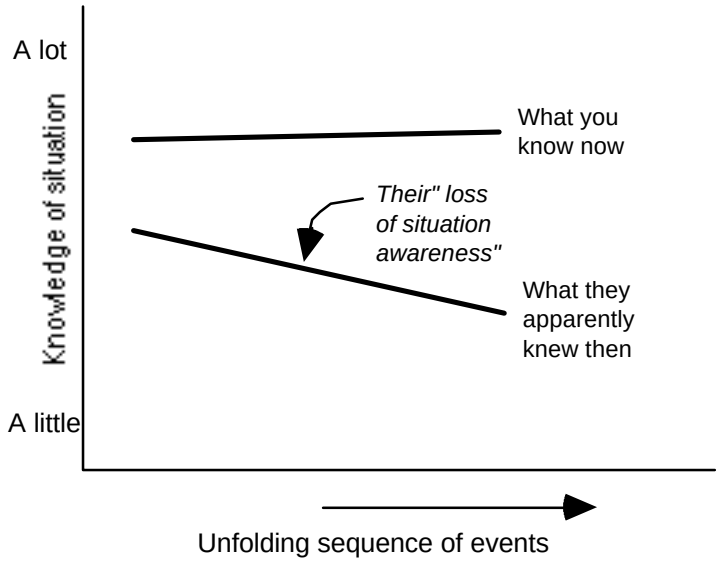


Fig. 2.2: Hindsight is everywhere. Here, "loss of situation awareness" as the difference between your knowledge today of which aspects in the situation were critical, and what people apparently knew then.

It is easy to show that people from another time and place did not know what you know today ("they should have known they were in open descent mode"). But it is not an explanation of their behavior.

You must guard yourself against mixing your reality with the reality of the people you are investigating. Those people did not know there was going to be a negative outcome (or they would have done something else). So it would have been impossible for them to assess—in the way that you can today—which data or decisions were critical in the light of it.

PROXIMAL

Focusing on people at the sharp end

Reactions to failure focus firstly and predominantly on those people who were closest to producing and to potentially avoiding the mishap. It is easy to see these people as the engine of action. If it were not for them, the trouble would not have occurred.

Someone called me on the phone, demanding to know how it was possible that a train drivers ran red lights. Britain had just suffered one of its worst rail disasters—this time at Ladbroke Grove near Paddington station in London. A commuter train had run head-on into a high-speed intercity coming from the other direction. Many travelers were killed in the crash and ensuing fire. The investigation returned a verdict of "human error". The driver of the commuter train had gone right underneath signal 109 just outside the station, and signal 109 had been red, or "unsafe". How could he have missed it? A photograph published around the same time showed sensationally how another driver was reading a newspaper while driving his train.

Blunt end and sharp end

In order to understand error, you have to examine the larger system in which these people worked. You can divide an operational system into a sharp end and a blunt end:

- At the sharp end (for example the train cab, the cockpit, the surgical operating table), people are in direct contact with the safety-critical process;
- The blunt end is the organization or set of organizations that supports and drives and shapes activities at the sharp end (for example the airline or hospital; equipment vendors and regulators).

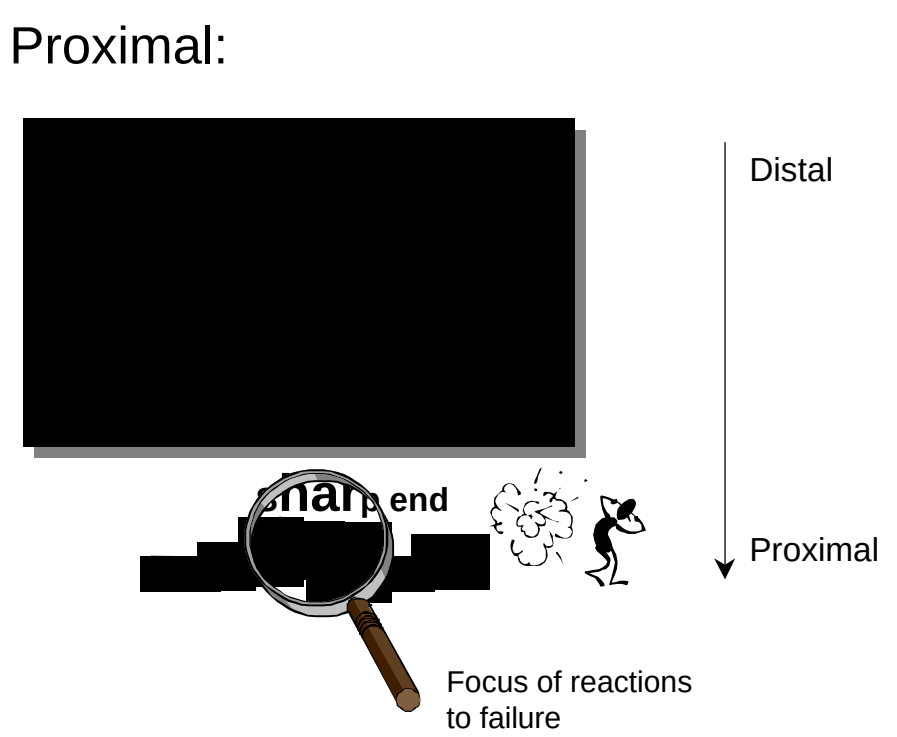


Fig. 2.3: Failures can only be understood by looking at the whole system in which behavior took place. But in our reactions to failure, we often focus on the sharp end, where people were closest to causing or potentially preventing the mishap.

The blunt end gives the sharp end resources (for example equipment, training, colleagues) to accomplish what it needs to accomplish. But at the same time it puts on constraints and pressures ("don't be late, don't cost us any unnecessary money, keep the customers happy"). Thus the blunt end shapes, creates, and can even encourage opportunities for errors at the sharp end. Figure 2.3 shows this flow of causes through a system. From blunt to sharp end; from upstream to downstream; from distal to proximal. It also shows where the focus of our reactions to failure is trained: on the proximal.

Why do people focus on the proximal?

Looking for sources of failure far away from people at the sharp end is counterintuitive. And it can be difficult. If you find that sources of failure lie really at the blunt end, this may call into question beliefs about the safety of the

entire system. It challenges previous views. Perhaps things are not as well-organized or well-designed as people had hoped. Perhaps this could have happened any time. Or worse, perhaps it could happen again.

The Ladbroke Grove verdict of "driver error" lost credibility very soon after it came to light that signal 109 was actually a cause célèbre among train drivers. Signal 109 and the entire cluttered rack on which it was suspended together with many other signals, were infamous. Many drivers had passed an unsafe signal 109 over the preceding years and the drivers' union had been complaining about its lack of visibility.

In trains like the one that crashed at Ladbroke Grove, automatic train braking systems (ATB) had not been installed because they had been considered too expensive. Train operators had grudgingly agreed to install a "lite" version of ATB, which in some sense relied as much on driver vigilance as the red light itself did.

Reducing surprise by pinning failure on local miscréants

Some people and organizations see surprise as an opportunity to learn. Failures offer them a window through which they can see the true internal workings of the system that produced the incident or accident. These people and organizations are willing to change their views, to modify their beliefs about the safety or robustness of their system on the basis of what the system has just gone through. This is where real learning about failure occurs, and where it can create lasting changes for the good. But such learning does not come easy. And it does not come often. Challenges to existing views are generally uncomfortable. Indeed, for most people and organizations, coming face to face with a mismatch between what they believed and what they have just experienced is difficult. These people and organizations will do anything to reduce the nature of the surprise.

It seems common among fighter pilots across the world to trash the reputation of a comrade who has just been killed in an accident. Sociologists have observed how his or her fellow pilots go to the bar and drink to the fallen comrade's misfortune, or more likely his or her screw-up, and put the drinks on his or her bar tab. This practice is aimed at highlighting or inventing evidence for why s/he wasn't such a good pilot after all. The transformation from "one of themselves" into "a bad pilot" psychologically shields those who do the same work from equal vulnerability to failure.

People and organizations often want the surprise in the failure to go away, and with it the challenge to their views and beliefs. The easiest way to do this is to see

the failure as something local, as something that is merely the problem of a few individuals who behaved in uncharacteristic, erratic or unrepresentative (indeed, locally "surprising") ways.

Potential revelations about systemic vulnerabilities were deflected by pinning failure on one individual in the case of Oscar November[1]. Oscar November was one of the airline's older Boeing 747 "Jumbojets". It had suffered earlier trouble with its autopilot, but on this morning everything else conspired against the pilots too. There had been more headwind than forecast, the weather at the destination was very bad, demanding an approach for which the co-pilot was not qualified but granted a waiver, while the co-pilot (and flight engineer) were actually severely afflicted by gastrointestinal infection. Air traffic control turned the big aircraft onto a tight final approach, which never gave the old autopilot enough time to settle down on the right path. The aircraft narrowly missed a building near the airport, which was shrouded in thick fog. On the next approach it landed without incident.

Oscar November's captain was taken to court to stand trial on criminal charges of "endangering his passengers" (something pilots do every time they fly, one fellow pilot quipped). The case centered around the crew's "bad" decisions. Why hadn't they diverted to pick up more fuel? Why hadn't they thrown away that approach earlier? Why hadn't they gone to another arrival airport? These questions trivialized or hid the organizational and operational dilemma's that confront crews all the time. The focus on customer service and image; the waiving of qualifications for approaches; putting more work on qualified crewmembers; heavy traffic around the arrival airport and subsequent tight turns; trade-offs between diversions in other countries or continuing with enough but just enough fuel. And so forth.

The vilified captain was demoted to co-pilot status and ordered to pay a fine. He committed suicide soon thereafter. The airline, however, had saved its public image by focusing on a single individual who—the court showed—had behaved erratically and unreliably.

Potentially disruptive lessons about the system as a whole are transformed into isolated hick-ups by a few uncharacteristically ill-performing individuals. It relieves the larger organization of any need to change views and beliefs, or associated policies or spending practices. The system is safe, if only it weren't for a few unreliable humans in it.

---

1 Wilkinson, S. (1994). The Oscar November Incident. Air & Space, February-March.

FACED WITH A BAD, SURPRISING EVENT, WE CHANGE THE EVENT OR THE PLAYERS IN IT—

RATHER THAN OUR BASIC BELIEFS ABOUT THE SYSTEM THAT MADE THE EVENT POSSIBLE

Instead of modifying our views in the light of the event, we re-shape, re-tell and re-inscribe the event until it fits the traditional and non-threatening view of the system. As far as organizational learning is concerned, the mishap might as well not have happened. The proximal nature of our reactions to failure makes that expensive organizational lessons can go completely unlearned.

The pilots of a large military helicopter that crashed on a hillside in Scotland in 1994 were found guilty of gross negligence. The pilots did not survive—29 people died in total—so their side of the story could never be heard. The official inquiry had no problems with "destroying the reputation of two good men", as a fellow pilot put it. Indeed, many other pilots felt uneasy about the conclusion. Potentially fundamental vulnerabilities (such as 160 reported cases of Uncommanded Flying Control Movement or UFCM in computerized helicopters alone since 1994) were not looked into seriously.[1]

COUNTERFACTUAL

Finding out what could have prevented the mishap

The outcome of a sequence of events is the starting point of your work as investigator. Otherwise you wouldn't actually be there. This puts you at a remarkable disadvantage when it comes to understanding the point of view of the people you're investigating. Tracing back from the outcome, you will come across joints where people had opportunities to "zig" instead of "zag"; where they could have directed the events away from failure. As investigator you come out on the other end of the sequence of events wondering how people could have missed those opportunities to steer away from failure.

---

[1] Sunday Times, 25 June 2000.

Accident reports are generally full of counterfactuals that describe in fine detail the pathways and options that the people in question did not take. For example, "The airplane could have overcome the windshear encounter if the pitch attitude of 15 degrees nose-up had been maintained, the thrust had been set to 1.93 EPR (Engine Pressure Ratio) and the landing gear had been retracted on schedule"[1]

Counterfactuals prove what could have happened if certain minute and often utopian conditions had been met. Counterfactual reasoning may thus be a fruitful exercise when recommending countermeasures against such failures in the future.

But when it comes to explaining behavior, counterfactuals contribute little. Stressing what was not done (but if it had been done, the accident wouldn't have happened) explains nothing about what actually happened, or why. Counterfactuals are not opportunities missed by the people you are investigating. Counterfactuals are products of your hindsight. Hindsight allows you to transform a uncertain and complex sequence of events into a simple, linear series of obvious options. By stating counterfactuals, you are probably oversimplifying the decision problems faced by people at the time.

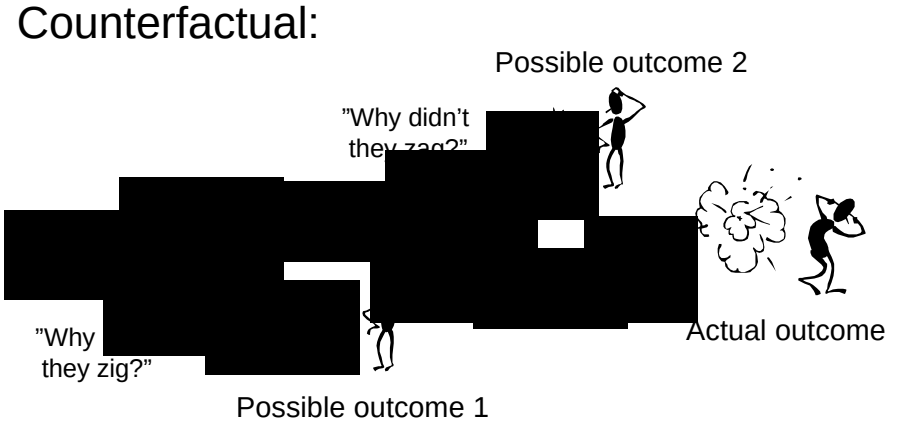


Fig. 2.4: Counterfactuals: Going back through a sequence, you wonder why people missed opportunities to direct events away from the eventual outcome. This, however, does not explain failure.

[1] National Transportation Safety Board (1995). *Aircraft Accident Report: Flight into terrain during missed approach USAir flight 1016, DC-9-31, N954VJ, Charlotte North Carolina, July 2, 1994*. Washington, DC: NTSB, page 119.

Forks in the road stand out so clearly to you, looking back. But when inside the tunnel, when looking forward and being pushed ahead by unfolding events, these forks were shrouded in the uncertainty and complexity of many possible options and demands; they were surrounded by time constraints and other pressures.

JUDGMENTAL

Saying what they should have done, or failed to do

To explain failure, we seek failure. In order to explain why a failure occurred, we look for errors, for incorrect actions, flawed analyses, inaccurate perceptions. When you have to explain failure, wrong judgments, inaccurate perceptions and missed opportunities would seem like a good place to start.

Yet these decisions, judgments, perceptions are bad or wrong or inaccurate only from hindsight—from your point of view as retrospective outsider. When viewed from the inside of a situation, decisions, judgments and perceptions are just that: decisions, judgments and perceptions.

Look at figure 2.5. The very use of the word "failure" in investigative conclusions (for example: "the crew failed to recognize a mode shift") indicates that you are still on the top line, looking down. It represents a judgment from outside the situation, not an explanation from people's point of view within.

The word failure implies an alternative pathway, one which the people in question did not take (for example, recognizing the mode change). Laying out this pathway is counterfactual, as explained above.

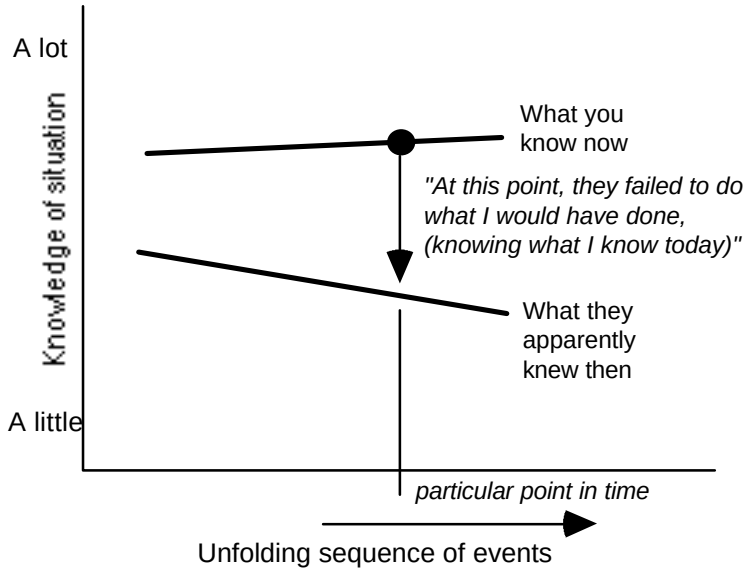


Fig. 2.5: Judgmental: saying that other people failed to do what they should have done (knowing what you know today) does not explain their behavior.

But by saying that people "failed" to take this pathway—in hindsight the right one—you judge their behavior according to a standard you can impose only with your broader knowledge of the mishap, its outcome and the circumstances surrounding it. You have not explained a thing yet. You have not shed light on how things looked on the inside of the situation; why people did what they did given *their* circumstances.

The literature on medical error describes how cases of death due to negligence may be a result of a judgment failure in the diagnostic or therapeutic process. Examples include a misdiagnosis in spite of adequate data, failure to select appropriate diagnostic tests or therapeutic procedures, and delay in diagnosis or treatment.[1]

Although they look like explanations of error, they are in fact judgments that carry no explanation at all. For example, the "misdiagnosis in spite of adequate data" was once (before hindsight) a reasonable diagnosis based on the data that seemed critical or relevant—otherwise it would not have been made by the physician in question. Calling it a misdiagnosis is an unconstructive, retrospective judgment that misses the reasons behind the actual diagnosis.

[1] Bogner, M: S. (Ed.) (1994). *Human error in medicine*. Hillsdale, N.J: Erlbaum.

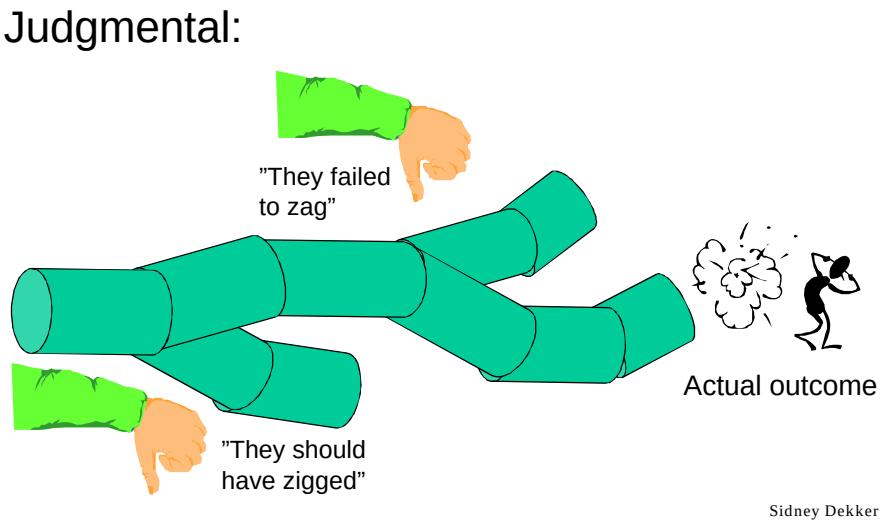


Fig. 2.6: Judgmental: by claiming that people should have done something they didn't, or failed to do something they should have, you do not explain their actual behavior.

The illusion of cause-consequence equivalence

One reason why people feel compelled to judge instead of explain—why they look for failure to explain failure—has to do with "cause-consequence equivalence".

BAD OUTCOME = BAD PROCESS

We assume that really bad consequences can only be the result of really bad causes. Faced with a disastrous outcome, or the potential for one, we assume that the acts leading up to it must have been equally monstrous. Once we know an outcome is bad, we can no longer look objectively at the process that led up to it.

But this automatic response is very problematic in complex worlds. Here even bad processes often lead to good outcomes. And good processes can lead to bad outcomes. Processes may be "bad" in the retrospecitve sense that they departed from routines you now know to have been applicable. But this does not necessarily lead to failure. Given their variability inherent to these worlds, they typically allow an envelope of options and pathways to safe outcomes. There is more than one way to success. Think of a rushed approach in an aircraft that becomes stabilized at the right time and leads to a safe landing. The opposite goes too. Good processes (in the sense that they do not depart from the drill), where people double-check and communicate and stick to procedures, can lead to disastrous

outcomes.

BAD PROCESSES MOSTLY LEAD TO GOOD OUTCOMES

GOOD PROCESSES SOMETIMES LEAD TO BAD OUTCOMES

Think for example of an inflight fire or other serious malfunction where pilots negotiate between landing overweight or dumping fuel (two things you simply can't do at the same time), while sorting through procedures that aim to locate the source of trouble—in other words, doing what the book and training and professional discipline tells them to do. If the fire or malfunction catches up with the pilots while they are still airborne, you may say that they should have landed instead of bothered with anything else. But it is only hindsight that allows you to say that. You cannot really judge the pilots' process of double-checking and negotiation to be bad by any other standard.

FAILURES AS THE BY-PRODUCT OF NORMAL WORK

What is striking about many accidents is that people were doing exactly the sorts of things they would usually be doing—the things that usually lead to success and safety. In the sequence of events leading up to failures, there is no "badness" in anybody's behavior by any objective measure. People are doing what makes sense given the situational indications, operational pressures and organizational norms existing at the time.

If this is the most profound lesson you and your organization can learn from a mishap, it is also the most frightening. The difficulty of accepting this reality lies behind our reactions to failure. Going beyond reacting to failure means acknowledging that failures are baked into the very nature of your work and organization; that they are symptoms of deeper trouble or by-products of systemic brittleness in the way you do your business. It means having to acknowledge that mishaps are the result of everyday influences on everyday decision making, not isolated cases of erratic individuals behaving unrepresentatively. Going beyond your reactions to failure means having to find out why what people did back there and then actually made sense given the organization and operation that surrounded them.

3. What Is The Cause?

What was the cause of the mishap? In the aftermath of failure, no question seems more pressing. There can be significant pressure from all kinds of directions to pinpoint a cause:

- People want to start investing in countermeasures;
- People want to know how to adjust their behavior to avoid the same kind of trouble;
- People may simply seek retribution, punishment, justice.

The problem is, there is no such thing as *the* cause of a mishap. And just sorting through the rubble will not necessarily guide you to one either.

THE CONSTRUCTION OF CAUSE

Look at two official investigations into the same accident. One was conducted by the airline whose aircraft crashed somewhere in the mountains. The other was conducted by the civil aviation authority of the country in which the accident occurred, and who employed the air traffic controller in whose airspace it took place.

The authority says that the controller did not contribute to the cause of the accident, yet the airline claims that air traffic control clearances were not in accordance with applicable standards and that the controller's inadequate language skills and inattention were causal. The authority counters that the pilot's inadequate use of flightdeck automation was actually to blame, whereupon the airline points to an inadequate navigational database supplied to their flight computers among the causes. The authority explains that the accident was due to a lack of situation awareness regarding terrain and navigation aids, whereas the airline blames lack of radar coverage over the area. The authority states that the crew failed to revert to basic navigation when flight deck automation usage created confusion and workload, whereupon the airline argues that manufacturers and vendors of flightdeck automation exuded overconfidence in the capabilities of their technologies and passed this on to pilots. The authority finally blames ongoing efforts by the flight crew to expedite their approach to the airport in order to avoid delays, whereupon the airline lays it on the controller for suddenly

inundating the flight crew with a novel arrival route and different runway for landing.[1]

Causes according to Authority:	Causes according to Airline:
Air Traffic Controller did not play a role	No standard phraseology, inadequate language and inattention by Controller
Pilots' inadequate use of automation	Inadequate automation database
Loss of pilots' situation awareness	Lack of radar coverage over area
Failure to revert to basic navigation	Overconfidence in automation sponsored by vendors
Efforts to hasten arrival	Workload increase because of Controller's sudden request

Table 3.1: Two statements of cause about the same accident

So who is right? The reality behind the controversy, of course, is that both investigations are right. They are both right in that all of the factors mentioned were in some sense causal, or contributory, or at least necessary. Make any one of these factors go away and the sequence of events would probably have led elsewhere. But this also means that both sets of claims are wrong. They are both wrong in that they focus on only a subset of contributory factors and pick and choose which ones are causal and which ones are not. This choosing can be driven more by socio-political and organizational pressures than by mere evidence found in the rubble. Cause is not something you find. Cause is something you construct. How you construct it and from what evidence appears to depend on where you look, what you look for, who you talk to, and likely on who you work for.

There is no "root" or "primary" cause

How come that there are so many causes to choose from in any mishap? This has to do with the fact that the kinds of systems that are vulnerable to human error are

[1] See: Aeronautica Civil (1996). *Aircraft Accident Report: Controlled flight into terrain American Airlines flight 965, Boeing 757-223, N851AA near Cali, Colombia, December 20, 1995*. Santafe de Bogota, Colombia: Aeronautica Civil Unidad Administrativa Especial, and American Airlines' (1996) Submission to the Cali Accident Investigation.

so well protected against it. The potential for danger in many industries and systems has been recognized long ago. And consequently, major investments have been made in protecting them from the breakdowns that we know or think can occur. These so-called "defenses" against failure contain human and engineered and organizational elements.

Flying the right approach speeds for landing while an aircraft goes through its subsequent configurations (of flaps and slats and wheels that come out), is safety-critical. As a result it has evolved into a well-defended process of double-checking and cross-referencing between crew members, speed booklets, aircraft weight, instrument settings, reminders and call-outs, and in some aircraft even by engineered interlocks.

Accidents in such systems can occur only if multiple factors succeed in eroding or bypassing all these layers of defense. The breach of any of these layers can be called "causal". For example, the crew opened the speed booklet on the wrong page (i.e. the wrong aircraft landing weight). But this fails to explain the entire breakdown, because other layers of defense had to be broken or side-stepped too. And there is another question. Why did the crew open the booklet onto the wrong page? In other words, what is the cause of that action? Was it their expectation of aircraft weight based on fuel used on that typical trip; was it a misreading of an instrument? And once pinpointed, what is the cause of that cause? And so forth.

Because of this investment in multiple layers of defense, we can find "causes" of failures everywhere—when they happen, that is. The causal web quickly multiplies and fans out, like cracks in a window. What you call "root cause" is simply the place where you stop looking any further. As far as the causal web is concerned, there are no such things as root or primary causes—there is in fact no end anywhere. If you find a root or primary cause, it was your decision to distinguish something in the dense causal pattern by those labels.

There is no single cause

So what is the cause of the accident? This question is just as bizarre as asking what *the* cause is of not having an accident. There is no single cause. Neither for failure, nor for success. In order to push a well-defended system over the edge (or make it work safely, for that matter), a large number of contributory factors are necessary and only jointly sufficient.

MULTIPLE FACTORS—EACH NECESSARY AND ONLY JOINTLY SUFFICIENT—ARE NEEDED TO PUSH A COMPLEX SYSTEM OVER THE EDGE OF BREAKDOWN

So where you focus in your search for cause is something that the evidence in a mishap will not necessarily determine for you. It is up to your investigation.

In a break with the tradition of identifying "probable causes" in aviation crashes—which oversimplify the long and intertwined pathway to failure—Judge Moshansky's investigation of the Air Ontario crash at Dryden, Canada in 1989 did not produce any probable causes.

The pilot in question had made a decision to take off with ice and snow on the wings, but, as Moshanky's commission wrote, "that decision was not made in isolation. It was made in the context of an integrated air transportation system that, if it had been functioning properly, should have prevented the decision to take off...there were significant failures, most of them beyond the captain's control, that had an operational impact on the events at Dryden...regulatory, organizational, physical and crew components...."

Instead of forcing this complexity into a number of probable causes, the Commission generated no less than 191 recommendations which pointed to the many "causes" or systemic failures underlying the symptomatic accident on that day in March 1989. Recommendations ranged in topic from the introduction of a new aircraft type to a fleet, to management selection and turn-over in the airline, to corporate take-overs and mergers in the aviation industry.[1]

Probable cause statements are of necessity:

- Selective. There are only so many things you can label "causal" before the word "causal" becomes meaningless.;
- Exclusive. They leave out factors that were also necessary and only jointly sufficient to "cause" the failure;
- Oversimplifications. They highlight only a few hotspots along a long, twisted and highly interconnected causal pathway that starts long before and far way from where the actual failure occurs.

If protocol prescribes that probable causes be identified, the best way to deal with that is to generate, as "probable cause", the shortest possible summary of the

---

[1] Moshansky, V. P. (1992). *Commission of inquiry into the Air Ontario accident at Dryden, Ontario* (Final report, vol. 1-4). Ottawa, ON: Minister of Supply and Services, Canada.

sequence of events that led up to the mishap. This description should start as high up in the causal chain as possible, and follow the meandering pathway to the eventual failure.

4. Human Error By Any Other Name

"A spokesman for the Kennedy family has declined to comment on reports that a federal investigation has concluded that pilot error caused the plane crash that killed John F. Kennedy Jr., his wife and his sister-in-law. The National Transportation Safety Board is expected to finish its report on last year's crash and release it in the next several weeks. Rather than use the words 'pilot error', however, the safety board will probably attribute the cause to Kennedy's becoming 'spatially disoriented', which is when a pilot loses track of the plane's position in the sky."[1]

UNDERSPECIFIED LABELS

"Human error" as explanation for accidents has become increasingly unsatisfying. As mentioned earlier, there is always an organizational world that lays the groundwork for the "errors", and an operational one that allows them to spin into larger trouble.

We also know there is a psychological world behind the errors—to do with people's attention, perception, decision making, and so forth. Since a number of decades, human factors has produced or loaned a number of terms that try to capture such phenomena. Labels like "complacency", "situation awareness", "crew resource management", "shared mental models", "stress", "workload", are such common currency today that nobody really dares to ask what they actually mean. The labels are assumed to speak for themselves; to be inherently meaningful. They get used freely as causes to explain failure. For example:

- "The crew lost situation awareness and effective crew resource

[1] International Herald Tribune, 24-25 June 2000.

management (CRM)" (which is why they crashed);
- "High workload led to a stressful situation" (which is why they got into this incident);
- "It is essential in the battle against complacency that crews retain their situation awareness" (otherwise they keep missing those red signals).

The question is: are labels such as complacency or situation awareness much better than the label "human error"? In one sense they are. They provide some specification; they appear to give some kind of reasons behind the behavior; they provide an idea of the sort of circumstances and manner in which the error manifested itself.

But if they are used as quoted above, they do not differ much from the verdict "human error" they were meant to replace. These labels actually all conclude that human error—by different names—was the cause:

- Loss of CRM is one name for human error—the failure to invest in common ground, to coordinate operationally significant data among crewmembers;
- Loss of situation awareness is another name for human error—the failure to notice things that in hindsight turned out to be critical;
- Complacency is also a name for human error—the failure to recognize the gravity of a situation or to follow procedures or standards of good practice.

DEALING WITH THE ILLUSION OF EXPLANATION

Human factors risks falling into the trap of citing "human error" by any other name. Just like "human error", other labels also hide what really went on—they try to say so much that they may end up saying very little. Indeed, these labels must not be mistaken for deeper insight into human factors issues. The risk occurs when these labels are applied by investigators without making explicit connections between:

- The label and the evidence for it. For example, exactly which interactions and miscoordinations in the sequence of events constituted a loss of effective crew resource management—based on available and accepted models of "effective crew resource management"?
- The label and how it "caused" the mishap. For example, "loss of effective crew resource management" may be cited in the probable

causes or conclusions. But how exactly did the behaviors that constituted its loss contribute to the outcome of the sequence of events?

If you reveal which kinds of behaviors in the sequence of events produced a "loss of effective crew resource management", these behaviors can themselves point to the outcome, without you having to rely on a label that obscures all the interesting bits and interactions.

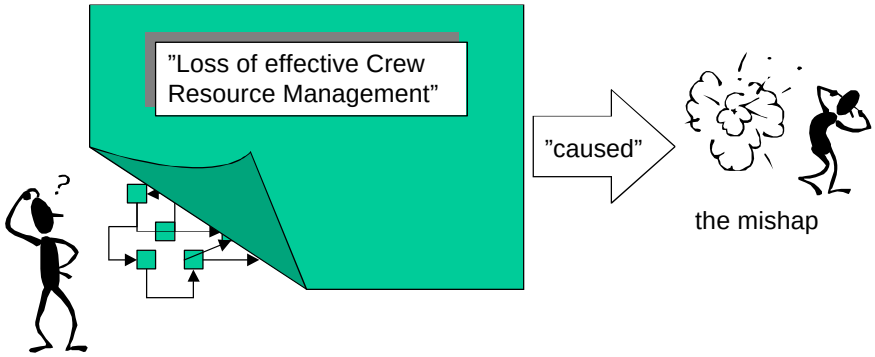


Fig. 4.1: The interesting mental dynamics take place *beneath* the large psychological label. The label in itself explains nothing.

To understand the mindset of someone caught up in an unfolding situation is not a matter of translating his or her behavior into big psychological terms. It's the mental dynamics *beneath* the labels that are interesting—for example:

- The ways people shift attention on the basis of earlier assessments of the situation or on the basis of future expectations;
- The trade-offs they have to make between various operational or organizational goals;
- How they activate and apply knowledge in context;
- How they recognize patterns of data on the basis of experience with similar circumstances;
- How they coordinate with various sources of expertise inside and outside their situation;
- How they deal with the clumsiness and complexity of the technology that surrounds them.

It's these mental and interpersonal processes that drive a sequence of

events in certain directions; it's these processes—if anything—that can be said to be "causal" in the sense that they help determine the outcome of a sequence of events. When you penetrate the evidence of your mishap to a level where you can start to see these processes at work, you will become able to connect them directly to the outcome that followed, bypassing or at least specifying the large label that would otherwise obscure all the interesting cognitive dynamics and causal links.

The use of large terms in investigative findings and explanations may be seen as the rite of passage into psychological phenomena. That is, for a human factors investigation to be taken seriously, it should contain its dose of situation awarenesses and stresses and workloads. But this is a misconception, and the rest of this chapter presents an alternative. More detailed analysis should be done behind the labels. This produces more specific, more meaningful insights into human behavior. It will render your investigation more accessible and verifiable for others too.

LOSS OF SITUATION AWARENESS

A label that has become very popular—loss of situation awareness—is also increasingly recognized to be problematic. The traditional idea is that we process information from the world around us and form a picture of what is going on on the basis it (see figure 4.2).

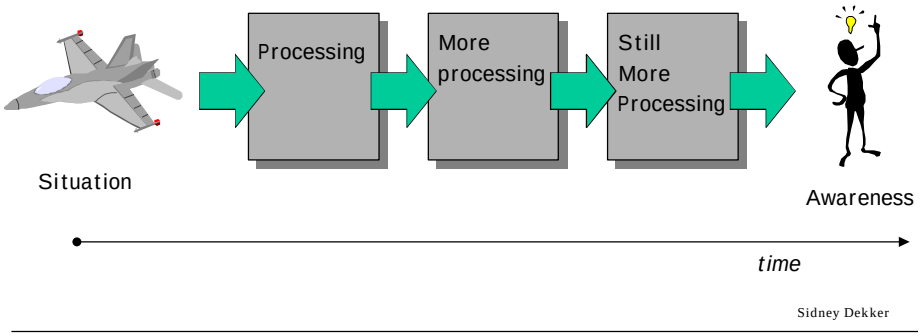


Fig. 4.2: The traditional notion of situation awareness: we process information from the world until we arrive at awareness, or a mental picture of what is going on.

Such information processing is typically thought to go through several stages (for example perceiving elements in the situation, processing

their meaning and understanding their future implications) before arriving at full situation awareness. A "loss of situation awareness" may occur when our information processing is hampered in some way, for example by high stress or workload. (see figure 4.3)
There are major problems with this notion. First, it portrays people as passive recipients of whatever the world throws at them, and everything is OK as long as our mental processing can keep up. In this model we make no active contribution to our understanding of the world, and no active contribution to changing the world itself—which we certainly do in reality. For example, we move around in the world; we change and tweak things to make it reveal more about itself; we influence it to make it slow down; we decide to look in some places rather than others.

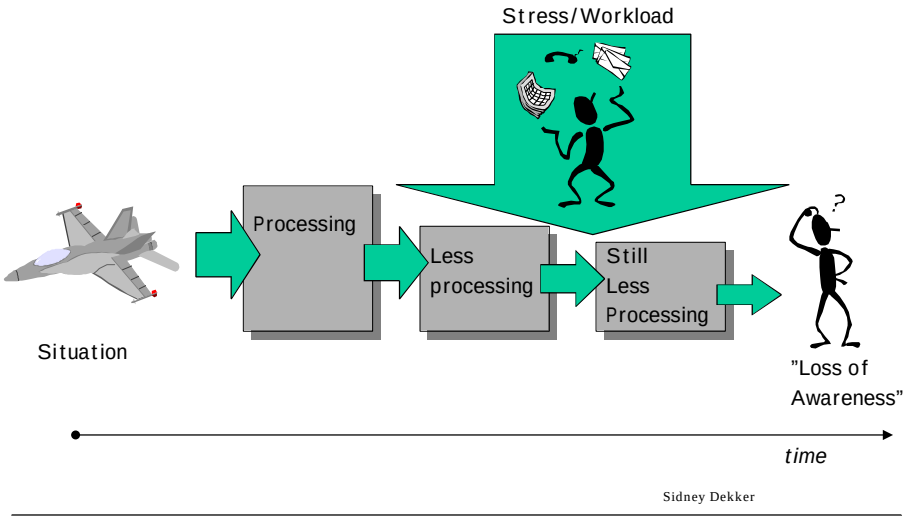


Fig. 4.3: In the traditional notion, a loss of situation awareness is presumed to occur through pressures and difficulties in processing information.

Another issue is that we do not perceive elements in a situation and only then set out to make sense of them by gradually adding meaning along some psychic information highway. If we would perceive individual "elements", we would get pummeled by the world. We would simply go crazy. In reality we perceive patterns, structures. We give meaning to the world simply by looking at it. We rely on our experience to recognize key patterns that indicate what is happening; where things are going.

A third problem is that we cannot "lose awareness" (other than by

becoming physically unconscious). There is no such thing as a mental vacuum. We are always forming *some* idea of where they are; of what our system and process is doing. We cannot help but give meaning to incoming cues. We interpret incoming data on the basis of what we already know; what we have just done to the system or process; what we have set out to do; what we expect to happen.

IF YOU LOSE SITUATION AWARENESS, WHAT REPLACES IT?

THERE IS NO SUCH THING AS A MENTAL VACUUM

Indeed, the question "what is happening now?" has such an idea behind it: people had expectations of what the system or process would do. By implication, people had some kind of mental model on which to form those expectations. The system or process did not behave according to their expectations, thus prompting the question.

The reality is that we could not live or survive without constantly building and maintaining and modifying our idea of the changing world around us, influencing our situation on the basis of it, and then receiving new information which updates our understanding once again.

Figure 4.4 portrays this cycle, in which situation awareness is not some end-product, but a constant process of assessments and actions that inform one another. Through this loop of continuous transactions with the world, we are remarkably good at creating a coherent and robust picture of our systems and processes, even when evidence is buggy, incomplete, shifting and uncertain. By going around and around through the cognitive cycle, we get and stay in tune with our circumstances, which enables us to function in a complex and constantly changing world.

Reverse engineering of situation awareness

When, in hindsight, you uncover a mismatch between how people understood their situation to be, and how you now know it really was (see figure 2.2) nothing was lost. The challenge for you as investigator is not to point out how people at another time and place did not know what you know today (calling it *their* "loss of situation awareness"). The challenge is to reconstruct how *they* understood the unfolding

situation—what they were looking at and how they gave meaning to incoming data and what they were expecting. People's understanding of the situation can become clear if you look at their actions. What were they driving at? And what made them focus on certain cues rather than others? What evidence did *they* find to cling onto a hypothesis that you now know was increasingly at odds with the real situation? Yet what in their situation would have made this focus reasonable?

DO NOT POINT OUT HOW OTHER PEOPLE DID NOT KNOW WHAT YOU KNOW TODAY

—CALLING IT THEIR LOSS OF SITUATION AWARENESS—

RECONSTRUCT HOW THEY UNDERSTOOD THEIR UNFOLDING SITUATION TO BE.

Sidney Dekker

Fig. 4.4: The new view of situation awareness: we make assessments about the world, updating our current understanding. This directs our actions in the world, which

change what the world looks like, which in turn updates our understanding, and so forth (Figure is modeled on Ulrich Neisser's perceptual cycle).

Chapter 7 is about the reconstruction of unfolding mindset. It starts with:

- The situation in which people found themselves—the various unfolding threads of data; the multiple demands and pressures;
- The tasks people had to carry out in this situation;
- The tools with which they were doing this work.

Chapter 7, together with the chapters that follow, helps you to reverse engineer people's constantly updated idea of how things were evolving around them.

LOSS OF EFFECTIVE CREW RESOURCE MANAGEMENT

In most complex worlds, people do not carry out their work alone. Work, and the error detection and recovery in it, is inherently distributed over multiple people, likely in different roles.

- These people need to coordinate to get the work done
- Thus, problems in coordination may mark a sequence of events towards failure.

Crew Resource Management has become a popular label—not only in aviation, but also in medicine and other domains—that covers the coordinative processes between teammembers who pursue a common operational goal. So what does "the loss of effective CRM" mean? Here are some places to look for more specifics:

Differences between teammembers' goals.

Complex operating environments invariably contain multiple goals that can all be active at the same time.

Take a simple flight from A to B: On-time arrivals, smooth rides through weather, slot allocation pressures, optimum fuel usage, availability of alternates, passenger convenience—these are all goals that can influence a

single assessment or decision.

Given that people in the same operational team have different roles, not everyone in a team may feel equally affected by, or responsible for, some of these goals. This can lead to mismatches between what individuals see as their, or the team's, dominant pursuit at any one time.

Differences between teammembers' interpretation

Divergences can exist and grow in how people with different backgrounds and roles can interpret their circumstances. Different assessments can lead to different goals being pursued, and different actions being taken.

Gary Klein tells an interesting story of an airliner with three generators—one on each of its engines. One of the generators failed early in a flight. This is not particularly unsafe: two generators can provide the electrical power the aircraft needs. But then another engine began to lose oil, almost forcing a shut-down. After some discussion, the crew decided to let the ailing engine run idle, so that its generator could be called upon if necessary. When asked after landing how many generators had just been available, the co-pilot (who was flying the aircraft at the time) said "two". The captain said "one and a half", meaning one good engine and one idle. But the flight engineer said "one"--since getting the idle engine up and running where it powers the generator takes a moment.[1]

Knowledge that did not make it into the crew consciousness.

The story above also shows how certain knowledge can remain in a team's pre-conscious—that is, locked in the heads of individuals without being made public, or conscious. There may be many reasons why individuals do not contribute their understanding of the situation or their knowledge to the common ground, including overbearing commanders or shy subordinates. But very often the lack of

---

[1] Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

coordination is a matter of people assuming that others have a similar understanding of the situation. Just like the flight engineer in the example above may have assumed that the two pilots knew how only one generator was available for at least a moment. Usually there are very good reasons for these assumptions, as they facilitate team coordination by not cluttering crew communication with redundant reminders and pointers.

When you encounter differences between people's goals, between people's interpretations and when you find missing communications in the rubble, it is easy to look at them as failures or losses. Failures of teamwork, for example. Or failures of leadership, or loss of crew resource management. But look behind the failure. Silence by one crewmember may in actually represent good teamwork—which includes knowing when not to interrupt.

Features of the operating environment

Features of the operating environment may make the sharing of assessments and actions difficult (ergnomic problems such as high noise levels or low lighting or clumsy seating arrangements already do this). Other, more subtle features of people's operating environment can also profoundly influence how well they can coordinate, and how well they can cross-check and catch errors made by others:

Modern airliners are equipped with flight management systems (FMS's) that basically fly the entire aircraft today. Pilots of these airliners each have individual access to the FMS through a separate interface—their private little workspace. Here they can make significant changes to the flight plan without the other pilot necessarily seeing, knowing, or understanding. The pilot only needs to press "execute" and the computer will do what s/he has programmed.

Airlines have of course devised procedures that require pilots to cross-check each other's computer entries, but in reality there are many circumstances in which this is impractical or unnecessary. The real coordination problem is not pilots' failure to follow procedures. It is a feature of the design that makes coordination very difficult, yet safety-critical.[1]

---

[1] Dekker, S. W. A., & Hollnagel, E. (Eds.) (1999). *Coping with computers in the cockpit*. Aldershot, UK: Ashgate.

COMPLACENCY

Confronted with failure, it can be easy to see people's behavior as deficient, as unmotivated, as not living up to what you may expect from operators in their position. One of the labels often given here is "complacency" or "negligence". Over time, people seem to have lost respect for the seriousness of their jobs—they start reading newspapers while driving their trains or flying their aircraft, they do not double-check before beginning an amputation.

Departures from the routine that become routine

Figure 4.5 shows what really may be going on here and why complacency or negligence is not only a judgment, but also an incomplete label.

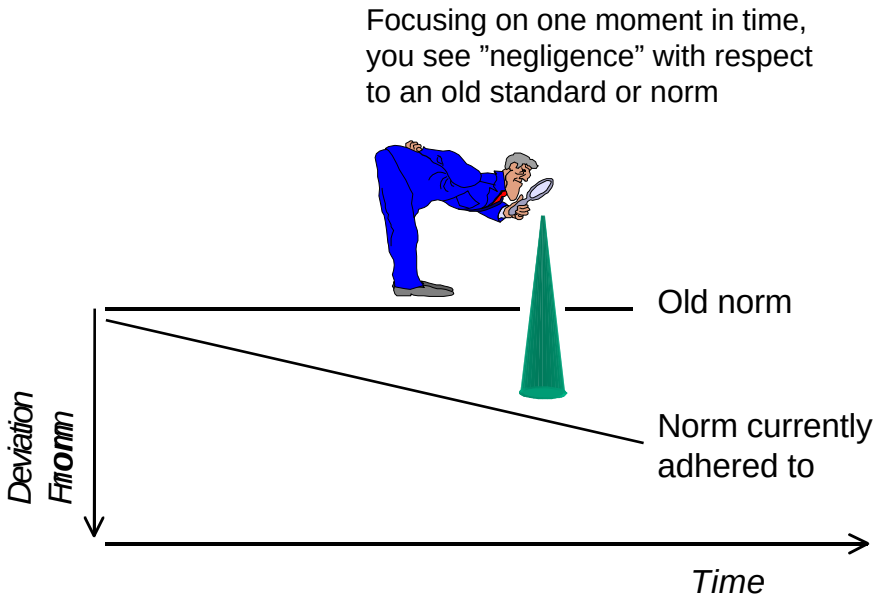


Fig. 4.5: At a particular moment in time, behavior that does not live up to some standard may look like complacency or negligence. But this focus ignores the history, and thus the explanation, behind the behavior in question. Deviance may have become the new norm across an entire operation or organization.

Departures from some standard or routine may at any one moment seem to occur because people are not motivated to do otherwise—, these people are seen as "complacent" or "negligent". But often, theirs is not the first departure from the routine. Departures from the routine that have become routine can include anything from superficial checklist reading, to cutting other corners to save time, to signing off equipment or people without all official criteria met.

Most pertinent to human error investigations is to find out what organizational history or pressures exist behind these routine departures from the routine. Take on-time departures, arrivals or deliveries—relevant to any organization that operates on a schedule:

- The rewards of on-time performance are immediate and tangible: happy customers, happy bosses, money made, and so forth.
- The potential risks (how much did you borrow from safety to operate on time?) are unclear, unquantifiable or even unknown.

Borrowing from safety

With rewards constant and tangible, departures from the routine may become routine across an entire operation or organization.

DEVIATIONS FROM THE NORM CAN THEMSELVES BECOME THE NORM

Without realizing it, people start to borrow from safety, and achieve other system goals because of it—production, economics, customer service, political satisfaction. Behavior shifts over time because other parts of the system send messages, in subtle ways or not, about the importance of these goals. In fact, organizations reward or punish operational people in daily trade-offs ("We are an ON-TIME operation!"), focusing them on goals other than safety. The lack of adverse consequences with each trade-off that bends to goals other than safety, strengthens people's tacit belief that it is safe to borrow from safety.

In "The Challenger Launch Decision", Diane Vaughan has carefully documented how an entire organization started borrowing from safety—reinforced by one successful Space Shuttle Launch after the other, even if O-rings in the

solid rocket boosters showed signs of heat damage. The evidence for this O-ring "blow-by" was each time looked at critically, assessed against known criteria, and then decided upon as "acceptable". Vaughan has called this repeated process "the normalization of deviance": what was deviant earlier, now became the new norm. This was thought to be safe: after all, there were two O-rings: the system was redundant. And if past launches were anything to go by (the most tangible evidence for success), future safety would be guaranteed. The Challenger Space Shuttle, launched in cold temperatures in January 1986, showed just how much NASA had been borrowing from safety: it broke up and exploded after lift-off because of O-ring blow-by.[1]

The problem with complex, dynamic worlds is that safety is not a constant. Past success while departing from a routine is not a guarantee for future safety. In other words, a safe outcome today is not a guarantee of a safe outcome tomorrow, even if behavior is the same. This means that[2]:

MURPHY'S LAW IS WRONG

WHAT CAN GO WRONG USUALLY GOES RIGHT, BUT THEN WE DRAW THE WRONG CONCLUSION

Circumstances change, and so do the safety treats associated with them. Doing what you do today (which could go wrong but did not) does not mean you will get away with it tomorrow. The dynamic safety threat is picture in figure 4.6.

---

1 Vaughan, D. (1996). *The Challenger launch decision*. Chicago, IL: University of Chicago Press.

2 The quote on Murphy's law comes in part from Langewiesche, W. (1998). *Inside the sky*. New York: Pantheon.

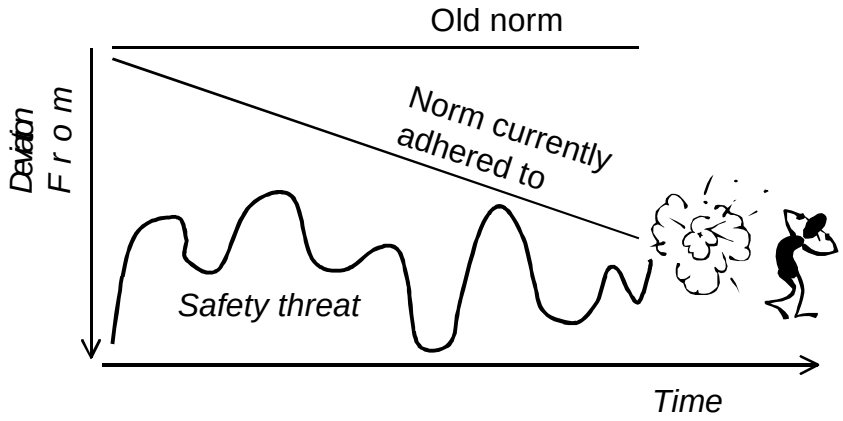


Fig. 4.6: Murphy's law is wrong. What can go wrong usually goes right, and over time we come to think that a safety threat does not exist or is not as bad. Yet while we adjust our behavior to accommodate other system pressures (e.g. on-time performance), safety threats vary underneath, setting us up for problems sometime down the line.

STRESS AND WORKLOAD

Stress has long been an important term, especially where people carry out dynamic, complex and safety-critical work. On a superficial reading of your mishap data, it may be easy to assert that people got stressed; that there was high workload and that things got out of hand because of it. But this does not mean or explain very much. Psychologists still debate whether stress is a feature of a situation, the mental result of a situation, or a physiological and psychological coping strategy that allows us to deal with demanding or threatening circumstances. This complicates the use of stress in any causal statement, because what produced what?

Demand-resource mismatch

What you can do on the basis of your data is make an inventory of the demands in a situation, and the resources that people had available to cope with these demands. This is a way to more specifically handle your available evidence. If you suspect that stress or high workload

may have been an issue, look for examples of demands and resources in your situation (see Table 4.1).

In studies of stress and workload, some have reported that a mismatch between demands and resources may mean different things for different kinds of operators. In a marine patrol aircraft, for example, people in the back are concerned with dropping sonobuoys (to detect submarines) out of the aircraft. The more sonobuoys in a certain amount of time, the more workload, the more stress. People in the front of the aircraft were instead concerned with more strategic questions. For them, the number of things to do had little bearing on their experience of stress and workload. They would feel stressed, however, if their model of the situation did not match reality, or if it had fallen behind actual circumstances.

Tunneling and regression

One of the reported consequences of stress is tunneling—the tendency to see an increasingly narrow portion of one's operating environment. This is generally interpreted as a shortcoming; as something dysfunctional that marks less capable operators and should be avoided if at all possible. Another consequence that has been noted is regression—the tendency to revert to earlier learned routines even if not entirely appropriate to the current situation.

Problem demands:	Coping resources:
Ill-structured problems	Experience with similar problems
Highly dynamic circumstances: things changing quickly over time	Other people contributing to assessments of what is going on
Uncertainty about what is going on or about possible outcomes	Knowledge or training to deal with the circumstances
Interactions with other people that generate more investment than return (in terms of offloading)	Other people to off-load tasks or help solve problems

Organizational constraints and pressures	Organizational awareness of such pressures and constraints
Conflicts between goals	Guidance about goal priorities
High stakes associated with outcome	Knowledge there is an envelope of pathways to a safe outcome
Time pressure	Workload management skills

Table 4.1: Finding a mismatch between problem demands and coping resources can help you make arguments about stress and workload more specific.

You can actually see both tunneling and regression as strategies in themselves; as a contributions from the human that are meant to deal with high demands (lots to pay attention to and keep track of) and limited resources (limited time too look around; limited mental workspace to integrate and deal with diverse and rapidly changing data). Tunneling (sometimes called "fixation", especially when people lock onto one explanation of the world around them) comes from the human strength to form a stable, robust idea of a shifting world with multiple threads that compete for attention and where evidence may be uncertain and incomplete. It gives us the stability of a framework to interpret and assess new data, and allows us to stay ahead of changes in the world by forming predictions about what will happen next.
If we were to jump on each new piece of data instead, and change tack and explanation right there, our ability to function in a changing world quickly disintegrates. This "mental vagabonding" sometimes happens, of course, and certain domains even have their own term for it, for example "falling behind the airplane". Without sufficient experience in handling a particular scenario, or with other complicating factors present, we can get to lag behind in the cognitive cycle. With every change in the world, attempts are made (but truncated by new changes) to update our understanding or catch up with responding actions. This makes it difficult to anticipate and influence future circumstances: behavior becomes driven by events more than the other way around. When confronted with evidence in this direction, ask yourself: what were the multiple pressures and attentional demands that made people fall behind developments in the world around them, forcing them to deal with newly emerging problems in an event-driven, haphazard, uncoordinated way?

In highly dynamic and complex situations, it would seem that tunneling is an (involuntary) strategy which allows people to track and

stay ahead of a limited number of threads out of a host of potential distracters. Similarly, regression to earlier learned routines frees up mental resources: we do not have to match current perceptions with consciously finding out what to do each time anew. Stress and workload, and people's own perception of it, will thus be affected by their ways of dealing with it.

5. Human Error— In The Head Or In The World?

The use of underspecified labels in human error investigations, covered in the previous chapter, has various roots. One reason for the use of large psychological terms is the confusion over whether you should start looking for the source of human error:

- In the head (of the person committing the error)
- Or in the situation (in which the person works)

The first alternative is used in various human error analysis tools, and in fact often implied in investigations. For example, when you use "complacency" as a label to explain behavior, you really look for how the problem started with an individual who was not sufficiently motivated to look closely at critical details of his or her situation.

As said in the previous chapters, such an approach to "explaining" human error is a dead-end. It prevents an investigation from finding enduring features of the operational environment that actually produce the controversial behavior (and that will keep producing it if left in place). And there is more. The assumption that errors start in the head also leaves an investigative conclusion hard to verify for others, as is explained below.

The alternative—look for the source of error in the world—is a more hopeful path for investigations. Human error is systematically linked to features of the world—the tasks and tools that people work with, and the operational and organizational environment in which people carry out that work. If you start with the situation, you can identify, probe and document the reasons for the observed behavior, without any need to resort to non-observable processes or structures or big labels in someone's head. This is the path that The Field Guide will take you along.

HUMAN ERROR—IT'S ALL IN THE HEAD

To "reverse engineer" human error, chapter 9 will encourage you to reconstruct how people's mindset unfolded and changed over time. You would think that reconstructing someone's unfolding mindset begins with the mind. The mind, after all, is the obvious place to look for the mindset that developed inside of it. Was there a problem holding things working memory? What was in the person's perceptual store? Was there trouble retrieving a piece of knowledge from long-term memory? These are indeed the kinds of questions asked in a variety of human error analysis tools and incident reporting systems.

A tool is being developed for the analysis of human errors in air traffic control. For each observed error, it takes the analyst through a long series of questions that are based on an elaborate information processing model of the human brain. It begins with perceptual processes and points the analyst to possible problems or difficulties there. Then it goes on along the processing pathway, hoping to guide the analyst to the source of trouble in a long range of psychological processes or structures: short term memory, long term memory, decision making, response selection, response execution, and even the controller's image of him or herself. For each observed error, the journey through the questions can be long and arduous and the final destination (the supposed source of error) dubious and hard to verify.

These kinds human error analyses deal with the complexity of behavior by simplifying it down to boxes; by nailing the error down in a psychological process or structure. For example, it was an error of vigilance, or one of working memory, or one of judgment or decision making, or one of response selection. The aim is to conclude that an error originated in a certain stage along a psychological processing pathway in our head. These approaches basically explain an error by taking it back to the brain from which it came.

The shortcomings, as far as investigating human error is concerned, are severe. These approaches hide an error back in the brain under a label that is not much more revealing or enlightening than "human error" is. The labels made popular in these approaches (such as working memory or response execution) are also little more than artifacts of the language of a particular psychological model. This model may not even be right, but it sure is hard to prove wrong. Who can prove the existence of short term memory? But who can prove that it does not exist?

This problem extends seriously into investigative practice. Explaining human error on the basis of internal mental structures will leave other people guessing as to whether you were right or not. Nobody can actually see things like short term memories or perceptual stores, and nobody can go back into the short term memories or perceptual stores of the people you are investigating to check your work. Nobody can really verify your conclusion that these things were responsible for the failures that occurred. Other people can only hope you were right when you categorized. By just relabeling it in more detailed psychological terms, human

error, and its investigation, remains locked in a practice where anyone can make seemingly justifiable yet unverifiable assertions. Investigations remain fuzzy and uncertain and inconclusive, and low on credibility.

HUMAN ERROR—MATTER OVER MIND

Things are different when you begin your investigation with the unfolding situation in which people found themselves. Methods that attribute human error to structures inside the brain easily ignore the situation in which human behavior took place, or they at least underestimate its importance. Yet it makes sense to start with the situation:

- Past situations can be objectively reconstructed to a great extent, and documented in detail;
- There are tight and systematic connections between situations and behavior; between what people did and what happened in the world around them.

These connections between situations and behavior work both ways:

- People change the situation by doing what they do; by managing their processes;
- But the evolving situation also changes people's behavior. An evolving situation provides changing and new evidence; it updates people's understanding; it presents more difficulties; it forecloses or opens pathways to recovery.

You can uncover the connections between situation and behavior, investigate them, document them, describe them, represent them graphically. Other people can look at the reconstructed situation and how you related it to the behavior that took place inside of it. Other people can actually trace your explanations and conclusions. Starting with the situation brings a human error investigation out in the open. It does not rely on hidden psychological structures or processes, but instead allows verification and debate by those who understand the domain. When a human error investigation starts with the situation, it sponsors its own credibility.

A large part of human error investigations, then, is not at all about the human behind the error. It is not about supposed structures in a human's brain; about psychological constructs that were putatively involved in causing mental hick-ups. A large part of human error investigations is about the situation in which the human was working; about the tasks he or she was carrying out; about the tools that were used.

THE RECONSTRUCTION OF MINDSET BEGINS NOT WITH THE MIND

IT BEGINS WITH THE CIRCUMSTANCES IN WHICH THE MIND FOUND ITSELF

To understand the situation that produced and accompanied behavior, is to understand the human assessments and actions inside that situation. This allows you to "reverse engineer" human error by showing:

- how the safety critical process changed over time;
- how people's assessments and actions evolved in parallel with their changing situation;
- how features of people's tools and tasks and their organizational and operational environment influenced their assessments and actions inside that situation.

This is what the reconstruction of unfolding mindset, the topic of chapter 8, is all about.

6. Put Data Into Context

Putting behavior back into the situation that produced and accompanied it is not easy. In fact, to make sense of behavior it is always tempting to go for a context that actually lies *outside* the accident sequence. Taking behavior out of context, and giving it meaning from the outside, is common in investigations. This chapter discusses two ways in which behavioral data is typically taken out of context, by:

- micro-matching them with a world you now know to be true, or by
- lumping selected bits together under one condition you have identified in hindsight ("cherry picking").

OUT OF CONTEXT I: HOLDING PERFORMANCE FRAGMENTS AGAINST A WORLD YOU NOW KNOW TO BE TRUE

One of the most popular ways by which investigators assess behavior is to hold it up against a world he or she *now* knows to be true. There are various ways in which after-the-fact-worlds can be brought to life:

- A procedure or collection of rules: People's behavior was not in accordance with standard operating procedures that were found to be applicable for that situation afterward;
- A set of cues: People missed cues or data that turned out to be critical for understanding the true nature of the situation;
- Standards of good practice: People's behavior fall short of standards of good practice in the particular industry.

The problem is that these after-the-fact-worlds have very little in common with the actual world that produced the behavior under investigation. They contrast people's behavior against the investigator's reality, not the reality that surrounded the behavior in question. Thus, micromatching fragments of behavior with these various standards explains nothing—it only judges.

Procedures

First, individual fragments of behavior are frequently compared with procedures or regulations, which can be found to have been applicable in hindsight. Compared with such written guidance, actual performance is often found wanting; it does not live up to procedures or regulations.

Take the automated airliner that started to turn towards mountains because of a computer-database anomaly. The aircraft ended up crashing into the mountains. The accident report explains that one of the pilots executed a computer entry without having verified that it was the correct selection, and without having first obtained approval of the other pilot, contrary to the airline's procedures.[1] Other commentators add how, in their assessments and actions, the flightcrew failed to adhere to Federal Aviation Regulation FAR 91.123(a).

Investigations invest considerably in organizational archeology so that they can construct the regulatory or procedural framework within which the operations took place or should have taken place. Inconsistencies between existing procedures or regulations and actual behavior are easy to expose in hindsight. Your starting point is a fragment of behavior, and you have the luxury of time and resources to excavate organizational records and regulations to find rules with which the fragment did not match.

But what have you shown? You have only pointed out that there was a mismatch between a fragment of human performance and existing guidance that you uncovered or highlighted after-the-fact. This is not very informative. Showing that there was a mismatch between procedure and practice sheds little light on the *why* of the behavior in question. And, for that matter, it sheds little light on the why of this particular mishap. Mismatches between procedure and practice are not unique ingredients of accident sequences. They are often a feature of daily operational life (which is where the interesting bit in your investigation starts).

---

1 The accident report is: Aeronautica Civil (1996). *Aircraft Accident Report: Controlled flight into terrain American Airlines flight 965, Boeing 757-223, N851AA near Cali, Colombia, December 20, 1995*. Santafe de Bogota, Colombia: Aeronautica Civil Unidad Administrativa Especial.

Available data

Second, to construct the world against which to evaluate individual performance fragments, investigators can turn to data in the situation that were not picked up by the operators but that, in hindsight, turned out to be critical.

Continue with the automated aircraft above. What should the crew have seen in order to notice the turn? They had plenty of indications, according to the manufacturer of their aircraft:

"Indications that the airplane was in a left turn would have included the following: the EHSI (Electronic Horizontal Situation Indicator) Map Display (if selected) with a curved path leading away from the intended direction of flight; the EHSI VOR display, with the CDI (Course Deviation Indicator) displaced to the right, indicating the airplane was left of the direct Cali VOR course, the EaDI indicating approximately 16 degrees of bank, and all heading indicators moving to the right. Additionally the crew may have tuned Rozo in the ADF and may have had bearing pointer information to Rozo NDB on the RMDI".[1]

This is a standard response after mishaps: point to the data that would have revealed the true nature of the situation. But knowledge of the "critical" data comes only with the privilege of hindsight. If such critical data can be shown to have been physically available, it is automatically assumed that it should have been picked up by the operators in the situation.

The problem is that pointing out that it should have does not explain why it was perhaps not, or why it was interpreted differently back then. There is a difference between data availability and data observability—between what can be shown to have been physically available and what would have been observable given the multiple interleaving tasks, goals, attentional focus, interests, and even culture of the person in question.

The mystery, as far as an investigation is concerned, is not why people could have been so unmotivated or stupid not to pick up the things that you can decide were critical in hindsight. The mystery is to find out what *was* important to them, and why.

---

[1] Boeing submission to the American Airlines Flight 965 Accident Investigation Board (1996). Seattle, WA: Boeing.

Other standards

Third, there are a number of other standards especially for performance fragments that do not easily match procedural guidance or for which it is more difficult to point out data that existed in the world and should have picked up.
This is often the case when a controversial fragment knows no clear pre-ordained guidance but relies on local, situated judgment. For example, a decision to accept a runway change, or continue flying into bad weather. For these cases there are always ”standards of good practice” which are based on convention and putatively practiced across an entire industry. One such standard in aviation is ”good airmanship”, which, if nothing else can, will cover the variance in behavior that had not yet been accounted for.

Cases for medical negligence can often be made only by contrasting actual physician performance against standards of proper care or good practice. Rigid, algorithmic procedures generally cannot live up to the complexity of the work and the ambiguous, ill-defined situations in which it needs to be carried out. Consequently, it cannot easily be claimed that this or that checklist should have been followed in this or that situation.

But which standards of proper care do you invoke to contrast actual behavior against? This is largely arbitrary, and driven by hindsight. After wrong-site surgery, for example, the standard of good care that gets invoked is that physicians have to make sure that the correct limb is amputated or operated upon.

As a physician, you are chanceless against such a judgment. You can only nod your head in approval at such motherhood exhortations, and think that—after all—these are the standards you try to follow all the time; in all the little and larger decisions and trade-offs you make daily. Finding appropriate standards in hindsight does nothing to elucidate the actual circumstances and systemic vulnerabilities which in the end allowed wrong-site surgery to take place.

By referring to procedures, physically available data or standards of good practice, investigators can micro-match controversial fragments of behavior with standards that seem applicable from their after-the-fact position. Referent worlds are constructed from outside the accident sequence, based on data investigators now have access to, based on facts they now know to be true. The problem is that these after-the-fact-

worlds may have very little relevance to the circumstances of the accident sequence. They do not explain the observed behavior. The investigator has substituted his own world for the one that surrounded the people in question.

OUT OF CONTEXT II: GROUPING SIMILAR PERFORMANCE FRAGMENTS UNDER A LABEL IDENTIFIED IN HINDSIGHT

There is a second way in which data are commonly taken out of context; in which they are given meaning from the outside. This is the grouping individual fragments of behavior that represent some common condition.

Consider this example, where diverse fragments of behavior are lumped together to build a case for haste as explanation of the bad decisions taken by the crew. The fragments are actually not temporally co-located. They are spread out over a considerable time, but that does not matter. According to the investigation they point to a common condition.

”Investigators were able to identify a series of errors that initiated with the flightcrew's acceptance of the controller's offer to land on runway 19…The CVR indicates that the decision to accept the offer to land on runway 19 was made jointly by the captain and the first officer in a 4-second exchange that began at 2136:38. The captain asked: 'would you like to shoot the one nine straight in?' The first officer responded, 'Yeah, we'll have to scramble to get down. We can do it.' This interchange followed an earlier discussion in which the captain indicated to the first officer his desire to hurry the arrival into Cali, following the delay on departure from Miami, in an apparent to minimize the effect of the delay on the flight attendants' rest requirements. For example, at 2126:01, he asked the first officer to 'keep the speed up in the descent'… The evidence of the hurried nature of the tasks performed and the inadequate review of critical information between the time of the flightcrew's acceptance of the offer to land on runway 19 and the flight's crossing the initial approach fix, ULQ, indicates that insufficient time was available to fully or effectively carry out these actions. Consequently, several necessary steps were performed improperly or not at all”. (Aeronautica Civil, 1996, p. 29)

As one result of the runway change and self-imposed workload the flight crew also ”lacks situation awareness”—an argument that is also constructed by grouping voice utterance fragments from here and there:

”…from the beginning of their attempt to land on runway 19, the crew exhibited a lack of awareness…. The first officer asked 'where are we', followed by 'so you want a left turn back to ULQ. The captain replied, 'hell no,

let's press on to... and the first officer stated 'well, press on to where though?'.... Deficient situation awareness is also evident from the captain's interaction with the Cali air traffic controller".[1]

It is easy to pick through the evidence of an accident sequence and look for fragments that all seem to point to a common condition. The investigator treats the voice record as if it were a public quarry to select stones from, and the accident explanation the building he needs to construct from those stones. Among investigators this practice is sometimes called "cherry picking"—selecting those bits that help their *a-priori* argument. The problems associated with cherry picking are many:

- You probably miss all kinds of details that are relevant to explaining the behavior in question;
- Each cherry, each fragment, is meaningless outside the context that produced it. Each of the bits that gets lumped together with other "similar" ones actually has its own story, its own background, its own context and its own reasons for being. When it was produced it may have had nothing to do with the other fragments it is now grouped with. The similarity is entirely in the eye of the retrospective beholder.
- Much performance, much behavior, takes place *in between* the fragments that the investigator selects to build his case. These intermediary episodes contain changes and evolutions in perceptions and assessments that separate the excised fragments not only in time, but also in meaning.

Thus, the condition that binds similar performance fragments together has little to do with the circumstances that brought each of the fragments forth; it is not a feature of those circumstances. It is an artifact of you as investigator. The danger is that you come up with a theory that guides the search for evidence about itself. This leaves your investigation not with findings, but with tautologies. What is the solution?

PUT DATA INTO CONTEXT

Taking data out of context, either by:

---

[1] Aeronautica Civil, op. cit., pages 33-34.

- micro-matching them with a world you now know to be true, or by
- lumping selected bits together under one condition identified in hindsight

robs data of its original meaning. And these data out of context are simultaneously given a new meaning—imposed from the outside and from hindsight. You impose this new meaning when you look at the data in a context you *now* know to be true. Or you impose meaning by tagging an outside label on a loose collection of seemingly similar fragments.

But to understand the actual meaning that data had at the time and place it was produced, you need to step into the past yourself.

Historian Barbara Tuchman put it this way: "Every scripture is entitled to be read in the light of the circumstances that brought it forth. To understand the choices open to people of another time, one must limit oneself to what they knew; see the past in its own clothes, as it were, not in ours."[1]

When left in the context that produced and surrounded it, human behavior is inherently meaningful. It also does not need to be placed in after-the-fact worlds made up of the things and rules people apparently did not take notice of. Behavior makes inherent sense when relocated in the stream of assessments, actions and circumstances of which it was a fundamental part.

To make sense, behavior also does not require large psychological labels tagged on from the outside. To make sense of controversial behavior, you must not start with a theory and then pick cherries from the evidence to support it—the risk of having it wrong, of missing the real explanation, are just too large. Instead, start with the situation in which the behavior took place, and put the controversial fragments back in there. The next chapter takes you through the steps necessary for such relocation. It takes you inside the "tunnel" of the situation in which other people found themselves.

---

[1] Tuchman, B. (1981). *Practicing history: Selected essays*. New York: Norton, page 75.

7. Human Error— The New View

PEOPLE CREATE SAFETY IN COMPLEX SYSTEMS

The new view on human error sees the complex, dynamic systems in which people work as not basically safe at all. In fact, these systems themselves are inherent contradictions between safety and all kinds of other pressures. There are economic pressures; pressures that have to do with schedules, slots, competition, customer service, public image.

An airline pilot who was fired after refusing to fly during a 1996 ice storm, was awarded 10 million dollars by a jury. The pilot, who had flown for 10 years for the airline, was awarded the money in a lawsuit contending that he had been fired for turning around his turboprop plane in a storm. The pilot said he had made an attempt to fly from Dallas to Houston but returned to the airport because he thought conditions were unsafe.[1]

A hero of the jury (themselves potential passengers probably), this pilot could have decided to press on. But if something had happened to the aircraft as a result of icing, the investigation would probably have returned the finding of "human error", saying that the pilot knowingly continued into severe icing conditions. His trade-off must be understood against the backdrop of a turboprop crash in his company only a few years earlier—severe icing was blamed in that case.

Trade-offs such as the one above have to made in circumstances where evidence is often unclear, or where it may be shifting.

Testing for prostate-specific antigen levels (PSA) in all men above the age of

[1] International Herald Tribune, 15 January 2000.

65 would be a really good idea—in fact it would probably help reveal prostate cancer at early stages in many cases. PSA is a relatively easy test, as it is based on a blood sample. However, prostrate cancer is actually more prevalent than fatal in this age group: obductions of men who died of other causes often show prostate cancer to some extent.

Yet catching the cases that could turn out fatal before they do any harm would have positive effects on both patients and those paying for their healthcare. It is easy to call an undiagnosed case of prostate cancer "human error", but the complexity that lies behind an undiagnosed case must be understood in terms of this trade-off. Do we test? Don't we test? Do we operate? Don't we operate? Economic pressures enter into this trade-off as well—blanket testing for PSA levels across populations is not cheap. And then there is the residual uncertainty: a high PSA level indicates a greater risk of prostate cancer, but is in itself no diagnosis.

Pressures and uncertainties do not just reside passively in an organization, to be decided upon by management. As dilemma's and complexities, they get pushed down into individual operating units—cockpits, operating rooms, ships bridges, truck cabs—for practitioners to sort out on the line. These pressures enter, unrecognizably or not, into thousands of little and larger decisions and trade-offs and considerations that practitioners make every day. Will we depart or won't we? Will we push on or won't we? Will we operate or won't we? Will we accept the direct or won't we? Will we accept this display or alarm as indication of trouble or won't we? What this means is that:

COMPLEX SYSTEMS ARE NOT BASICALLY SAFE

PEOPLE HAVE TO CREATE SAFETY BY NEGOTIATING AMONG MULTIPLE SYSTEM GOALS

In the new view on human error:

- People are vital to creating safety. They are the only ones who can negotiate between safety and other pressures in actual operating conditions;
- Human errors do not come unexpectedly. They are the inevitable by-product of human expertise—the human ability to conduct these negotiations while faced with uncertain evidence and uncertain outcomes.

The new view on the role of technology

How does the new view on human error look at the role of technology? New technology does not remove the potential for human error, but changes it. New technology can give a system and its operators new capabilities, but inevitably brings new complexities too. New technology can lead to an increase in operational demands by allowing the system to be driven faster; harder; longer; more precisely or minutely; in lousier weather. Although first introduced as greater protection against failure (more precise approaches to the runway with a Head-Up-Display, for example), the new technology allows a system to be driven closer to its margins, eroding the safety advantage that was gained.

New technology is also often ill-adapted to the way in which people do or did their work, or to the actual circumstances in which people have to carry out their work, or to other technologies that were already there. New technology often forces practitioners to tailor it in locally pragmatic ways, to make it work in real practice. New technology shifts the ways in which systems break down. It asks people to acquire more knowledge and skills, to remember new facts. It adds new vulnerabilities that did not exist before. It can open new and unprecedented doors to system breakdown. The new view of human error maintains that:

- People are the only ones who can hold together the patchwork of technologies introduced into their worlds; the only ones who can make it all work in actual practice;
- It is never surprising to find human errors at the heart of system failure because people are at the heart of making these systems work in the first place.

INVESTIGATIONS AND THE NEW VIEW ON HUMAN ERROR

In the new view, investigations are driven by one unifying principle:

HUMAN ERRORS ARE SYMPTOMS OF DEEPER TROUBLE

Investigations are not interested in human error per se. They are inter-

ested in what the error points to. What are the sources of people's difficulties? Investigations target what lies behind the error—the organizational trade-offs pushed down into individual operating units; the effects of new technology; the complexity buried in the circumstances surrounding human performance; the nature of the mental work that went on in difficult situations; the way in which people coordinated or communicated to get their jobs done; the uncertainty of the evidence around them.

Why are investigations in the new view interested in these things? Because this is where the action is. If people want to learn anything of value about the systems they operate, they will look at human errors as:

- A window on a problem that every practitioner in the system might have;
- A marker in the system's everyday behavior, and an opportunity to learn more about organizational, operational and technological features that create error potential.

Recommendations in the new view:

- Are hardly ever about individual practitioners, because their errors are a symptom of systemic problems that everyone may be vulnerable to;
- Do not rely on tighter procedures because humans need the discretion to deal with complex and dynamic circumstances for which pre-specified guidance is badly suited;
- Do not get trapped in promises of new technology. Although it may remove a particular error potential, new technology will likely open new doors to system breakdown;
- Try to address the kind of systemic trouble that has its source in organizational decisions, workplace conditions or technological features.

PROGRESS ON SAFETY

The new view of human error holds the key to progress on safety. Investigations according to the new view lead to underlying difficulties in the way and circumstances in which people work, and in the tools they operate. Error is no longer seen as a thing in itself, as something that is alien to the system. Errors are symptoms of deeper trouble in

the way people do the work they do every day: pursuing system goals like schedule, customer service, economics, while negotiating with safety.

So what investigations in the new view see is behavior—not error. They see people's everyday behavior. And they typically discover how this behavior was reasonable given the goals that people were pursuing, the evidence and knowledge they had available, the trade-offs they faced, the strategies they had developed, the pressures that existed around them. This means that:

THE POINT OF AN INVESTIGATION IS NOT TO FIND WHERE PEOPLE WENT WRONG

IT IS TO UNDERSTAND WHY THEIR ASSESSMENTS AND ACTIONS SEEMED RIGHT AT THE TIME

In the new view, "human error" is little more than an artifact of our hindsight. It is no more than a label that we put on certain fragments of behavior after the fact. The logical conclusion of the new view is that there is no such thing as human error. The fragments of behavior we call "error" in hindsight have no identifiable counterpart in the actual situation in which that behavior occurred. There, in that situation, behavior was locally rational—it made sense given what people were trying to accomplish, and given the circumstances in which they were doing their work.

The Field Guide intends to help you with investigating human error according to the new view. It intends to help you identify how people's assessments and actions actually made sense (or at least some sense) given the circumstances. To do so, it intends to help you find the connections between these assessments and actions on the one hand, and features of peoples tasks, tools and environment on the other. Because that is where the action is. In complex, dynamic systems, that is where the real sources of trouble lie.

8. Reconstruct The Unfolding Mindset

To reconstruct people's unfolding mindset, you have to understand:

- how their process and other circumstances unfolded around them;
- how people's assessments and actions evolved in parallel with their changing situation;
- how features of people's tools and tasks and their organizational and operational environment influenced their assessments and actions inside that situation.

This chapter and the next two take you through such reconstruction. They lay out five steps towards the reconstruction of unfolding mindset, help you get the human factors data you need, and guide you through the rubble to find the right pieces of evidence.

FIVE STEPS TO RECONSTRUCTION

Remember the tunnel of chapter 2? Figure 8.1 shows what it looked like. You want to attain the perspective of people inside the tunnel; the people whose assessments and actions you are now investigating. You want to see the unfolding world from their local point of view.

The tunnel makes two points about the unfolding mindset of the one inside of it:

- The assessments and decisions that people on the inside make, are based on what they see on the inside of the tunnel. In other words, assessments and actions make sense on the basis of the circumstances surrounding them; they get made on the basis of how the world looked there and then.
- The meandering of the tunnel reflects how behavior stretches over time, and people's assessments and actions are usually not isolated one-shot solutions to single problems. Rather, each assessment and each action relies in part on earlier understandings of a situation. And each assessment points forward to how it is expected to develop in the future.

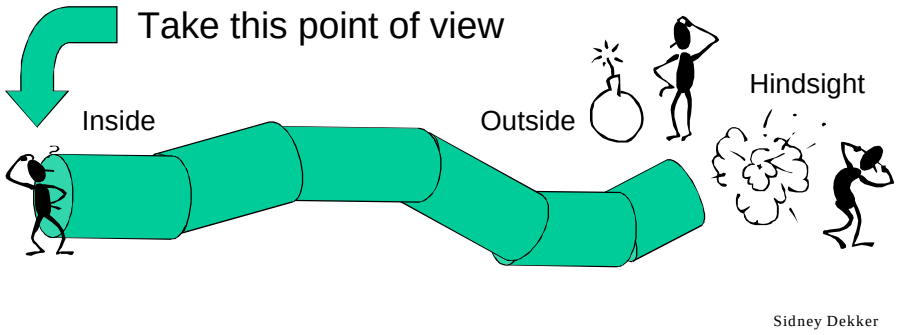


Fig. 8.1: See the unfolding world from the point of view of people inside the situation—not from the outside or from hindsight.

How do you get to the completed tunnel? You reconstruct it by going through five steps. These steps interact and inform one another; it would be impossible to make just one pass through them and be done. To gradually reconstruct a tunnel whose inside looks like the reality of the people who were in it at the time, you may have to jump back and forth between these steps. You may have to loop back around, or repeat sub-parts. Here are the five steps:

1. Mark the beginning and the end of the sequence of events you want to investigate.
2. Lay out the junctures in this sequence of events where things took a different turn or could have taken a different turn.
3. Reconstruct the situation around each juncture as it would have surrounded people on the inside, for example in terms of process indications that would have been available and any operational and organizational pressures that existed.
4. Identify the tasks people were carrying out while crossing these junctures, and what goals they were pursuing. This reveals which of the available indications would actually have mattered, and how operational demands would have received most attention.
5. See how features of people's tools and tasks and their organizational and operational environment influenced their assessments and actions at each of the junctures.

1. MARK THE BEGINNING AND END OF A SEQUENCE OF EVENTS

It may seem an obvious step to take in any analysis—bound the event under investigation by marking the start and the finish. Yet many investigations do not explicitly say where in a sequence of events their work really begins and where it

ends. The issue is often decided implicitly by the availability of evidence.

For example, the beginning of a cockpit voice recording may be where investigative activities start for real, and the end of the recording where they end. Or the beginning is contained in the typical 72-hour or 24-hour histories of what a particular practitioner did and did not do (play tennis, sleep well, wake up early, etc.) before embarking on the fatal journey or operation. Of course even these markers are arbitrary, and the reasons for them are seldom made clear.

One reason for not explicitly indicating the start of an investigation is the inherent difficulty in deciding what counts as the beginning (especially the beginning—the end of a sequence of events often speaks for itself). This difficulty was explained in the discussion on causes and the fallacy of the root cause in chapter one. There is no such thing as a root cause—so technically there is no such thing as the beginning of a mishap.

Yet as an investigator you need to start somewhere. Making clear where you start and explaining this choice is the first step toward a structured, well-engineered human error investigation. Take as your beginning the first assessment, decision or action by people close to the mishap—the one that, according to you, set the sequence of events in motion. Such a decision may be the pilot's acceptance of a runway change that led to trouble later on; the resident surgeon's decision to accept an emergency tracheotomy.

These assessments and actions can be seen as a trigger for the unfolding series of events that follows. Of course the trigger itself has a reason, a background, that extends beyond the mishap sequence —both in time and in place. The whole point of taking a proximal assessment or action as starting point is not to ignore these backgrounds, but to identify concrete points to begin your investigation into them. This also allows you to deal with any controversy that may surround your choice of starting point.

Was the pilot's acceptance of a runway change the trigger of trouble? Or was it the air traffic controller's dilemma of having too many aircraft converge on the airport at the same time—something that necessitated the runway change?

Someone can always say that another decision or action preceded the one you marked as your starting point. This is a reminder of what to take into account when analyzing the decision or action you have marked as the beginning. What went on before that? Whatever your choice of beginning, make it explicit. From there you can reach back into history, or over into surrounding circumstances, and find explanations for the decision or action that, according to you, set the

sequence of events in motion. Look at figure 8.2. This is what you want to end up with—a marked beginning and end to the sequence of events you wish to investigate.

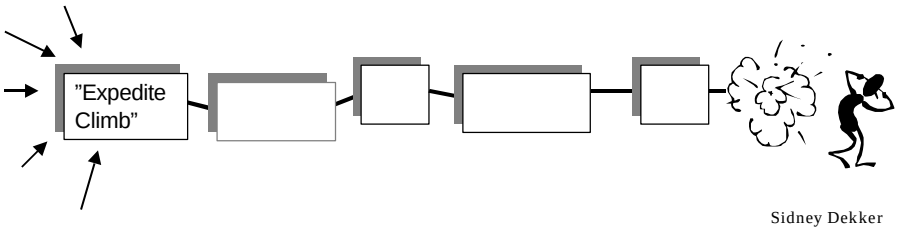
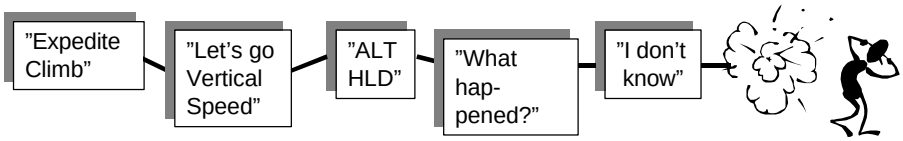


Fig. 8.2: Marking the beginning and end of a sequence of events. Note the various causal influences (which other people could see as triggers or beginnings) on what is marked as beginning here: the request to speed up a climb.

2. LAY OUT THE JUNCTURES IN A SEQUENCE OF EVENTS

The mind is not the starting point for understanding the human contribution to a sequence of events that led up to failure. The unfolding situation in which the human mind found itself is. This means you need to reconstruct the situation as it evolved around the people you are investigating.

But there is a prior step. What do you organize your reconstruction around; what do you base it on? Take the beginning and end of the sequence of events from step 1. Then lay out what happened in between. Find all the important assessments, decisions, actions and changes in the process and lay them out in order. What you will probably see (indeed with the benefit of your hindsight) is that this series of actions and decisions and changes does not head straight for the outcome. It meanders, it twists and turns, just like the tunnel in the previous chapter. People may change course when they assess the evidence unfolding around them; they make decisions to go this way or that. Or the process they are managing takes a turn towards or away from the outcome itself (e.g. an engine flames out due to high angle of attack; the automation reverts to a different operating mode). See in figure 8.3 how this would look—for a shortened, hypothetical sequence of events.



Sidney Dekker

Fig. 8.3: Laying out the complete sequence of events, including people's assessments and actions and changes in the process itself (here for example an automation mode change to altitude hold mode).

In this meandering towards the outcome you can locate junctures. These are points where:

- The sequence of events took a turn towards the outcome.
- The sequence of events momentarily veered away from the outcome.
- The sequence of events could have taken a turn away from the outcome altogether but did not.

How do you locate junctures?

Junctures are places, or stretches of time, where either people or the processes they manage contribute critically to the outcome that followed. Junctures in a sequence of events are places where people did something or could have done something to influence the direction of events. But junctures are also places where the process did something or could have done something to influence the direction of events—whether as a result from human inputs or not.

Junctures are starting points for investigating the backgrounds, reasons and histories behind them. Where did decisions come from? What pushed them one way rather than the other? In other words, these junctures form the organizing thread, for reconstructing the situation that surrounded the people you are investigating.

As a rule, what people did and what their processes did is tightly interconnected—the two rarely develop independently from one another. Where the process makes its contributions (e.g. an automation mode change) people can get different insights, come to different conclusions or move towards particular decisions. Which in turn may influence how the process is managed. This means that discovering changes in one may lead you onto a juncture in the other. Junctures in a sequence of events towards failure can be identified by cross-examining people's decision, cognitive resets, shifts in behavior or strategy, actions to influence the process, and changes in the process itself. More about these follows below.

Decisions

Decisions can be obvious junctures, particularly when they are made in the open and talked about. For instance, the decision to accept a runway change leaves a trace of communication between air traffic control and pilots, and likely among pilots themselves, and is accompanied by shifts in the process (the aircraft changing course, descent rate, etc.). A decision to not accept a runway change is still a decision, and may also serve as a marker in the sequence of events towards breakdown.

Re-evaluations and cognitive resets

In domains where people's work is dynamic and where evidence about the world around them can shift and change over time, people routinely re-evaluate their circumstances. Are we still on course? What are we headed towards now? Am I going to achieve my goal here? Sometimes these re-evaluations can lead to fundamental insights. Where people thought they were, is not at all where they really were. How safe they thought they were is not at all how safe they really were.

Under certain circumstances these insights can come even on the inside of the tunnel, and will be marked by suddenly different behavior or different strategies (see the next point). If so, such junctures can be marked as "cognitive resets", points where people realized the situation was different from what they believed before.

Shifts in behavior or strategy

Cognitive resets are often accompanied by shifts in people's behavior or in their strategy. For example, a pilot may go from normal to hard braking when he notices he is not going to make a planned runway exit. He may switch off the autopilot when he notices the aircraft is not automatically capturing a localizer or changing to an expected mode.

Such shifts in human behavior, such changes in how people manage their process, can themselves be markers where you want to start looking for cognitive resets, for people's re-evaluations and for the evidence on which they were based. And changes in behavior themselves can in turn be linked to changes you had first noticed in the process. For example, you make a plot of brake pressure and see a spike somewhere, where braking was obviously increased significantly.

Actions to influence the process

Rather than shifts in behavior or strategy that were the result of a realization that things were not as people first believed, actions to influence the process may come from people's own intentions. For example, a pilot may type a particular waypoint

in his or her flight management computer, in order to get the aircraft to fly this way or that. He or she may switch systems, dial radio's or begin a descent.

Evidence for these actions may not originate in the actions themselves, but in process changes that follow from them. That is, you may not have any data record of pilots typing, but you may have readouts of what the autopilot was told to do at a certain time. These actions can serve as important junctures. They not only refer to themselves; they also give you a strong clue about the human's current understanding of the situation—where the human thought he or she was; how he or she wanted to proceed; what evidence or which indications he or she probably relied on.

Changes in the process

Any significant change in the process that people manage must serve as juncture. Not all changes in a process managed by people actually come from people. In fact, increasing automation in a variety of workplaces has led to the potential for autonomous process changes almost everywhere—for example:

- Automatic shut-down sequences or other interventions;
- Alarms that go off because a parameter crossed a threshold;
- Uncommanded mode changes;
- Autonomous recovery from undesirable states or configurations.

But even if they are autonomous, these process changes do not happen in a vacuum. They always point to human behavior around them; behavior that preceded it and behavior that followed it. People may have helped to get the process into a configuration where autonomous changes were triggered. And when changes happen, people notice them or not; people respond to them or not.

The connection between autonomous process changes and people's behavior gives you strong clues about what people understood their current circumstances to be. It can give you clues about people's preferences and priorities—about how they for example integrated operational pressures or historical evidence into their responses to alarms and warnings. The nature of people's reactions can also tell you what system knowledge people may or may not have possessed.

The junctures that were no junctures

Human decisions, actions and assessments can also be less obvious. For example, people seem to decide, in the face of evidence to the contrary, to not change their course of action; to continue with their plan as it is. With your hindsight, you may see that people had opportunities to recover from their misunderstanding of the situation, but missed the cues, or misinterpreted them.

These "decisions" to continue, these opportunities to revise, may look like clear candidates for junctures to you. And they are. But they are junctures only in

hindsight. To the people caught up in the sequence of events there was not any compelling reason to re-assess their situation or decide against anything. Or else they would have. They were doing what they were doing because they thought they were right; given their understanding of the situation; their pressures.

As a juncture in the sequence of events you are laying out, the challenge for you becomes to understand how this was not a juncture to the people you were investigating. How their "decision" to continue was nothing more than continuous behavior—reinforced by their current understanding of the situation, confirmed by the cues they were focusing on, and reaffirmed by their expectations of how things would develop in the near future.

3. RECONSTRUCT THE SITUATION AT EACH JUNCTURE

When the people you are investigating did what they did, they inhabited a certain world. A world was unfolding around them. It showed indications about the status of their processes. Parameters were changing over time, both as a result of human influences and of the process moving along—changing pressures, ratios, settings, altitudes, quantities, modes, rates. The values of these parameters were likely available to people in all kinds of ways—dials, displays, knobs that pointed certain ways, sounds, mode annunciations, alarms, warnings.

Identifying the connections between these changing parameters on the one hand and what people thought and decided and did on the other, gets you toward coupling behavior and situation—toward putting the observed behavior back into the situation that produced and accompanied it. Step three is about reconstructing this coupling in its most direct sense, by using indications that were directly available in the world the people inhabited at the time. Step three is about taking the junctures from step two and relating them to how you know the world was unfolding around people at those times.

Laying out how some of the critical parameters changed over time is nothing new to investigations. Many accident report appendices contain read-outs from data recorders, which show the graphs of known and relevant process parameters. But building these pictures is often where investigations stop today. Tentative references about connections between known parameters and people's assessments and actions are sometimes made, but never in a systematic, or graphic way.

The point of step three is to marry all the junctures you have identified above with the unfolding process—to begin to see the two in parallel, as an inextricable, causal *dance-a-deux*. The point of step three is to build a picture that shows these connections; to create a common ground for you and other investigators to enter and begin your probe.

Choosing among datatraces

Many complex, dynamic processes are data-rich. They may leave a huge electronic footprint of parameters behind. This can produce your own data overload. How could everything possibly be relevant to your investigation? Or, if you were to pick some, how could you be sure you were not leaving out critical cues? The problem is to decide which—of all the parameters—counted as a stimulus for the behavior under investigation, and which did not. Which of these indications or parameters, and how they evolved over time, were actually instrumental in influencing the behavior in your mishap sequence? The answer lies in the nature of events itself.

Here are a few examples from the world most richly endowed with devices for tracking and recording process parameters—commercial aviation: If the outcome of the sequence of events was a stall warning, then airspeed, and what it did over time, becomes a relevant parameter to include. If the outcome involves a departure from the hard surface of a runway, then brake pressure is a parameter to focus on. If the outcome was an automation surprise, then the various mode changes the automation went through, including their annunciations, are what you want to get down.

Of course you cannot consider any of these parameters in isolation. People do more than tracking airspeed or braking or watching mode changes, and they will likely have been looking at other things that may have relevance to your sequence of events. You will have to use knowledge of the people involved, or of people like them, or of yourself, to understand what else may have been relevant in this context.

When are you sure you have covered the parameters you need? After going through the reconstruction of people's unfolding mindsets, you may be left with gaps in your explanation of people's assessments and actions. If so, it is time to start looking for some more parameters that could have served as critical stimuli to influence people's understanding and behavior—parameters that did not seem obvious before.

Connecting process and behavior

Once you have decided which process parameters to track in their journey towards the outcome, the next stage is relatively easy. Build a picture. Build a picture of the critical parameters around the junctures you recovered in step two.

The sort of representation does not really matter, and what works best may depend on the kind of sequence of events you are investigating. You can draw graphs of relevant process parameters that go up and down and up again as people decide and act their way towards the outcome. You can draw representations of instruments as they must have looked at the various junctures. If relevant, you can

draw a geographical map of the area in which things played out, and record people's positions and assessments and actions in it over time.
Whatever the representational form, in it you must account for all the decision points, the cognitive resets, the (in hindsight) missed opportunities you have identified previously—in short, all the junctures. This means you also have to insert the changes that emanated from the process—the mode reversions, the alarms, the automatic resets.

With this picture, connections can start to emerge between how the world looked and what people did. You have graphically tied the relevant process parameters to the human assessments and actions that evolved in concert with them. This is where one may begin to explain the other—and vice versa. See figure 8.4 for an example.

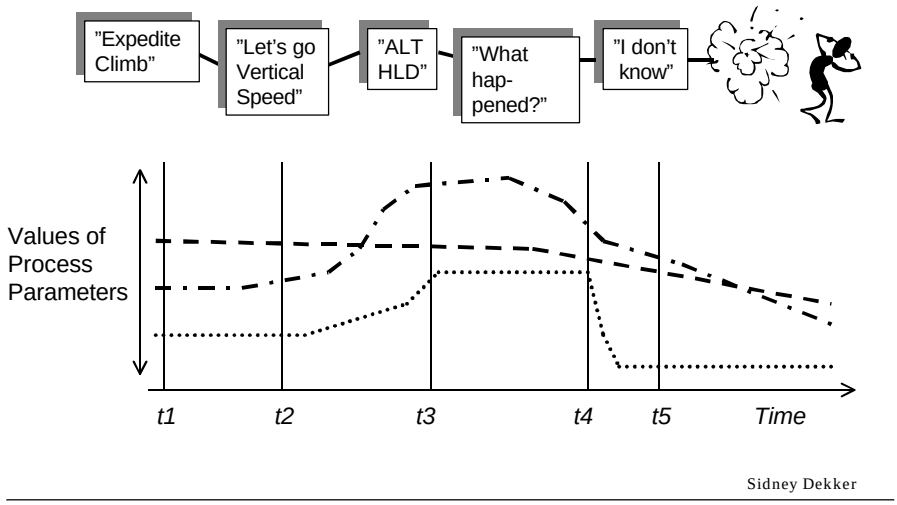


Fig. 8.4: Connecting critical process parameters to the sequence of people's assessments and actions and other junctures.

A question remains here. Out of the critical parameters you have selected and drawn up, what did people actually notice? Where did they look? There is a systematic answer to this, covered in step four.

4. IDENTIFY TASKS AND GOALS

Step three reveals only how relevant process data were physically available. This is relatively easy to show—once you have decided which data traces to follow and

can recover how they evolved over time. But step three—showing data availability—is only one step towards a full explanation of human error. Why did people miss certain things that we know were there? Why did they focus on one indication and not the other?

These questions are not answered by just showing that data were physically available and by recanting counterfactually that people should have noticed them. They have to be answered by an additional step in the reconstruction of unfolding mindset—the drawing of a thread of tasks and goals through the junctures and surrounding situation.

People do not wander through situations aimlessly, simply receiving inputs and producing outcomes as they go along. They are there to get a job done, to accomplish tasks, to pursue goals. If there is anything that determines where people look and how they interpret what they see, it is the goals that they have at the time, and the tasks they are trying to accomplish.

Finding what tasks people were working on does not need to be difficult. It often connects directly to how process parameters were unfolding around them. Setting the navigation systems up for an approach to the airport, for example, is one task that stretches both into what people were saying and doing and to what was happening with the process they managed. Changing a flight plan in the flight management computer is another. To identify what task people were trying to accomplish at any juncture, ask yourself the following questions:

- What is canonical, or normal at this time in the operation? Tasks relate in systematic ways to stages in a process. You can find these relationships out from your own knowledge or from that of (other) expert practitioners in the field.
- What was happening in the managed process? Starting from your record of parameters from step three, you can see how systems were set or inputs were made. These changes obviously connect to the task people were carrying out.
- What were other people in the operating environment doing? People who work together on common goals often divide the necessary tasks among them in predictable or complementary ways. There may be standard role divisions, for example between pilot flying and pilot not-flying, that specify the tasks for each. What one pilot was doing may give some hints about what the other pilot was doing.

If you find that pictures speak more clearly than text, create a graphical representation of the major tasks over time, and if necessary, of who was carrying out what. This picture can also give you a good immediate impression of the kind of workload associated with the sequence of events. See figure 8.5 for an example.

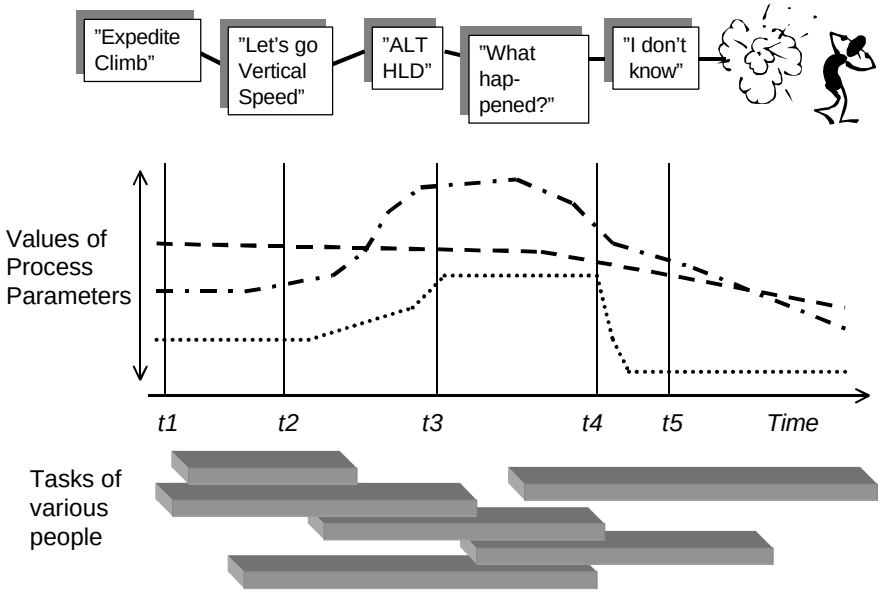


Fig. 8.5: Laying out the various (overlapping) tasks that people were accomplishing during the sequence of events

You can lay the tasks out underneath the picture that emerged from the three previous steps. This combination indicates what people were occupied with during the junctures and changes in process parameters. And once you have an idea what people were occupied with, you can begin to discern what they probably looked at, which parameters they would have found interesting, and which would have been irrelevant or secondary. You can also begin to get an idea of how evidence about an unfolding situation got interpreted in relation to the task people were accomplishing.

It can be more difficult to identify the larger goals people were pursuing. In aviation, you hope such an overriding goal is "flight safety". But how do these goals translate to concrete assessments and actions? Sometimes local decisions and actions seem contrary to these goals.

For example, a pilot may do everything to stay visual with an airport where he has just missed an approach. This can lead to all kinds of trouble, for example getting close to terrain, being forced lower by shifting cloud ceilings, getting in conflict with other aircraft, losing bearings, and so forth. So why would anyone do it? In the context in which the pilot was operating, it may actually be an action that lies closest to the goal of flight safety. What kind of country was the airport in? How reliable were the navigation aids around it? How good or understandable were the controllers? How much other traffic was around?

How familiar was the pilot with the area? Was there severe turbulence in the clouds? Given this context, the goal of flight safety takes on a different meaning. Achieving flight safety translates to different assessments and actions under different circumstances—ones that may at first seem counterintuitive or counterprocedural.

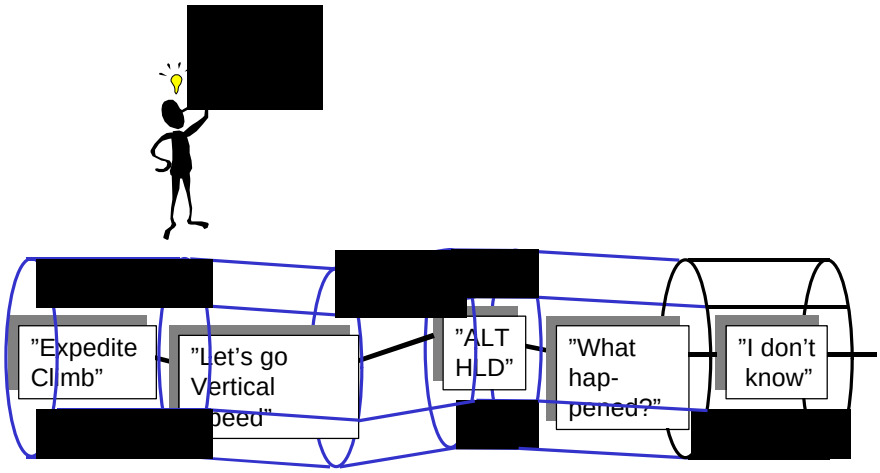
So understanding the goals people were pursuing and how they got reflected in concrete assessments and actions is another way to resituate behavior in the situation that surrounded it—the overall aim of reconstructing unfolding mindset. Tasks and goals pull threads through multiple junctures; they connect junctures in meaningful, coherent and reasonable ways with one another. This continuation means that what people did at any one moment, or at any one juncture, was determined not just by what they saw or thought there and then, but also by:

- past assessments of the situation and what they (thought they) were doing about it;
- expectations of how it would develop in the (near) future, given what they were doing about it

Assessments and actions at one juncture may refer back to those at an earlier junction. Or they may point ahead to what people were going to do, or how they understood their situation to become. Taking the view of someone inside the situation, then, also means giving yourself the ability to look backward and forward inside of it. This may help you comprehend why what people did actually made sense.

5. IDENTIFY OTHER INFLUENCES ON ASSESSMENTS AND ACTIONS

Remember that the target of reconstructing unfolding mindset is to find out why actions and assessments made sense to people at the time. If you are lucky, this is partly a solved problem by now. Around each of the junctures in the sequence of events, you have reconstructed what the process looked like. You have been doing what is shown in figure 8.6—covering the tunnel with bits and pieces you have found in the rubble; reconstructing the world as it looked to people on the inside. You have recovered the tasks people were pursuing; the goals they had. All of this may have led you to a better understanding of why people did what they did.



Sidney Dekker

Fig. 8.6: Trying to rebuild the tunnel, the way it looked on the inside: reconstructing the situation that surrounded people's assessments and actions and other changes in the process.

Yet in many ways, these steps are just that: steps along the way to reconstructing unfolding mindset. You may still be left with large gaps in the explanation of behavior. One major reason is that, so far, these steps reconnect observed behavior only with directly obvious, more easily available factors—the parameters that were physically available in the operating environment of the people you are investigating.

Human behavior is of course determined by many more factors than process parameters. As a rule, however, other influences are less visible and more difficult to recover from the rubble. Take organizational pressures to choose schedule over safety, for example. Such pressures exist and exert a powerful influence on the many little local trade-offs people make. Yet especially in the aftermath of failure, these factors easily get rationalized away as being irrelevant or insignificant. As in: real professionals should not be susceptible to those kinds of pressures.

Such reactions, however, reveal a profound shortcoming in the understanding of human error. The next chapter is all about reconstructing how the situation looked farther away from people's proximal assessments and actions. It is all about finding the less obvious connections between people's behavior and features of the circumstances in which it took place.

9. Clues In The Rubble

The previous chapter was about reconstructing the unfolding reality that surrounded the people you are investigating. What did their world look like? How did it determine or influence their assessments and actions? And what did these people do to influence the situation in turn; to re-direct the sequence of events?

This chapter takes you into the reconstruction of the deeper, wider situation surrounding these people. It helps you find and probe the factors and reasons that are less immediately visible, but that exert a powerful influence on human behavior. This chapter directs your attention to:

- The history of operations: Have similar situation occurred before?
- The organization: how did it influence trade-offs and decisions?
- The technology: How did its features shape human performance?

LOOK IN HISTORY

Dress rehearsals

The period before a mishap may contain sequences of events that look like the one in the actual accident or incident, but without the same bad outcome. These could be called "dress rehearsals".

In January 1992, a highly automated aircraft crashed into a mountain close to Strasbourg airport in eastern France. Confusion between two automation modes that could each manage the aircraft's descent turned out to have been central in the crash. The pilots intended to make and automatic approach at a flight path angle of 3.3 degrees towards the runway. Due, however, to an internal connection between horizontal and vertical automation modes in the aircraft's computer systems, the aircraft was not in flight path angle mode, but had slipped into vertical speed mode. Pilots have to use the same knob in either mode, so dialing 3.3 resulted in a descent rate of 3300 feet per minute down—much steeper than 3.3 degrees.

During the years preceding this accident, various airlines had had similar sequences of events: pilots flying in Rate of Descent instead of Flight Path Angle mode. In these cases, go-arounds could be made. One airline had even developed some ad-hoc specific

preventative training to avoid just this sort of event, even though it commented that pilots on this fleet were reluctant to admit there might be an ergonomic shortcoming in this cockpit.

Dress rehearsals tell you to look for more systemic contributors to the behavior in question. What are the commonalities? What is the trap that everybody seems to fall into? The contrast between dress rehearsal and actual mishap also shows what it takes to push a system over the edge, and what prevented a complete breakdown earlier.

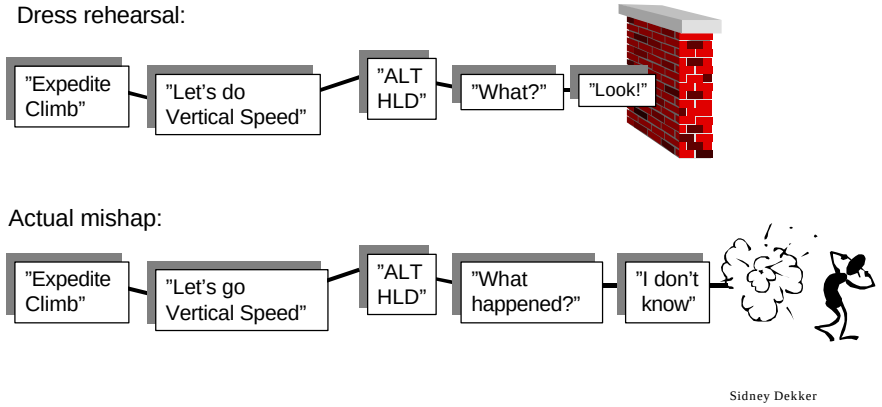


Fig. 9.1: Dress rehearsals for the real mishap can help reveal more fundamental conditions that contribute to this kind of failure

---

The Strasbourg crash happened at night, in snow. It is likely that the dress rehearsals took place in better conditions, where pilots had eye contact with the ground. Also, the airline going into Strasbourg had elected not to install Ground Proximity Warning Systems in its aircraft because of the high false alarm rate in the systems at that time, and the fact that it flew many short missions in mountainous terrain—exacerbating the false alarm problem. One dress rehearsals was kept from disaster by a Ground Proximity Warning .

That dress rehearsals can occur locally without subsequent investments in serious countermeasures also gives you a clue about an industry's perception of risk and danger, and reveals vulnerabilities in its way of sharing safety-critical information with other operators.

Contrast cases

Other mishaps, whether in the same organization or industry or not, can function as contrast cases. These are situations which are largely similar, but where people behaved slightly differently—making other assessments or decisions. This difference is a powerful clue to the reasons for behavior embedded in your situation.

An airliner was urgently requested by air traffic control to use a rapid exit taxiway from the runway on which it had just landed, because of traffic tightly behind it. The airliner could not make the final turn and momentarily slid completely off the hard surface. It reentered another taxiway and taxied to the gate under its own power. Although no procedures existed at the time to tell them otherwise, the airline wondered why the pilots continued taxiing, as the aircraft may have suffered unknown damage to wheels, brakelines, and so forth (although it turned out to be undamaged).

Not long before, the airline had had another incident where a similar aircraft had left the hard surface. This, however, occurred at a small provincial airport, late at night, after the aircraft's and pilots' last flight of the day. Theirs was the only aircraft on the airport. The pilots elected not to taxi to the gate by themselves, but disembarked the passengers right there and had the aircraft towed away. The control tower was involved in the entire operation.

This contrasted sharply with the other case, which happened at the airline's major hub. Many passengers had connecting flights, as did the pilots and their aircraft. It rained heavily, and the wind blew hard, making disembarkation on the field extremely undesirable. People in the control tower seemed not to have noticed the event. Moreover, for the time it would have taken to get busses and a tow truck out to the field, the aircraft would have blocked a major taxiway, all but choking the movements of aircraft landing behind it every two minutes.

LOOK IN THE ORGANIZATION

The number of ways in which organizational features can contribute to failure is unlimited. For example, there can be contributions from:

- The division of responsibilities
- Organizational culture
- Maintenance
- Supervision
- Rules and procedures
- Staff and/or departmental communication

- Contractors
- Planning
- Morale
- Time of day and scheduling
- Commercial and operating pressures
- Training and selection
- Understaffing

Some of these potential sources of error have been treated in other places in the Field Guide. The remainder of this section presents further pointers and examples[1].

Resources and constraints

The ability of people at the sharp end to assess and decide as they see fit given the circumstances, is actually influenced and constrained by resources and pressures that come from the organizational context in which they work. In other words, the blunt end has significant influence on people's abilities to perform well at the sharp end:

A woman was hospitalized with severe complications of an abdominal infection. A few days earlier, she had seen a physician with complaints of aches, but was sent home with the message to come back in eight days for an ultrasound scan if the problem persisted. In the meantime, her appendix burst, causing infection and requiring major surgery. The woman's physician had been under pressure from her managed care organization, with financial incentives and disincentives, to control the costs of care and avoid unnecessary procedures.[2] The problem is that a physician might not know that a procedure is unnecessary before doing it, or at least doing part of it. Pre-operative evidence may be too ambiguous. Physicians end up in difficult double binds, created by the various organizational pressures.

Goal conflicts

Although "safety" is almost always cited as an organization's overriding goal, it is never the only goal (and in practice not even a measurably overriding goal), or the organization would have no reason to exist. People who work in these systems

---

1 See for a more thorough discussion: Reason, J. T. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate.

2 International Herald Tribune, 13 June 2000.

have to pursue multiple goals at the same time, which often results in goal conflicts. The trade-off between safety and schedule is often mentioned as prime example. But goal conflicts can also arise from the nature of the work itself:

Anesthesiology presents interesting inherent goal conflicts. On the one hand, anesthesiologists want to protect patient safety and avoid being sued for malpractice afterward. This maximizes their need for patient information and pre-operative workup. But hospitals continually have to reduce costs and increase patient turnover, which produces pressure to admit, operate and discharge patients on the same day. Other pressures stem from the need to maintain smooth relationships and working practices with other professionals (surgeons, for example), whose schedules interlock with those of the anesthesiologists.[1]

The complexity of these systems, and of the technology they employ, can also mean that one kind of safety needs to be considered against another:

The space shuttle Challenger broke up and exploded shortly after lift-off in 1986 because hot gases bypassed O-rings in the booster rockets. The failure has often been blamed on the decision that the booster rockets should be segmented (which created the need for O-rings) rather than seamless "tubes". Segmented rockets were cheaper to produce—an important incentive for an increasingly cash-strapped operation.

The apparent trade-off between cost and safety hides a more complex reality where one kind of safety had to be traded off against another—on the basis of uncertain evidence and unproven technology. The seamless design, for example, could probably not withstand predicted prelaunch bending moments, or the repeated impact of water (which is where the rocket boosters would end up after being jettisoned from a climbing shuttle). Furthermore, the rockets would have to be transported (probably over land) from manufacturer to launch site: individual segments posed significantly less risk along the way than a monolithic structure filled with rocket fuel.[2]

Defenses breached

---

1 See: Woods, D. D., Johanssen, L. J., Cook, R. I., & Sarter, N. B. (1994). Behind human error: Cognitive systems, computers and hindsight. Dayton, OH: CSERIAC, page 63.

2 Vaughan, D. (1996). *The Challenger lauch decision.* Chicago, IL: University of Chicago Press.

An increasingly popular way to think about pathways to failure is in the form of the breaching or by-passing of defenses. As explained in chapter 3, safety-critical organizations invest heavily in multiple layers of defense against known or possible failure trajectories. If failures do happen, then something has to be wrong with these layers of defense.

The story of the escape of huge amounts of methyl isocyanate (MIC) from Union Carbide's pesticide plant in Bhopal, India, in 1984 is one of many by-passed, broken, breached or non-existent defenses. For example, instrumentation in the process control room was inadequate: among other things had its design not taken extreme conditions into account: meters pegged (saturated) at values far below what was actually going on inside the MIC tank. Defenses that could have stopped or mitigated the further evolution of events were compromised or simply not there. For example, none of the plant operators had ever taken any emergency procedures training. The tank refrigeration system had been shut down and was now devoid of liquid coolant; the vent gas scrubber was designed to neutralize escaping MIC gasses of quantities 200 less and at lower temperatures than what was actually escaping; the flare tower (that would burn off escaping gas and was itself intact) had been disconnected from the MIC tanks because maintenance workers had removed a corroded pipe and never replaced it. Finally, a water curtain to contain the gas cloud could reach only 40 feet up into the air, while the MIC billowed from a hole more than 100 feet up.

Investigating which layers of defense were breached or by-passed reveals more than just the reasons for a particular failure. The existence of defenses (or the holes you find in them) carry valuable information about the organization's current beliefs, and the nature of its understanding about vulnerabilities that threaten safety. This can open up opportunities for more fundamental countermeasures (see chapter 11).

Contributions from regulators

Most safety-critical industries are regulated in some way. With the specific data of an accident in hand, it is always easy to find gaps where the regulator "failed" in its monitoring role. This is not a very meaningful finding, however. Identifying regulatory oversights in hindsight does not explain the reasons for those—what now look like—obvious omissions. Local workload, the need to keep up with ever-changing technologies and working practices and the fact that the narrow technical expertise of many inspectors can hardly foresee the kinds of complex, interactive sequences that produce real accidents, all conspire against a regulator's ability to exercise its role. If you feel you have to address the regulator in your investigation, do not look for where they went wrong. As with investigating the assessments and actions of operators, find out how the regulator's trade-offs, perceptions and judgments made local sense at the time; why what they were doing or looking at was the right thing given their goals, resources, and

understanding of the situation.

Another complaint often leveled against regulators is that they collude with those they are supposed to regulate, but this is largely a red herring (and, interestingly, almost universally disagreed with by those who are regulated. Independent of claims to collusion, they often see regulators as behind the times, intrusive and threatening). To get the information they need, regulators are to a large extent dependent on the organizations they regulate, and likely even on personal relationships with people in those organizations. The choice, really, is between creating an adversarial atmosphere in which it will be difficult to get access to required sources of safety-related information, or one in which a joint investment in safety is seen as in everybody's best interest.

LOOK AT THE TECHNOLOGY

Human work in safety-critical domains has almost without exception been work with technology. Today, it is more and more work with computers. This means that human-computer interaction is an increasingly dominant source of error. Computer technology has shaped and influenced the way in which people make errors. It has also affected people's opportunities to detect or recover from the errors they make and thus, in cases, accelerated their journeys towards breakdown.

As is the case with organizational sources of error, human-computer errors are not random. They too are systematically connected to features of the tools that people work with and the tasks they have to carry out. Here is a guide[1], first to some of the "errors" you may typically find in the rubble of the human error mishap. Then a list of computer features from which these errors originate, and then a list of some of the cognitive consequences of computerization that lie behind the creation of those errors. This chapter concludes with some observations about the connection between organizational pressures and new technology.

Typical errors

If people were interacting with computers in the events that led up to the mishap, look for the possibility of the following "errors":

- **Mode error**. The user thought the computer was in one mode, and did the

[1] Much material for this section comes from Woods, D. D., Johanssen, L. J., Cook, R. I., & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers and hindsight*. Dayton, OH: CSERIAC, and Dekker, S. W. A., & Hollnagel, E. (Eds.) (1999). *Coping with Computers in the Cockpit*. Aldershot, UK: Ashgate.

right thing had it been in that mode, yet the computer was actually in another mode;

- **Getting lost** in display architectures. Computers often have only one or a few displays, but a potentially unlimited number of things you can see on them. Thus it may be difficult to find the right page or data set;
- **Not coordinating computer entries**. Where people work together on one (automated) process, they have to invest in common ground by telling one another what they tell the computer, and double-checking each other's work. Under the pressure of circumstances or constant meaningless repetition, such coordination may not happen consistently
- **Overload**. Computers are supposed to off-load people in their work. But often the demand to interact with computers concentrates itself on exactly those times when there is already a lot to do; when other tasks or people are also competing for the operator's attention. You may find that people were very busy programming computers when other things were equally deserving of their attention;
- **Data overload**. People were forced to sort through a large amount of data produced by their computers, and were unable to locate the pieces that would have revealed the true nature of their situation. Computers may also spawn all manner of automated (visual and auditory) warnings which clutter a workspace and proliferate distractions.
- **Not noticing changes**. Despite the enormous visualization opportunities the computer offers, many displays still rely on raw digital values (for showing rates, quantities, modes, ratios, ranges and so forth). It is very difficult to observe changes, trends, events or activities in the underlying process through one digital value clicking up or down. You have to look at it often or continuously, and interpolate and infer what is going on;
- **Automation surprises** are often the end-result: the system did something that the user had not expected. Especially in high tempo, high workload scenarios, where modes change without direct user commands and computer activities are hard to observe, people may be surprised by what the automation did or did not do.

Computer features

What are some of the features of today's computer technology that contribute systematically to the kinds of errors discussed in the section above?

- Computers can make things "invisible"; they can hide interesting changes and events, or system anomalies. The presentation of digital values for critical process parameters contributes to this "invisibility". The practice of showing only system *status* (what mode it is in?) instead of *behavior* (what is the system actually doing, where is it going?) is another reason. The interfaces can look simple, but they really hide a lot of complexity.

- Computers, because they only have one or a few interfaces (this is called the "keyhole problem), can force people to dig through a series of display pages to look for, and integrate, data that really are required for the task in parallel. A lot of displays is not the answer to this problem of course, because then navigation across displays becomes an issue. Rather, each computer page should present aids for navigation (How did I get here? How do I get back? What is the related page and how do I get there?). If not, input or retrieval sequences may seem arbitrary, and people will get lost.
- Computers can force people into managing the interface (How do I get to that page? How do we get it into this mode?) instead of managing the safety-critical process (something the computer was promised to help them do). These extra interface management burdens often fall together with periods of high workload;
- Computers can change mode autonomously or in other ways that are not commanded by the user (these mode changes can for example result from pre-programmed logic, much earlier inputs, inputs from other people or parts of the system, and so forth).
- Computers ask people typically in the most rudimentary or syntactic ways to verify their entries (Are you sure you want to go to X? We'll go to X then) without addressing the meaning of their request and whether it makes sense given the situation. And when people tell computers to proceed, it may be difficult to make them stop. All this limits people's ability to detect and recover from their own errors.
- Computers are smart, but not that smart. Computers and automation can do a lot for people—they can almost autonomously run a safety-critical process. Yet computers typically know little about the changing situation around them. Computers assume a largely stable world where they can proceed with their pre-programmed routines even if inappropriate; they dutifully execute user commands that make no sense given the situation; they can interrupt people's other activities without knowing they are seriously bothering.

Cognitive consequences of computerization

The characteristics of computer technology discussed above shape the way in which people assess, think, decide, act and coordinate, which in turn determines the reasons for their "errors":

- Computers increase demands on people's memory (What was this mode again? How do we get to that page?);
- Computers ask people to add to their package of skills and knowledge for managing their processes (How to program, how to monitor, and so forth). Training may prove no match to these new skill and knowledge requirements: much of the knowledge gained in formal training may remain inert (in the head, not practically available) when operators get confronted with the kinds of complex situations that call for its application;

- Computers can complicate situation assessment (they may not show system behavior and lots of digital values) and undermine people's attention management (how you know where to look when);
- By new ways of representing data, computers can disrupt people's traditionally efficient and robust scanning patterns;
- Through the limited visibility of changes and events, the clutter of alarms and indications, extra interface management tasks and new memory burdens, computers increase the risk of people falling behind in high tempo operations;
- Computers can increase system reliability to a point where mechanical failures are rare (as compared with older technologies). This gives people little opportunity for practicing and maintaining the skills for which they are, after all, partly still there: managing system anomalies;
- Computers can undermine people's formation of accurate mental models of how the system and underlying process works, because working the safety-critical process through computers only exposes them to a superficial and limited array of experiences;
- Computers can mislead people into thinking that they know more about the system than they really do, precisely because the full functionality is hardly ever shown to them (either in training or in practice). This is called the knowledge calibration problem;
- Computers can force people to think up strategies (programming "tricks") that are necessary to get the task done. These tricks may work well in common circumstances, but can introduce new vulnerabilities and openings to system breakdown in others.

New technology and operational pressures

Are new technology and operational pressures related to one another? The answer is yes. The introduction of new technology can increase the operational requirements and expectations that organizations impose on people. Organizations that invest in new technologies often unknowingly exploit the advances by requiring operational personnel to do more, do it more quickly, do it in more complex ways, do it with fewer other resources, or do it in less and less favorable conditions.

Larry Hirschorn talks about a law of systems development, which is that every system always operates at its capacity. Improvements in the form of new technology get stretched in some way, pushing operators back to the edge of the operational envelope from which the technological innovation was supposed to buffer them.

In operation Desert Storm, during the Gulf War, much of the equipment employed was designed to ease the burden on the operator, reduce fatigue, and simplify the tasks

involved in combat. Instead these advances were used to demand more from the operator. Almost without exception, technology did not meet the goal of unencumbering the military personnel operating the equipment. Weapon and support systems often required exception human expertise, commitment and endurance. The Gulf War shows that there is a natural synergy between tactics, technology and human factors: effective leaders will exploit every new advance to the limit.[1]

---

[1] Cordesman, A. H., & Wagner, A. R. (1996). The lessons of modern war, Vol. 4: The Gulf war. Boulder, CO: Westview Press.

10. Human Factors Data

The previous chapter has pointed to parts of the evidence that may be promising or interesting. But how do you get to those parts? In other words, how do you get human factors data? This chapter discusses three commonly used routes of access—each with its promises and problems:

- Third-party and historical sources;
- Debriefings of participants themselves;
- Recordings of people's and process performance.

THIRD PARTY AND HISTORICAL SOURCES

Data about people's performance and the reasons behind it can for example be derived from:

- Interviewing peers or others who can give opinions about the people under investigation;
- Scrutinizing training-or other relevant records;
- Documenting what people did in the days or hours leading up to the mishap.

Finding personal shortcomings

In many investigations, these routes to data are used mainly as a process of "elimination"; as a background check to rule out longer-standing vulnerabilities that were particular to the people in question. Most safety-critical systems, however, invest heavily in selection of personnel as well as in on-going monitoring, training and proficiency checking. This means that personal shortcomings on part of individual operators are all but ruled out before they can even touch the controls of any process.

Using third party and historical sources to find out about people's individual features can in fact fuel the bad apple theory—the belief that the system is itself basically safe and has nothing to do with the failure being investigated. The system only contains a few bad apples or broken components, and evidence must be found to show that they were ready to snap anytime. Hindsight seriously biases the search for evidence about people's personal shortcomings. You know where people failed, so you know what to look for, and with enough digging you can probably find it too (real or imagined). This, however, is not very informative. It will trick people into believing the event is only a local hick-up, and divert attention away from more systemic problems that every mishap is bound to contain.

Finding systemic shortcomings

Local shortcomings of individual operators can instead be used as a starting point for probing deeper into the systemic conditions of which their problems are a symptom. Here are some examples:

- From their 72-hour history preceding a mishap, individual operators can be found to have been fatigued. This may not just be a personal problem, but a features of their operation and scheduling—thus affecting a larger proportion of operators;
- Training records may sometimes reveal below average progress or performance by the people who are later caught up in a mishap. But it is only hindsight that connects the two, that enables you to look back from a specific incident and cherry pick putatively associated shortcomings from a historical record at leisure. Finding real or imagined evidence is almost pre-ordained because you come looking for it from a backward direction. But this does not prove any specific causal link with actions or assessments in the sequence of events. Training records are a much more interesting source when screened for the things that all operators got trained on, and how and when, as this explains local performance much better. For example, how were they trained to recognize a particular warning that played a role in the mishap sequence? When were they last trained on this? Answers to these questions my reveal more fundamental mismatches between the kind of training people get and the kind of work they have to do;
- Operators may be found to have been overly concerned with, for example, customer satisfaction. In hindsight this tendency can be associated with a mishap sequence: individuals should have zigged (gone around, done it again, diverted, etc.) instead of zagged

(pressed on because of on-time desires). Colleagues can be interviewed to confirm how customer oriented these operators were. But rather than branding an individual with a particular bias, such findings point to the entire organization that, in subtle or less subtle ways, has probably been sponsoring the trade-offs that favor other system goals over safety—keeping the practice alive over time.

DEBRIEFINGS OF PARTICIPANTS

What seems like a good idea—ask the people involved in the mishap themselves—also carries a great potential for distortion. This is not because operators necessarily have a desire to bend the truth when asked about their contribution to failure. In fact, experience shows that participants are interested in finding out what went wrong and why, which generally makes them forthright about their actions and assessments. Rather, problems arise because of the inherent features of human memory:

- Human memory does not function like a videotape that can be rewound and played again;
- Human memory is a highly complex, interconnected network of impressions, for which it quickly becomes impossible to separate actual events and cues that were observed from later inputs;
- One reason is that the human brain has the propensity to order and structure events more than what they were in the real world; to make events and stories more plausible.

Gary Klein has spent many years refining methods of debriefing people who were caught up in critical incidents: firefighters, pilots, nurses, and so forth. Insights from these methods are valuable to share with investigators of human error mishaps here[1].

The aim of a debriefing

Debriefings of mishap participants are foremost intended to help build

[1] See: Klein, G. (1998). *Sources of power: How people make decisions.* Cambridge, MA: MIT Press.

the tunnel from chapter 8; to reconstruct the situation that surrounded people at the time and to get their point of view on that situation. Some investigations may have access to a re-play of how the world (for example: cockpit instruments, process control panel) looked during the sequence of events, which may look like a wonderful tool. It must be used with caution, however, in order to avoid memory distortions. Klein proposes the following debriefing order:

1. First have participants tell the story from their point of view, without presenting them with any replays that will "freshen up their memory";
2. Then tell the story back to them as investigator. This is an investment in common ground, to check whether you understand the story as the participants understood it;
3. If you had not done so already, identify (together with participants) the critical junctures in the sequence of events (see chapter 8);
4. Progressively probe and rebuild how the world looked to people on the inside of the situation at each juncture. Here it is appropriate to show a re-play (if available) to fill the gaps that may still exist, or to show the difference between data that were available and data that were actually observed.

At each juncture in the sequence of events, you want to get to know:

- Which cues were observed (what did he or she notice/see or did not notice what he or she had expected to notice?)
- What knowledge was used to deal with the situation? Did participants have any experience with similar situations that was useful in dealing with this one?
- What expectations did participants have about how things were going to develop, and what options did they think they have to influence the course of events?
- How did other influences (operational or organizational) help determine how they interpreted the situation and how they would act?

Some of Klein's questions to ask

Here are some questions Gary Klein and his researchers typically ask to find out how the situation looked to people on the inside at each of the critical junctures:

Cues What were you seeing?

	What were you focusing on?
	What were you expecting to happen?
Interpretation	If you had to describe the situation to your fellow crewmember at that point, what would you have told?
Errors	What mistakes (for example in interpretation) were likely at this point?
Previous experience/ knowledge	Were you reminded of any previous experience? Did this situation fit a standard scenario? Were you trained to deal with this situation? Were there any rules that applied clearly here? Did you rely on other sources of knowledge to tell you what to do?
Goals	What goals governed your actions at the time? Were there conflicts or trade-offs to make between goals? Was there time pressure?
Taking action	How did you judge you could influence the course of events? Did you discuss or mentally imagine a number of options or did you know straight away what to do?
Outcome	Did the outcome fit your expectation? Did you have to update your assessment of the situation?

Debriefings need not follow a tightly scripted set of questions, as their relevance depends very much on the event and its investigation. But these suggestions may help you with your walkthrough of the steps above

RECORDINGS OF PERFORMANCE DATA

One thing that human error investigations are almost never short of is wishes for more recorded data, and novel ideas and proposals for capturing more performance data. This is especially the case when mishap participants are no longer available for debriefing.

Advances in recording what people did have been enormous—there has been a succession of recording materials and strategies, data transfer technologies; everything up to proposals to permanently mount video camera's in cockpits and other critical workplaces. In aviation, the electronic footprint that a professional pilot leaves at every flight is huge through automated monitoring systems now

installed in almost every airliner.

Getting these data, however, is only one side of the problem. Our ability to make sense of these data, to reconstruct how people contributed to an unfolding sequence of events, has not kept pace with our growing technical ability to register traces of their behavior. The issue that gets buried easily in people's enthusiasm for new data technologies is that recordings of human behavior—whether through voice (for example Cockpit Voice Recorders) or process parameters (for example Flight Data Recorders)—are never the real or complete behavior.

Recordings represent partial data traces: small, letterbox-sized windows onto assessments and actions that all were part of a larger picture. Human behavior in rich, unfolding settings is much more than the data trace it leaves behind. Data traces point beyond themselves, to a world that was unfolding around the people at the time, to tasks, goals, perceptions, intentions, and thoughts that have since evaporated. The burden is on investigators to combine what people did with what happened around them, but various problems conspire against their ability to do so:

Conventional restrictions

Investigations may be formally restricted in how they can couple recorded data traces to the world (e.g. instrument indications, automation mode settings) that was unfolding around the people who left those traces behind. Conventions and rules on investigations may prescribe how only those data that can be factually established may be analyzed in the search for cause (this is, for example, the case in aviation). Such provisions leave a voice or data recording as only factual, decontextualized and impoverished footprint of human performance.

Lack of automation traces

In many domains this problem is compounded by the fact that today's recordings may not capture important automation-related traces—precisely the data of immediate importance to the problem-solving environment in which many people today carry out their jobs. Much operational human work has shifted from direct control of a process to the management and supervision of a suite of automated systems, and accident sequences frequently start with small problems in human-machine interaction.

Not recording relevant traces at the intersection between people and technology represents a large gap in our ability to understand hu-

man contributions to system failure. For example, flight data recorders in many automated airliners do not track which navigation beacons were selected by the pilots, what automation mode control panel selections on airspeed, heading, altitude and vertical speed were made, or what was shown on either of the pilots' moving map displays. This makes it difficult to understand how and why certain lateral or vertical navigational decisions were made, something that can hamper investigations into CFIT accidents (Controlled Flight Into Terrain—an important category of aircraft mishaps).

THE PROBLEM WITH HUMAN FACTORS DATA

One problem with a human error investigation is the seeming lack of data. You may think you need access to certain process or performance parameters to get an understanding not only of what people did, but why. Solutions to this lack may be technically feasible, but socially unpalatable (e.g. video cameras in workplaces), and it actually remains questionable whether these technical solutions would capture data at the right resolution or from the right angles.

This means that to find out about critical process parameters (for instance, what really was shown on that left operator's display?) you will have to rely on interpolation. You must build evidence for the missing parameter from other data traces that you *do* have access to. For example, there may be an utterance by one of the operators that refers to the display ("but it shows that it's to the left..." or something to that effect) which gives you enough clues when combined with other data or knowledge about their tasks and goals.

Recognize that data is not something absolute. There is not a finite amount of data that you could gather about a human error mishap and then think you have it all. Data about human error is infinite, and you will often have to reconstruct certain data from other data, cross-linking and bridging between different sources in order to arrive at what you want to know.

This can take you into some new problems. For example, investigations may need to make a distinction between factual data and analysis. So where is the border between these two if you start to derive or infer certain data from other data? It all depends on what you can factually establish and how factually you establish it. If there is structure behind your inferences—in other words, if you can show what you did and why you concluded what you concluded—it may not at all be unacceptable to present well-derived data as factual evidence.

11. Writing Recommendations

Coming up with human factors recommendations can be one of the more difficult tasks in an investigation. Often only the shallowest of remedies seem to lie within reach. Tell people to watch out a little more carefully. Write another procedure to regiment their behavior. Or just get rid of the particular miscréants altogether. The limitations of such countermeasures are severe and deep, and well-documented:

- People will only watch out more carefully for so long, as the novelty and warning of the mishap wears off;
- A new procedure will at some point clash with operational demands or simply disappear in masses of other regulatory paperwork;
- Getting rid of the miscréants doesn't get rid of the problem they got themselves into. Others always seem to be waiting to follow in their footsteps.

A human error investigation should ultimately point to changes that will truly remove the error potential from a system—something that places a high premium on meaningful recommendations.

RECOMMENDATIONS AS PREDICTIONS

Coming up with meaningful recommendations may be easier if you think of them as predictions, or as a sort of experiment. Human error is systematically connected to features of the tasks and tools that people work with, and to features of the environment in which they carry out their work. Recommendations basically propose to change some of these features. Whether you want new procedures, new technologies, new training, new safety interlocks, new regulations, more managerial commitment—your recommendations essentially propose to re-tool or re-shape parts of the operational or organizational environment in the hope of altering the behavior that goes on within it.

In this sense your recommendations are a prediction, a hypothesis. You propose to modify something, and you implicitly predict it will have a certain effect on human behavior. The strength of your prediction, of course, hinges on the credibility of the connection you have shown earlier in your investigation:

between the observed human errors and critical features of tasks, tools and environment. With this prediction in hand, you challenge those responsible for implementing your recommendations to go along in your experiment—to see if, over time, the proposed changes indeed have the desired effect on human performance.

High-end or low-end recommendations

So what about those changes? What kinds of changes can you propose that might have some effect on human performance? A basic choice open to you is how far up the causal chain you want your recommended changes to have an impact.

Typical of reactions to failure is that people start very low or downstream. Recommendations focus on those who committed the error, or on other operators like them. Recommendations low in the causal chain aim for example at retraining individuals who proved to be deficient, or at demoting them or getting rid of them in some other way. Other low-end recommendations may suggest to tighten procedures, presumably regimenting or boxing in the behavior of erratic and unreliable human beings.

Alternatively, recommendations can aim high—upstream in the causal chain—at structural decisions regarding resources, technologies and pressures that people in the workplace deal with. High-end recommendations could for example suggest to re-allocate resources to particular departments or operational activities.

This choice—upstream or downstream—is more or less yours as an investigator. And this choice directly influences:

- the ease with which your recommendation can be implemented;
- the effectiveness of your recommended change.

The ease of implementation and the effectiveness of an implemented recommendation generally work in opposite directions. In other words: the easier the recommendation can be sold and implemented, the less effective it will be (see Figure 11.1).

Generally, recommendations for changes low on the causal chain are not very sweeping. They concentrate on a few individuals or a small subsection of an organization. These recommendations are satisfying for people who seek retribution for a mishap, or people who want to "set an example" by coming down on those who committed the errors.

But after implementation, the potential for the same kinds of error is left in the organization or operation. The error is almost guaranteed to repeat itself in some shape or form, through someone else who finds him-or herself in a similar situation. Low-end recommendations really deal with symptoms, not with causes. After their implementation, the system as a whole has not become much wiser or better.

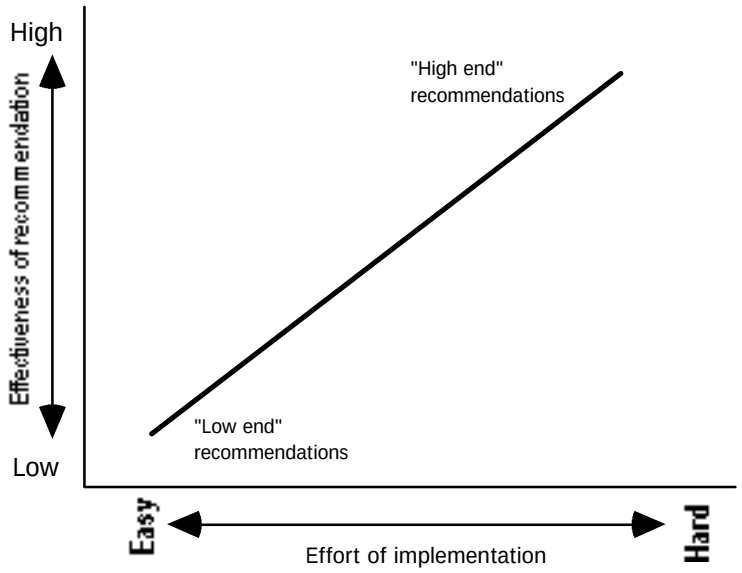


Fig. 11.1: The trade-off between recommendations that will be easier to implement and recommendations that will actually have some lasting effect.

One reason for the illusion that low-end or other narrow recommendations will prevent recurrence is the idea that failure sequences always take a linear path: Take any step along the way out of the sequence, and the failure will no longer occur (see figure 11.2).

In complex, dynamic systems, however, this is hardly ever the case. The pathway towards failure is seldom linear or narrow or simple. Mishaps have dense patterns of causes, with contributions from all corners and parts of the system, and typically depend on many subtle interactions. Putting one countermeasure in place somewhere along

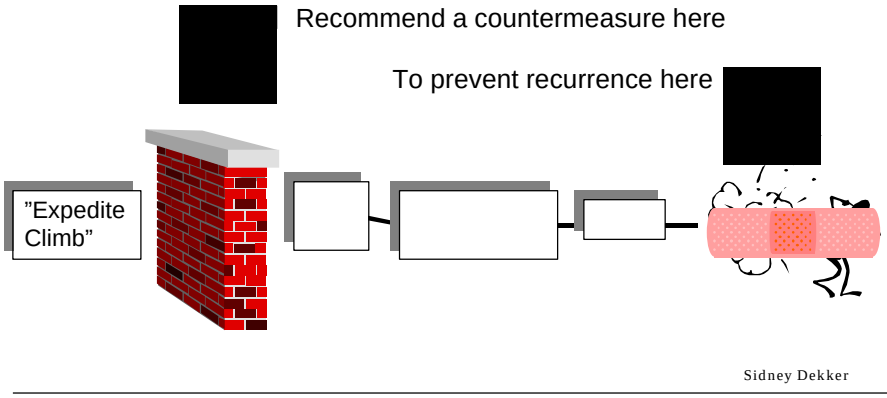


Fig. 11.2: We may believe that blocking a known pathway to failure somewhere along the way will prevent all similar mishaps.

(what you thought was like) a line may not be enough. In devising countermeasures it is crucial to understand the vulnerabilities through which entire parts of a system (the tools, tasks, operational and organizational features) can contribute to system failure under different guises or conditions (see figure 11.3).

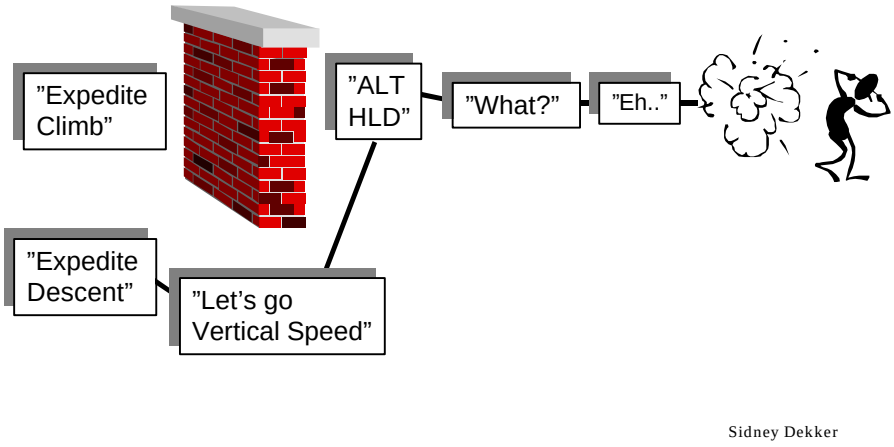


Fig. 11.3: Without understanding and addressing the deeper and more subtle vulnerabilities that drive pathways towards failure, we leave opportunities for recurrence open.

Difficulties with high-end recommendations

The higher you aim in a causal chain, the more difficult it becomes to find acceptance for your recommendation. The proposed change will likely be substantial, structural or wholesale. It will almost certainly be more expensive. And it may concern those who are so far removed from any operational particulars that they can easily claim to bear no responsibility in causing this event or in helping to prevent the next one. Short of saying that it would be too expensive, organizations are good at finding reasons why structural recommendations do not need to be implemented, for example:

- "We already pay attention to that"
- "That's in the manual"
- "This is not our role"
- "We've got a procedure to cover that"
- "This recommendation has no relevance to the mishap"
- "People are selected and trained to deal with that"
- "This is not our problem"

It is easy to be put off as investigator before you even begin writing any recommendations. In fact, many recommendations that aim very high in the causal chain do not come out of first investigations, but out of re-opened inquiries, or ones re-submitted to higher authorities after compelling expressions of discontent with earlier conclusions.

One such case was the crash of a DC-10 airliner into Mount Erebus on Antarctica. The probable cause in the Aircraft Accident Report was the decision of the captain to continue the flight at low level toward an area of poor surface and horizon definition when the crew was not certain of their position. The kinds of recommendations that follow from such a probable cause statement are not difficult to imagine. Tighten procedures; exhort captains to be more careful next time around.

A subsequent Commission of Inquiry determined that the dominant cause was the mistake by airline officials who programmed the aircraft computers—a mistake directly attributable not so much to the persons who made it, but to the administrative airline procedures which made the mistake possible. The kinds of recommendations that follow from this conclusion would be different and aim more at the high end. Review the entire operation to Antarctica and the way in which it is prepared and managed. Institute double checking of computer programming. And so forth.[1]

The case for including or emphasizing high-end recommendations in a first investigation is strong. If anything, it is discouraging to have to investigate the same basic incident or accident twice. Structural changes are more likely to have an effect on the operation as a whole, by removing or foreclosing error traps that

---

[1] See: Vette, G. (1983). *Impact Erebus.* Auckland, NZ: Hodder & Stoughton.

would otherwise remain present in the system.

Remember from chapter 4 that Judge Moshansky's investigation of the Air Ontario crash generated 191 recommendations. Most of these were high-end. They concerned for example[1]:

- Allocation of resources to safety versus production activities;
- Inadequate safety management by airline and authority alike;
- Management of organizational change;
- Deficiencies in operations and maintenance;
- Deficient management and introduction of new aircraft;
- Deficient lines of communication between management and personnel;
- Deficient scheduling (overcommitting this particular aircraft);
- Deficient monitoring and auditing;
- Deficient inspection and control and handling of information;
- Inadequate purchasing of spares;
- Low motivation and job instability following airline merger;
- Different corporate cultures;
- High employee turnover;
- Poor support to operational personnel;
- Inadequate policy making by airline and authority.

These are just some of the areas where recommendations were made. With a serious human error investigation, many of these kinds of conditions can probably be uncovered in any complex system. The ability to generate structural recommendations that aim high up in a causal chain is a reflection of the quality and depth of your understanding of human error.

SEARCHING THE EVIDENCE FOR COUNTERMEASURES

The kind and content of your recommendations depends, of course, on the kind and content of the mishap you are investigating. But to come up with high-end recommendations it may be useful to re-visit some of the organizational contributions to failure from chapter 9. For example:

- The re-allocation of resources that flow from the blunt end, and the

---

[1] Moshansky, V. P. (1992). *Commission of inquiry into the Air Ontario accident at Dryden, Ontario* (Final report, vol. 1-4). Ottawa, ON: Minister of Supply and Services, Canada.

alleviation of constraints that are imposed on operators' local decisions and trade-offs;
- Making goal conflicts explicit and turning them into topics for discussion among those involved;
- Re-invest in the defenses that turned out to be brittle or broken or non-existent;
- Make regulatory access more meaningful through a re-examination of the nature and depth of the relationship between regulator and operator.

Get help from the participants

If possible, it can be fruitful to build on the list above by talking to the participants themselves. These are some of the question that Gary Klein and his researchers ask participants when looking for countermeasures against recurrence of the mishap:

- What would have helped you to get the right picture of the situation?
- Would any specific training, experience, knowledge, procedures or cooperation with others have helped?
- If a key feature of the situation would have been different, what would you have done differently?
- Could clearer guidance from your company have helped you make a better trade-offs between conflicting goals?

Not only can answers to these questions identify countermeasures you perhaps had not yet thought of. They can also serve as a reality check. Would the countermeasures you think about proposing have any effect on the kind of situation you are trying to avoid? Asking the participants themselves, who after all have intimate knowledge of the situation you are investigating, may be a good idea.

12. Learning From Failure

The point of any investigation is to learn from failure. Mishaps, in this regard, are a window of opportunity. The immediate aftermath of a mishap typically creates an atmosphere in which:

- Parts of an organization may welcome self-examination more than before;
- Traditional lines between management and operators, between regulators and operators, may be temporarily blurred in joint efforts to find out what went wrong and why;
- People and the systems they work in may be open to change—even if only for a short while;
- Resources may be available that are otherwise dedicated to production only, something that could make even the more difficult recommendations for change realistic.

Just doing the investigation, however, does not guarantee success in capitalizing on this window of opportunity. Learning from failure is about more than picking over the evidence of something gone wrong. Learning is about modifying an organization's basic assumptions and beliefs. It is about identifying, acknowledging and influencing the real sources of operational vulnerability. This can actually be done even before real failures occur, and the remainder of this chapter is about the opportunities and difficulties of organizational learning—before as well as after failures.

INVESTING IN A SAFETY CULTURE

Safety typically comes to the foreground only at certain moments—the frightening, surprising and generally expensive moments of mishaps. But it does not need to be that way. Signs about safety (or the lack thereof) exist in an organization and operation at any time, and can be identified. The most frequently mentioned key to this is a safety culture.

A SAFETY CULTURE IS ONE THAT ALLOWS THE BOSS TO HEAR BAD NEWS

The "easy" and "hard" problem of a safety culture

Creating a safety culture, however, presents an organization with two problems: an easy one and a hard one. The easy problem (by no means easy, actually, but comparatively straightforward) is to make sure that bad news reaches the boss. Many organizations have instituted safety reporting systems that do exactly that: identifying and addressing problems before they can develop into incidents or accidents.

The hard problem is to decide what is bad news. Chapter 4, which discusses "complacency" as one label for human error, shows that an entire operation or organization can shift its idea of what is normative, and thus shift what counts as bad news. On-time performance can be normative, for example, even if it means that operators unknowingly borrow from safety to achieve it. In such cases, the hurried nature of a departure or arrival is not bad news that is worth reporting (or worth listening to, for that matter). It is the norm that everyone tries to adhere to since it satisfies other important organizational goals (customer service, financial gain) without obviously compromising safety.

Outside audits are one way to help an organization break out of the perception that its safety is uncompromised. In other words, neutral observers may better be able to spot the "bad news" among what are normal, everyday decisions and actions to people on the inside.

SIGNS OF NOT LEARNING FROM FAILURE:

Most organizations aim to learn from failures, either after they have happened or before they are about to happen. The path to learning from failure is generally paved with intentions to embrace the new view of human error; to see human error as a symptom of deeper, systemic trouble. But many obstacles get in the way, frustrating attempts to learn—either after a serious failure or on the way towards one. Here are some signs of people not learning—all are ways in which organizations try to limit the need for fundamental change:

"To err is human"

Although it is a forgiving stance to take, organizations that suggest that "to err is simply human" may normalize error to the point where it is no longer interpreted as a sign of deeper trouble.

"There is one place where doctors can talk candidly about their mistakes. It is called the Morbidity and Mortality Conference, or more simply, M. & M. Surgeons, in particular, take M. & M. seriously. Here they can gather behind closed doors to review the mistakes, complications and deaths that occurred on their watch, determine responsibility, and figure out what to do differently next time."

A sophisticated instrument for trying to learn from failure, M. & M.'s assume that every doctor can make errors, yet that no doctor should—avoiding errors is largely a matter of will. This can truncate the search for deeper, error-producing conditions. In fact, "the M & M takes none of this into account. For that reason, many experts see it as a rather shabby approach to analyzing error and improving performance in medicine. It is isn't enough to ask what a clinician could or should have done differently so that he and others may learn for next time. The doctor is often only the final actor in a chain of events that set him or her up to fail. Error experts, therefore, believe that it's the process, not the individuals in it, which requires closer examination and correction."[1]

"Setting examples"

Organizations that believe they have to "set an example" by punishing or reprimanding individual operators are not learning from failure. The illusion is there, of course: if error carries repercussions for individuals, then others will learn to be more careful too.

The problem is that instead of making people avoid errors, an organization will make people avoid the reporting of errors, or the reporting of conditions that may produce such errors.

In one organization it is not unusual for new operators to violate operating procedures as a sort of "initiation rite" when they get qualified for work on a new machine. By this they show veteran operators that they can handle the new machine just as well. To be sure, not all new operators take part, but many do. In fact, it is difficult to be sure how many take part. Occasionally, news of the violations reaches management, however. They respond by punishing the individual violators (typically demoting them), thus "setting examples".

The problem is that instead of mitigating the risky initiation practice, these organizational responses entrench it. The pressure on new operators is now not only to violate rules, but to make sure that they aren't caught doing it—making the initiation rite even more of a thrill for everyone. The message to operators is: don't get caught violating the rules. And if you do get caught, you deserve to be punished—not because you violate the rules, but because you were dumb enough to get caught.

A proposal was launched to make a few operators—who got caught violating rules even more than usual—into teachers for new operators. These teachers would be able to

[1] Gawande, A. (1999). When doctors make mistakes. *The New Yorker*, February 1, pages 40-55.

tell from their own experience about the pressures and risks of the practice and getting qualified. Management, however, voted down the proposal because all operators expected punishment of the perpetrators. "Promoting" them to teachers was thought to send entirely the wrong message: it would show that management condoned the practice.

Compartmentalization

One way to deal with information that threatens basic beliefs and assumptions about the safety of the system is to compartmentalize it; to contain it.

In the organization described above, the "initiation rite" takes place when new operators are qualifying for working on a new machine. So, nominally, it happens under the auspices of the training department. When other departments hear about the practice, all they do is turn their heads and declare that it is a "training problem". A problem, in other words, of which they have no part and from which they have nothing to learn.

The problem is that compartmentalization limits the reach of safety information. The assumption beneath compartmentalization is that the need to change—if there is a need at all—is an isolated one: it is someone else's problem. There is no larger lesson to be learned (about culture, for example) through which the entire organization may see the need for change. In the example above, were not all operators—also all operators outside the training department—once new operators, and thus maybe exposed to or affected by the pressures that the initiation rite represents?

What seems to characterize high reliability organizations (ones that invest heavily in learning from failure) more than anything is the ability to identify commonalities across incidents. Instead of departments distancing themselves from problems that occur at other times or places and focusing on the differences and unique features (real or imagined), they seek similarities that contain lessons for all to learn.

OBSTACLES TO LEARNING

We may be able to recognize the signs of organizations not learning from failure. But why don't they? Apart from the reasons already mentioned in chapter one (resource constraints, reactions to failure, hindsight bias, limited human factors

knowledge), there are more institutionalized obstacles to learning from failure.

Management were operators themselves

What characterizes many safety-critical organizations is that senior managers were often operators themselves—or still are (part-time). For example, in hospitals, physicians run departments, in airlines pilots do. On the one hand this provides an opportunity. Managers can identify with operators in terms of the pressures and dilemmas that exist in their jobs, thus making it easier for them to get access to the underlying sources of error.

But it can backfire too. The fact that managers were once operators themselves may rob them of credibility when it comes to proposing fundamental changes that affect everyone.

The organization in the examples above is one where senior management is made up of operators or ex-operators. What if management would want to reduce the risks associated with the initiation practice, or eliminate it altogether? They were once new operators themselves and very likely did the same thing when getting qualified. It is difficult for them to attain credibility in any proposal to curb the practice.

Over-zealous safety management

Sometimes the formal process of investigating mishaps and coming up with recommendations for change may itself stand in the way of learning from failure. In the aftermath of failure, the pressure to come up with findings and recommendations quickly can be enormous—depending on the visibility of the industry. An intense concern for safety (or showing such concern) can translate into pressure to reach closure quickly, something that can lead to a superficial study of the mishap and its deeper sources.

Also, concern for safety in a company or across an industry can promote the creation of safety departments and safety specialists. There have been cases where safety professionals have become divorced from daily operations to an extent where they only have a highly idealized view of the actual work processes and are no longer able to identify with the point of view of people who actually do the safety critical work every day.

Statistics and the 70% myth

One thing that incident reporting systems create is the illusion of statistical rationality. Across industries, cases of human error are counted and tabulated, categorized and put together. The assumption is that all "erratic" human behavior

is of the same sort or same origin. The idea is that human errors are homogenous. The assertion that at least 70% of mishaps are due to human error is particularly stable, and consistent across industries. It gives the bad news about system safety both a concrete source and a number. The persistently documented "human error problem" sponsors the false idea that the dominant safety threat today is one of human unreliability in basically safe systems.

Tabulation of errors may have worked once upon a time, when tightly controlled laboratory studies were set up to investigate human performance. In these lab studies, human tasks and opportunities to err were shrunk to a bare minimum, and singular, measurable errors could be counted as a basic unit of human performance. This kind of experimentation left the scientist with spartan but quantifiable results. Yet when it comes to human error "in the wild"—to human error as it occurs in natural complex settings—such tabulation and percentages obscure many things and muffles learning from failure:

- They ignore the fact that complex interactions between human and various other contributions are typically necessary to move a system towards breakdown today. These 70% human errors do not occur as erratic slips or brain bloopers in the vacuum of a perfectly engineered or rationally organized world. In real tales of failure, the actions and assessments we call "errors" are intermixed with breakdowns of many other kinds: mechanical, organizational. The bad news lies not in the 70% human errors, but in the interactions between human behavior and features and vulnerabilities of their operating worlds.
- Percentages hide the wide diversity of human error in the wild. As symptoms of deeper problems, the expression of human error is context-dependent. The kind of error is determined in large part by features of the circumstances in which it takes place. The details of why tasks and tools and working environments are vulnerable to errors—or why they may even invite a large percentage of errors in the first place—get lost under the large label of "human error".

Litigation

It is becoming increasingly normal—and very worrying to large segments of the safety community—that operators involved in mishaps get sued or charged with (criminal) offenses.

Valujet flight 592 crashed after take-off from Miami airport because oxygen generators in its cargo hold caught fire. The generators had been loaded onto the airplane by employees of a maintenance contractor, who were subsequently prosecuted. The editor of Aviation Week and Space Technology "strongly believed the failure of SabreTech employees to put caps on oxygen generators constituted willful negligence that led to the killing of 110

passengers and crew. Prosecutors were right to bring charges. There has to be some fear that not doing one's job correctly could lead to prosecution."[1]

But prosecution of individuals misses the point. It shortcuts the need to learn fundamental lessons, if it acknowledges that fundamental lessons are there to be learned in the first place. In the SabreTech case, the lowly maintenance employees inhabited a world of boss-men and sudden firings, stumbled through an operation that did not supply safety caps for expired oxygen generators and in which the airline was as inexperienced and under as much financial pressure as people in the maintenance organization supporting it. It was also a world of language difficulties—not only because many were Spanish speakers in an environment of English engineering language:

"Here is what really happened. Nearly 600 people logged work time against the three Valujet airplanes in SabreTech's Miami hangar; of them 72 workers logged 910 hours across several weeks against the job of replacing the "expired" oxygen generators—those at the end of their approved lives. According to the supplied Valujet work card 0069, the second step of the seven-step process was: 'If the generator has not been expended install shipping cap on the firing pin.'

This required a gang of hard-pressed mechanics to draw a distinction between canisters that were 'expired', meaning the ones they were removing, and canisters that were not 'expended', meaning the same ones, loaded and ready to fire, on which they were no expected to put nonexistent caps. Also involved were canisters which were expired and expended, and other which were not expered but were expended. And then, of course, there was the simpler thing—a set of new replacement canisters, which were both unexpended and unexpired."[2]

These were conditions that existed long before the Valujet accident, and that exist in many places today. Fear of prosecution stifles the flow of information about such conditions. And information is the prime asset that makes a safety culture work. A flow of information earlier could in fact have told the bad news. It could have revealed these features of people's tasks and tools; these long-standing vulnerabilities that form the stuff that accidents are made of. It would have shown how human error is inextricably connected to how the work is done, with what resources, and under what circumstances and pressures.

---

1 North, D. M. (2000). Let judicial system run its course in crash cases. *Aviation Week and Space Technology, May 15*, page 66.

2 Langewiesche, W. (1998). *Inside the sky*. New York: Random House, page 228.