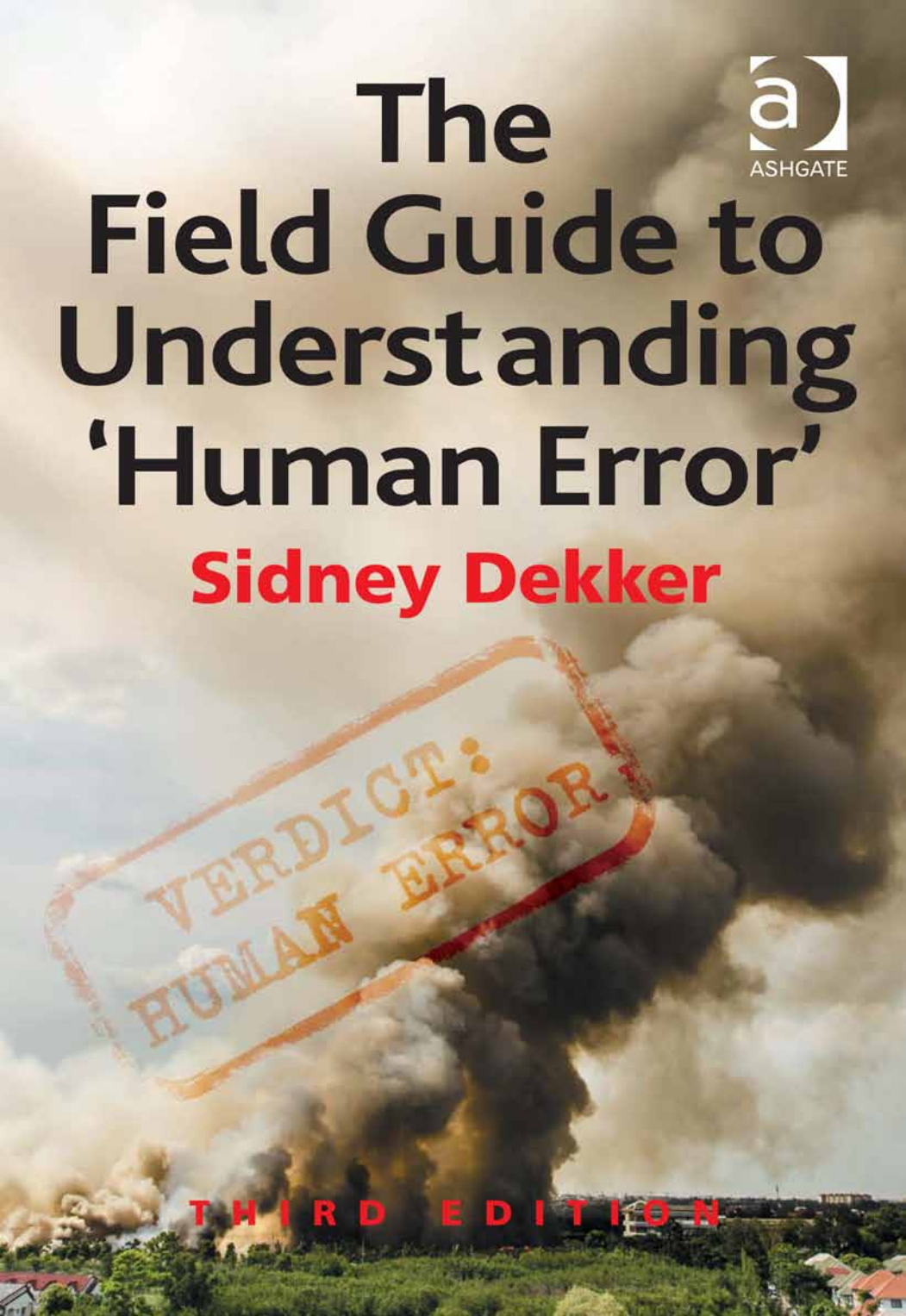
The Field Guide to Understanding 'Human Error'

Sidney Dekker



THIRD EDITION

ASHGATE

THE FIELD GUIDE TO UNDERSTANDING 'HUMAN ERROR'

It's in the world's best interest to read Dekker's book. The Field Guide *is nothing short of a paradigm shift in thinking about 'human error', and in my domain of software and Internet engineering, it should be considered required reading. This Third Edition is much better than the Second, and the layout of the material is far more accessible.*

John Allspaw, SVP, Infrastructure and Operations, Etsy

If you design equipment or operating procedures, if you investigate accidents or deal with safety, this is an essential book. Sidney Dekker, a leading world authority on 'human error' has enhanced his already exceptional "Field Guide" to be a concise, readable guide to both design of equipment and procedures and also the analysis of mishaps. The label 'human error' is misleading and its use prevents discovery and correction of the true underlying causes of incidents. So read about hindsight bias, about the difference between the view from inside the system rather than from outside, and about difference between the blunt end (where you should do your work) and the sharp end (where people tend to focus). Read, learn, and put these ideas into practice. The results will be fewer incidents, less damage, less injury.

Don Norman, author of *The Design of Everyday Things*

The Third Edition of Sidney Dekker's Field Guide to Understanding 'Human Error' *provides a solid practical framework for anyone wanting to make sense of safety science, human factors analysis, and the New View approach to conducting investigations of incidents and accidents. The trademark direct and passionate style that is common in Dekker's work focuses on the circumstances of frontline operators and managers working in complex systems, as well as the challenges of the safety investigator. Dekker does not mince his words ("Any human factors investigation that does not take goal conflicts seriously does not take human work seriously") and is clearly supportive both of sharp end workers, who are tasked with creating safety in the face of resource constraints in complex systems, as well as the investigators, charged with making sense of events that often seem surprising and unpredictable. Several new topics are introduced and enrich the earlier versions of* The Field Guide*—for instance the chapter on creating a safety department presents important principles for those with the courage to take on such a daunting task. This will be an invaluable resource for any organization serious about understanding and improving the safety of their operations.*

Dr Robert Robson, Principal Advisor,
Healthcare System Safety and Accountability, Inc.

When things go wrong in organisations, one thing is almost always found in the post-mortem: 'human error' (in various guises). But one only needs to scratch the surface of system failures to understand that things are not so straightforward. What seems to make sense as a causal catch-all for our everyday slips and blunders snaps when stretched; it fails to explain the context and complexity of our work and systems.

There is a better way. In this important book, Sidney Dekker conveys a practical approach for life after 'human error'. It is both humanistic and systemic; it treats people holistically and non-judgementally, while considering system conditions and dynamics in context. If you are prepared to suspend your own preconceptions and reactions to failure this book will repay you with a practical, highly readable and deeply humane approach to dealing with failure.

Steven Shorrock, European Safety Culture Program Leader, EUROCONTROL

Comments on the Second edition:

Next time I'm lecturing students, I'll be recommending The Field Guide *as required reading! Well done.*

Barry Kirwan, System Safety and Human Error, Eurocontrol, France; Co-editor of *Changing Regulation: Controlling Risks in Society and Human Factors Impacts in Air Traffic Management*

It is accessible, practical, eminently readable and will be of great use to safety practitioners whatever their background.

Health & Safety at Work, July 2007

This past year I read your book The Field Guide to Understanding Human Error *based on a recommendation of a colleague. I must admit it is one of the best book that I have read on accident prevention and safety. I have been practicing as a construction safety professional for 17 years and have struggled to accurately and completely articulate the concepts you so eloquently describe in your book. Although it draws many examples from an aviation safety standpoint, your book stands up brilliantly as a framework for understanding human error and accident prevention in any industry. Subsequently, I am using it as the text for my course "Safety in the Construction Industry" here at Columbia this fall.*

The construction industry is so very stuck in the world of the "Old View." Convincing construction management professional that removing bad apples is not the answer is a tough sell. Your book is making my job quite a bit easier. Thank you.

Ray Master, Columbia University, USA

I have every executive in the entire Department of Energy reading The Field Guide *as we speak.*

Todd Conklin, Los Alamos National Laboratory

No matter if the reader is an upper level executive in an aerospace company, a member of an accident investigation team, a safety engineer, or a university student, Sidney's Field Guide *is equally as useful. This book presents important ideas for those who regulate human factors investigation and research, making it an essential read for the academician, the research analyst, and the government regulator.*

International Journal of Applied Aviation Studies, Vol. 7, No. 2

The Field Guide to Understanding 'Human Error'

Third Edition

SIDNEY DEKKER
Griffith University, Australia

ASHGATE

Published by
Ashgate Publishing Limited
Wey Court East
Union Road
Farnham
Surrey, GU9 7PT
England

Ashgate Publishing Company
110 Cherry Street
Suite 3-1
Burlington, VT 05401-3818
USA

www.ashgate.com

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

The Library of Congress has cataloged the printed edition as follows:
Dekker, Sidney.
The field guide to understanding 'human error' / by Sidney Dekker. -- Third edition.
pages cm
Includes bibliographical references and index.
ISBN 978-1-4724-3904-8 (hardback) -- ISBN 978-1-4724-3905-5 (pbk.) -- ISBN 978-1-4724-3907-9 (epub) 1. System failures (Engineering) 2. Human engineering. 3. Industrial accidents. I. Title.
TA169.5.D45 2014
620.8--dc23

2014028900

ISBN: 978-1-4724-3904-8 (HBK
978-1-4724-3905-5 (PBK
978-1-4724-3906-2 (EBK)
978-1-4724-3907-9 (EPUB)

Contents

This page has been left blank intentionally

List of Figures

This page has been left blank intentionally

List of Tables

This page has been left blank intentionally

Acknowledgments

The Field Guide is inspired by the work of colleagues in 'error' and safety research. I have been fortunate enough to talk and work with many of them over the past decades, including David Woods, Erik Hollnagel, Richard Cook, James Reason, Nancy Leveson, Ed Hutchins, Diane Vaughan, Gene Rochlin, John Flach and Judith Orasanu. The foundations for much of what we all do in 'human error' and organizational safety today were laid by pioneers such as Alphonse Chapanis, Paul Fitts, Jens Rasmussen and Barry Turner.

I want to thank the practitioners and many readers of earlier versions of *The Field Guide* to give me the inspiration to write a new edition. It is through all your work across the world that the message of *The Field Guide* comes alive.

This page has been left blank intentionally

Preface

Welcome to the latest edition of *The Field Guide to Understanding 'Human Error.'* Like its two predecessors, *The Field Guide* helps you distinguish between the "Old View" of 'human error' and the "New View:"

- In the Old View, 'human errors' are the cause of most of your trouble. People, in this view, are a problem to control. People's *behavior* is something to control, something that you have to modify. You might believe you have to start with people's attitudes, because those influence their behavior. So you try to shape those attitudes with posters and campaigns and sanctions, which you hope will impact their behavior and reduce their errors. You might even elect to sanction some people under your 'just culture' policy (even though there is generally no evidence that any of this works).
- In the New View, the behavior which we call 'human error' is not a cause of trouble. It is the consequence, the effect, the symptom of trouble deeper inside your organization. The New View assumes that people do not come to work to do a bad job. So when there are bad outcomes, you must look beyond those people, at the conditions in which they worked at the time. You and your organization may well have helped create those conditions. Leave those conditions in place, and the same bad outcome may happen again—no matter how many sanctions you impose, posters you put up or safety attitude campaigns you launch.

"Old" versus "new" is perhaps a rather binary or simplistic way to think about such a complex problem as 'human error.' But, it might stretch the space of your thinking about that problem. It shows what is on either end of that space. And it provides you with a language, or concepts, to verbalize what is going on at those ends and in between. Don't allow yourself to be persuaded too easily by any view or perspective. You probably have a lot of experiences and stories that suggest your view may have been correct all along.

But do promise yourself this:

- Discipline yourself to think critically about your 'human error' problem. Make sure you ask all the questions that need asking. And then ask some more.

- Do not celebrate closure on your 'human error' problem too early. Don't take the easy answers to be the correct ones.
- Make sure you ask "is that so?" whenever you, or other people, seem to have reached a conclusion about 'human error.' Do so relentlessly. Never stop questioning.
- Use the contrast between the Old and New View to look at problems in your own setting and environment anew. Use it to map where other people might be in their thinking about 'human error.'
- Use the contrast to contribute actively to a more productive dialogue about 'human error'—not one where you necessarily want to be right all the time, but one that gets other people to stop, question and think.

Table P.1 The contrast between the Old View and the New View of 'human error'

Old View	New View
'Human error' is the *cause* of trouble	What we call 'human error' is a *symptom* of deeper trouble
'Human error' is a separate category of behavior, to be feared and fought	'Human error' is an attribution, a judgment that *we* make after the fact
'Human error' is the target; people's behavior is the problem we need to control	Behavior is systematically connected to features of people's tools, tasks and operating environment
'Human error' is something to declare war on. People need to practice perfection	'Human error' is information about how people have learned to cope (successfully or not) with complexities and contradictions of real work
'Human error' is a simple problem. Once all systems are in place, just get people to pay attention and comply	A 'human error' problem is at least as complex as the organization that helps create it
With tighter procedures, compliance, technology and supervision, we can reduce the 'human error' problem	With better understanding of the messy details of people's daily work, we can find ways to make it better for them
We can, and must, achieve zero errors, zero injuries, zero accidents	We can, and must, enhance the resilience of our people and organization

Promise yourself and your organization to do better on this 'human error' problem than you, and they, have been doing. A way of looking at things (a view) sometimes gets mistaken for a way of being (a reality). This leads people to say "We can't change. This is the way we are. This is the way things are." But of course you can change. And they can change. It begins with looking at things differently. It beings with talking about things differently, with using a different language, taking a different perspective.

The Maritime and Port Authority of Singapore (MPA) has concluded its investigations into a three separate collisions involving commercial vessels in and around Singapore that occurred over a two week period earlier this year, finding that human error and poor judgement were the main cause of all three.

Following the three collision incidents, each of which resulted in oil spills, the Maritime and Port Authority of Singapore (MPA) conducted investigations to determine the causes of the incidents and to see if there were any systemic issues that led to the spate of incidents. All three collisions resulted in oil being spilled, but no injuries to the crews.

The MPA found that in all three events, there was lack of situational awareness of the bridge teams, including the pilots, even though MPA's Port Operations Control Centre (POCC) had provided advisories and warnings of the traffic situation to the bridge teams. The MPA said that bridge teams also did not make use the Automatic Identification System (AIS), Automatic Radar Plotting Aid (ARPA), Radar, and Electronic Chart Display and Information System (ECDIS) to avoid the collisions.

Appropriate disciplinary actions will be taken against the members of the bridge teams, including the pilots, for contravening the relevant regulations, the MPA said. Existing systems and procedures put in place by the MPA have helped to keep the incident rates low otherwise.

In presenting their findings, the MPA organized a dialogue session with more than 150 representatives from the shipping community to update them on the on the investigation and measures to enhance the safety of navigation in the region.

"MPA places a strong emphasis on the safety of navigation and takes a serious view of any incidents in Singapore waters," said the MPA Chief Executive. "Moving forward, we will work more closely with all our industry partners

to review our safety management procedures and implement additional measures to enhance navigational safety. We will also not hesitate to take appropriate actions against those who infringe our safety regulations."[i]

What is New About the Third Edition?

This edition has not only been extended, but also re-organized and simplified from the previous one. Rather than a large number of chapters, each about their own micro-topic, you will find that this edition is driven by the broad concerns that readers of *The Field Guide* share. How, for example, can you distinguish different views on 'human error?' What do you need for a New View 'human error' investigation? What does a useful safety department look like, and how do you help create a culture of safety in your organization?

You can help change the way your organization thinks about 'human error.' This begins with asking questions, looking at things from another perspective, and using a different language.

In this new edition, you will find extended discussions about the sorts of hard questions that readers from previous editions have written and asked me about. Here is just a grab from them:

- The New View is nice. But what about the idiots in my organization?
- Can I really buy a "just culture" program from some consultant and make it work in my organization?
- How do I hold people accountable without making them fearful?
- What can I do to keep people engaged in safety when nothing risky has happened for a long time?
- What can we do to improve safety beyond managing the risk of negative events?
- What does resilience mean?

These and many more questions are covered in this third edition. You will also find guidance on what to do or where to go next in learning about the New View.

---

i Schuler, M. *Human error and poor judgment caused string of Singapore ship collisions*. gCaptain.com, 30 May 2014.

The quotation marks around 'human error'

Notice that 'human error' is in quotation marks (albeit single ones, in deference to a clean front cover design). The quotation marks should have been there all along. 'Human error,' after all, is no more than a label. It is a judgment. It is an attribution that we make, after the fact, about the behavior of other people, or about our own.

One of the first studies into 'human error,' in 1947, already put the label in quotation marks. Paul Fitts and Richard Jones, building on pioneering work by people like Alphonse Chapanis,[1] wanted to get a better understanding of how the design and location of cockpit controls influenced the kinds of errors that pilots made. Using recorded interviews and written reports, they built up a corpus of accounts of errors in using aircraft controls.[2] The question asked of pilots from the Air Materiel Command, the Air Training Command, and the Army Air Force Institute of Technology, as well as former pilots, was this: "Describe in detail an error in the operation of a cockpit control, flight control, engine control, toggle switch, selector switch, trim tab, etc. which was made by yourself or by another person whom you were watching at the time" (p. 332). Practically all Army Air Force pilots, they found, regardless of experience and skill, reported that they sometimes made errors in using cockpit controls. Here is one of those pilot reports:

> *"We made a normal takeoff in a heavily loaded F-13. After leaving the ground, I gave the signal to the copilot for gear up. The aircraft began to settle back toward the ground as the flaps were retracted instead of the gear. Elevator control was sufficient to prevent contacting the ground again and flying speed was maintained. The copilot, in a scramble to get the gear up, was unable to operate the two safety latches on the gear switch and I was forced to operate this switch to raise the gear. Everything would have been all right had the number-one engine not quit at this point. Full rudder and aileron would not maintain directional control and the airspeed and altitude would not permit any retarding of power on the opposite side. As the gear retraced, the airspeed built up sufficiently to enable me to maintain directional control and finally to gain a few feet of altitude. The combination of difficulties almost caused a crash. I believe that the error of raising the flaps instead of the gear was caused by inexperience of the copilot and the location of the two switches. Numerous instances of this error were reported by members of my squadron while overseas, although no accidents were known to be caused by it. However, it could result in very dangerous situations especially when combined with some other failure."*

'Human error' is no more than a label. It is a judgment. It is an attribution that we make, after the fact, about the behavior of other people, or about our own.

Fitts and Jones called their paper *Analysis of factors contributing to 460 "pilot-error" experiences in operating aircraft controls*. "Pilot error" was in quotation marks—denoting the researchers' suspicion of the term. This insight has since been replicated many times. The attribution of 'human error' depends on the perspective you take. What is a 'human error' to some people, is normal behavior to others, or even an investment in resilience.

Consider a study reported by Hollnagel and Amalberti in 2001.[3] The purpose was to test a new error counting tool in air traffic control. The method asked observers to count errors and categorize those errors using a taxonomy proposed by the developers.

It was tested in a field setting by pairs of psychologists and air traffic controllers who studied air traffic control work in real time. The practitioners and psychologists were trained to use the tool in the same way. But despite this common indoctrination, there were big differences between the numbers and the kinds of errors each of the two groups observed. Only a small number of errors were observed by both. Air traffic controllers typically relied on external working conditions (for example, interfaces, personnel and time available) to refer to and categorize errors. Psychologists, in contrast, preferred to locate the error somewhere in presumed quarters of the mind (working memory, for instance) or in some mental state (for example, attentional lapses).

Most importantly, air traffic controllers who actually did the work could tell the error coders that they both had it wrong. Observed 'human errors' were not errors to those "committing" them. Instead, the behavior that those others—who were not doing the work at the time—saw as errors turned out to be deliberate strategies intended to manage subtle operational problems or foreseen situations that the error counters had neither seen nor understood as such if they had. What was, at first, simplistic sight, a 'human error,' was an expression of human expertise. These were markers of resilience: where operators in the midst of things recognized, absorbed and adapted to challenges that they knew were going to pressurize them elsewhere, or erode margins, or create trouble later. None of this stuff—really critical for the creation of safety—would ever have been visible in a tool that set out to count 'human errors.'

What should you do about your 'human error' problem?

Go back to the 1947 Fitts and Jones study. This is how their paper opened: "It should be possible to eliminate a large proportion of so-called 'pilot-error' accidents by designing equipment in accordance with human requirements."[4] 'Pilot error' was again put in quotation marks, and once it had been investigated properly, the solutions to it became rather obvious. The point was not the 'pilot error.' That was just the symptom of trouble, not the cause of trouble. It was just the starting point. The remedy did not lie in telling pilots not to make errors. Rather, Fitts and Jones argued, we should change the tools, fix the environment in which we make people work, and by that we can eliminate the errors of people who deal with those tools. Skill and experience, after all, had little influence on "error" rates: getting people trained better or disciplined better would not have much impact. Rather change the environment, and you change the behavior that goes on inside of it. Note also how Fitts and Jones did not call these episodes "failures." Instead, they used the neutral term "experiences." We could all learn a lot from their insights, their understanding, their open-mindedness and their moral maturity.

I had a student who was involved as pilot union representative in the investigation of a runway accident. The accident had taken the lives of 49 people. At the time of the accident, the airport had been undergoing construction. "Lights are out all over the place," one pilot had remarked to the other as they got in the night before. Also, the ground navigation charts that the pilots had were not consistent with temporary taxiway layouts. Taxiway signage in the frequently shifting topography of the airport had become misleading and depictions of taxiways on the charts available in the cockpit were wrong. In the darkness of early morning, the two pilots started their takeoff roll from a runway that was close to their intended one. But the runway they had ended up on was too short.

A later state audit report that covers the year during which the accident occurred showed that more than half a million dollars had disappeared into questionable or unsupported expenses by airport managers. This included gym equipment, artwork that subsequently went missing, a hefty tab of a strip club, DVDs, video games and reimbursements for expenses that managers had already got paid for. When you are rebuilding your airport, half a million dollars may not seem like much. But it might have provided a few new taxi light bulbs and signs to guide crews to the correct runway.

> *None of this ever made it into the investigation, however. Instead, investigators determined that "the probable cause of this accident was the flight crewmembers' failure to use available cues and aids to identify the airplane's location on the airport surface during taxi and their failure to cross-check and verify that the airplane was on the correct runway before takeoff. Contributing to the accident were the flight crew's nonpertinent conversation during taxi, which resulted in a loss of positional awareness."*[5]

We sometimes seem to unlearn or disinherit everything that Fitts and Jones taught us more than half a century ago:

- Put 'human error' in quotation marks, because it is merely an attribution after the fact.
- Understand that this attribution is the starting point, not the conclusion of an investigation.
- Do not use the word "failure" but rather "experience" to describe the episode where things went wrong.

Instead, today's investigations, media and others frequently come up with new labels for 'human error' (for example: "loss of positional awareness"), and they stop when they have satisfied their desire to blame the frontline operator for their "failures" to do this, that or the other thing (which, in hindsight, is so obvious).

A Focus on 'Human Error' Makes You Do All the Wrong Things

When it comes to safety, a focus on 'human error' gets you and your organization to do almost all the wrong things. You can, however, use a perception of a 'human error' problem to your advantage. If people are worried about a 'human error' problem, you can grab the opportunity to lead your organization and others toward more constructive ways of thinking about safety and risk. Here are some of the reasons for leading people away from the label 'human error:'

- The focus on 'human error' very quickly becomes a focus on humans as the cause of safety trouble, and on humans as the targets for intervention. But this has long been shown to be a limited safety endeavor, as getting rid of one person does not remove the conditions that gave rise to the trouble they got into.

- A focus on 'human error' tends to feed on itself. We find novel names by which to call those 'human errors.' As one manager said to me, most of the injuries on her plant are caused by workers' carelessness, lack of attention, non-compliance and distractions. Many investigations now say that people "lost situation awareness" rather than more blatantly blaming operator error. One popular model is organized around workers' "unsafe acts." None of this says anything new. Those are all ways of saying 'human error.' Another way in which 'human error' feeds on itself is to find other people to blame for error. Some 'human error' management tools or books, for example, will tell you to go look for deficient supervision, inadequate maintenance, fallible decision makers or incompetent managers. Fair enough—these see 'human error' at the operational end as an effect, rather than a cause of trouble. But the effect of what exactly? Of other 'human errors,' by other defective humans. Such logic is a bit imprisoning. It doesn't get you very far.
- 'Human error' requires a standard. For the attribution to make any sense at all, it requires the possibility of actions or assessments that are not, or would not have been, erroneous. That standard often becomes visible only with knowledge of outcome, in hindsight. It is the outcome that allows us to say that other ways of working would have been smarter or better. If that outcome would have been different, the assessments and actions that are now deemed erroneous would likely have remained invisible (as normal work) or might even have been deemed constructive, heroic, resilient, innovative.
- The growing legitimacy of 'human error' as a psychological research concept has driven the development of tools and techniques for its measurement and management. What these tools do, though, is simply count and tabulate attributions—how often somebody makes a negative judgment about somebody else's performance.
- 'Human error' is not, as we once thought, a separate category of human performance. Yet this is the way that psychology has thought about it for a long time, and the media and politicians, too. And investigators. Even in popular imagination, 'human error' is often a fascinating and separate category of behavior—something to be feared, and fought. As Steve Shorrock commented, "*the popularisation of the term 'human error' has provided perhaps the biggest spur. When something goes wrong, complexity is reduced to this simple, pernicious, term. 'Human error' has become a shapeshifting persona that can morph into an explanation of almost any unwanted event. It is now almost guaranteed to be found in news stories pertaining to major accidents. Interestingly, some reports specify that 'human error' was not the cause. The reverse*

implication being that 'human error' would otherwise have been the cause (for example, "Paris train crash: 'human error' not to blame," Telegraph, *13 July 2013). Since the term suffices as explanation, little or no mention of findings in psychology or human factors, including the context and conditions of performance, is required."*[6]

- This leads to nonsensical statements like "85 percent of our incidents or accidents are due to 'human error'." 'Human error' is not only seen as the cause of trouble here. Such statements assume that these 'human errors' indeed represent a countable generic category of behavior—rather than attributions that we put on observations of behavior after the fact. They assume that there is no problem adding apples and oranges: all 'human errors' are essentially the same: countable, comparable. They can add up to 85 percent. And people then assume that 'human error' can be meaningfully separated from the otherwise blameless context (organizational, engineered) in which these varieties of behavior occurred.
- Safety and risk are made and broken the whole time, throughout your organization. You are not the custodian of an otherwise safe system that you need to protect from erratic human beings. A focus on 'human error' simplifies the enormously complex story of how people everywhere help create safety. How people have learned to cope, mostly successfully, with the pressures, contradictions and complexities of real work. The interesting story, which is masked by all this focus on 'human error,' is how in the resource-constrained and goal-conflicted worlds in which people work, their activities mostly contribute to success rather than failure.

Why a Field Guide?

This is a field guide because it helps you in those places (in your field) where 'human error' is seen as an important issue. It is a guide that you can pick up, share and use for a more fruitful dialogue with your colleagues about safety and risk. It may also guide you if you are a student and plan to work in such a field, or plan to support and advise people who already do. This may be in operations, supervision, safety, management, regulation, investigation and much more.

In order to lead people away from the label 'human error,' you need to understand that label, and its attraction, first. Much about 'human error' is about understanding yourself. And about understanding your organization: how managers, regulators, staff, colleagues, look at 'human errors' around

them. *The Field Guide* gives plenty of pointers and clues for how to do this. It helps you look at and begin to question the practices and languages that your organization might have in place to deal with its 'human error' problem. For example:

- If your organization has investigations and disciplinary actions resorting under the same department, or the same group of people, then that is a pretty strong clue that the Old View reigns.
- If your organization has adopted a behavior modification program to deal with its 'human error' problem, then that is just as strong a clue.
- If the safety manager goes to conferences where she or he mainly hears about motivation, attitude, personality and individual mindfulness, that's a clue.
- If your organization fires people for safety infractions, then that is more than a clue.
- If your organization has made the HR (Human Resources) Department responsible for safety, and they treat safety investigations as if they were individual performance reviews, then that is not even a clue. It is a call for change.

If the HR department is involved in safety in your organization, then an incident investigation can quickly degenerate from learning opportunity into performance review.

All these practices tend to stop the conversation; they stop the search for deeper sources of trouble. Sure, they provide managers and others the illusion of control. But they won't deal with the 'human error' problem. Because all they are doing is suppressing its symptoms.

The safety manager of a multinational company approached me not long ago in a bit of despair. He had just received a corporate communication which said that in recent months, there had been too many near misses, injuries and serious incidents. Safety audits, work observations, near miss reports and incident reports showed that employees were not taking the time to stop and think about the task before engaging with it. Employees were not giving safety their highest priority. The most common causes of these injuries and incidents, it went on, were carelessness, lack of awareness, disregard for safety procedures and distraction. 'Human errors' in other words.

The corporate communication announced that the company expected managers and supervisors to get their employees to be focused and vigilant,

and to get them to follow procedures in detail. It also said that there would be a renewed emphasis on accountability (whatever that meant…).

The week before, the operational arm of the same company had encouraged employees to report more near misses, the safety manager said. Even minor reports were interesting, as they could point to systemic issues that needed addressing.

But then, he lamented, this new message came down from corporate…

Targeting the "human" part of the perceived 'human error' problem is still quite popular. As in: keep beating the human until the error goes away. But your 'human error' problem is in all likelihood more about the organization than it is about the human. So *The Field Guide* is only in part about the supposed 'human errors' in the performance of others. It is only in part about understanding why people's assessments and actions made sense to them at the time. Because understanding and working on your 'human error' problem is very much about understanding *your own* reactions to failure; about recognizing and influencing your own organization's tendencies to blame and simplify. This *Field Guide* helps by providing a different language for you, and for others, to understand the 'human error' problem by. Such a different language offers a new repertoire of countermeasures: less punitive, less simplistic, less short-sighted. And more sustainable.

Some believe that they need to keep beating the "human" until the 'human error' goes away.

Subsequent editions of books tend to take on weight. To counter that, this edition of the Field Guide has been trimmed down somewhat from the previous one, notwithstanding the useful additions mentioned in the beginning of the Preface. Brevity is the reader's best friend—particularly if it is to guide you in the field where you work. That said, it is also updated and extended. You will find a chapter on creating a safety culture, for instance, and one on changing the way you do investigations. This Field Guide also suggests where to take your thinking and your organization next—after you have successfully engaged with its content.

Notes

1 Roscoe, S.N. The adolescence of engineering psychology. In: Casey S.M., editor. *Volume 1, Human factors history monograph series*. Santa Monica, CA: Human Factors and Ergonomics Society, 1997: 1–9.

2 Fitts, P.M., Jones, R.E. Analysis of factors contributing to 460 "pilot error" experiences in operating aircraft controls. Dayton, OH: Aero Medical Laboratory, Air Material Command, Wright-Patterson Air Force Base, 1947.

3 Dekker, S.W.A., editor. The emperor's new clothes: Or whatever happened to 'human error'? 4th international workshop on human error, safety and systems development; 2001; Linköping, Sweden. Linköping University.

4 Fitts, *op cit*.

5 NTSB. Attempted takeoff from wrong runway Comair flight 5191, Bombardier CL-600-2B19, N431CA, Lexington, Kentucky, August 27, 2006. Springfield, VA: National Transportation Safety Board, 2007.

6 Shorrock, S.T. The use and abuse of 'human error'. In: Hummerdal D, editor. *Safety differently*. Brisbane: Daniel Hummerdal, 2013: 2–13.

This page has been left blank intentionally

1 Two Views of 'Human Error'

There are basically two ways of looking at 'human error.' The first view is known as the Old View, or 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: more than two-thirds of them.
- Failures come as unpleasant surprises. They are unexpected and do not belong in the system. Failures are introduced to the system through the inherent unreliability of people.

The Old View maintains that safety problems are the result of a few Bad Apples in an otherwise safe system. These Bad Apples don't always follow the rules, they don't always watch out carefully. They undermine the organized and engineered system that other people have put in place. This, according to some, creates safety problems:[1]

> *"It is now generally acknowledged that human frailties lie behind the majority of accidents. Although many of these have been anticipated in safety rules, prescriptive procedures and management treatises, people don't always do what they are supposed to do. Some employees have negative attitudes to safety which adversely affect their behaviors. This undermines the system of multiple defenses that an organization constructs" to prevent injury and incidents.*

This embodies all of the tenets of the Old View:

- Human frailties lie behind the majority of accidents. 'Human errors' are the dominant cause of trouble.
- Safety rules, prescriptive procedures and management treatises are supposed to control erratic human behavior.
- But this control is undercut by unreliable, unpredictable people who still don't do what they are supposed to do.

- Some Bad Apples have negative attitudes toward safety, which adversely affects their behavior. So not attending to safety is a personal problem, a motivational one, an issue of individual choice.
- The basically safe system, of multiple defenses carefully constructed by the organization, is undermined by erratic or unreliable people.

Notice also what solutions are implied here. In order to *not* have safety problems, people should do as they are told. They should be compliant with what managers and planners have figured out for them. Indeed, managers and others above them are smart—they have put in place those treatises, those prescriptive procedures, those safety rules. All the dumb operators or practitioners need to do is follow them, stick to them! How hard can that be? Apparently it can be really hard. But the reason is also clear: it is because of people's negative attitudes which adversely affect their behaviors. So more work on their attitudes (with poster campaigns and sanctions, for example) should do the trick.

This view, the Old View, is limited in its usefulness. In fact, it can be deeply counterproductive. It has been tried for decades, without noticeable effect. Safety improvement comes from abandoning the idea that errors are causes, and that people are the major threat to otherwise safe systems. Progress on safety comes from embracing the New View.

> *A Boeing 747 Jumbo Jet crashed when taking off from a runway that was under construction and being converted into a taxiway. The weather at the time was bad—a typhoon was about to hit the country: winds were high and visibility low. The runway under construction was close and parallel to the intended runway, and bore all the markings, lights and indications of a real runway. This while it had been used as a taxiway for quite a while and was going to be officially converted at midnight the next day—ironically only hours after the accident.*
>
> *Pilots had complained about potential confusion for years, saying that not indicating that the runway was not really a runway was "setting a trap for a dark and stormy night." Moreover, at the departure end there was no sign that the runway was under construction. The first barrier stood a kilometer down the runway, and behind it a mass of construction equipment—all of it hidden in mist and heavy rain. The chief of the country's aviation administration, however, claimed that "runways, signs and lights were up to international requirements" and that "it was clear that 'human error' had led to the disaster." So 'human error' was simply the cause. To him, there was no deeper trouble of which the error was a symptom.*

Bad People In Safe Systems, Or Well-Intentioned People In Imperfect Systems?

At first sight, stories of error seem so simple:

- somebody did not pay enough attention;
- if only somebody had recognized the significance of this indication, or of that piece of data, then nothing would have happened;
- somebody should have put in more effort;
- somebody thought that making a shortcut was no big deal.

So telling other people to try harder, to watch out more carefully, is thought to deal with the 'human error' problem:

> *The ministry of transport in Tokyo issued an order to all air traffic controllers to step up their vigilance after an incident that happened to a JAL flight that ended up injuring 42 people.*

Given what you know after the fact, most errors seem so preventable. It might prompt you, or your organization to do the following things:

- get rid of Bad Apples;
- put in more rules, procedures and compliance demands;
- tell people to be more vigilant (with posters, memos, slogans);
- get technology to replace unreliable people.

But does that help in the long run—or even the short run? It doesn't. In fact, these countermeasures are not just neutral (or useless, if you want to put it that way). They have additional negative consequences:

- Getting rid of Bad Apples tends to send a signal to other people to be more careful with what they do, say, report or disclose. It does not make 'human errors' go away, but does tend to make the evidence of them go away; evidence that might otherwise have been available to you and your organization so that you could learn and improve.
- Putting in more rules, procedures and compliance demands runs into the problem that there is always a gap between how work is imagined (in rules or procedures) and how work is done. Pretending that this gap does not exist is like sticking your head in the sand. And trying to force the gap to close with more compliance demands and threats of sanctions will drive real practice from view.

- Telling people to be more vigilant (with posters, memos, slogans) does nothing to remove the problem, certainly not in the medium or longer term. What it does do, is put your ignorance about the problem on full display. If all you are seen to be able to do is ask everybody else to try harder, what does that say about you? You obviously have made up your mind about what the source of the problem is (it's those operators or practitioners who don't try hard enough). Such preconceived judgments generally do not help your credibility or your standing among your people. First you should do the hard work to understand why it made sense for your people to do what they did, given the conditions in which they worked. And you need to ask what *your* role and your organization's role has been in creating those conditions.
- Getting technology to replace unreliable people is an attractive idea, and is wide-spread. But technology introduces new problems as well as new capacities. Rather than replacing human work, it changes human work. New technology may lead to new kinds of 'human errors' and new pathways to system breakdown.

So the apparent simplicity of 'human error' is misleading. Underneath every seemingly obvious, simple story of error, there is a second, deeper story. A more complex story.

> *A most colorful characterization of this comes from James Reason: "Rather than being the main instigators of an accident, operators tend to be the inheritors of system defects created by poor design, incorrect installation, faulty maintenance and bad management decisions. Their part is usually that of adding the final garnish to a lethal brew whose ingredients have already been long in the cooking."*[2]

This second story is inevitably an organizational story, a story about the system in which people work, about its management, technology, governance, administration and operation:

- Safety is never the only goal. Organizations exist to provide goods or services (and often to make money from it).
- People do their best to reconcile different goals simultaneously (for example, service or efficiency versus safety).
- A system isn't automatically safe: people actually have to create safety through practice at all levels of the organization.
- The tools or technology that people work with create error opportunities and pathways to failure.

Production expectations and pressures to be efficient influence people's trade-offs, making normal or acceptable what was previously perhaps seen as irregular or unsafe. In fact, this may include practices or things that you would never have believed your people would do. When you discover such things, be careful not to jump on them and remind your people to comply, to not make shortcuts, to always be careful and vigilant. Such reminders can sound so hollow if you haven't first looked at yourself and your organization—at the many expectations (some communicated very subtly, not written down), the resource constraints and goal conflicts that you help push into people's everyday working life. Remember that the shortcuts and adaptations people have introduced into their work often do not serve their own goals, but yours or those of your organization!

Underneath every simple, obvious story about 'human error,' there is a deeper, more complex story about the organization.

The second story, in other words, is a story of the real complexity in which people work. Not a story about the apparent simplicity. Systems are not basically safe. People have to create safety despite a system that places other (sometimes contradictory) expectations and demands on them.

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."

But this was never about a few bad individuals. Under pressure to perform daily work in a highly cumbersome context of checking, double-checking and registering the use of sensitive materials, such "negligence" 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, but a way to get the work done given all its constraints, pressures and expectations. The practice had grown over time, accommodating production pressures from which the laboratory owed its existence.

Table 1.1 Contrast between the Old View and New View of 'human error'

Old View	New View
Asks *who* is responsible for the outcome	Asks *what* is responsible for the outcome
Sees 'human error' as the *cause* of trouble	Sees 'human error' as a *symptom* of deeper trouble
'Human error' is random, unreliable behavior	'Human error' is systematically connected to features of people's tools, tasks and operating environment
'Human error' is an acceptable *conclusion* of an investigation	'Human error' is only the *starting point* for further investigation

People who work in these systems learn about the pressures and contradictions, the vulnerabilities and pathways to failure. They develop strategies to not have failures happen. But these strategies may not be completely adapted. They may be outdated. They may be thwarted by the complexity and dynamics of the situation in which they find themselves. Or vexed by their rules, or nudged by the feedback they get from their management about what "really" is important (often production and efficiency). In this way, safety is made and broken the whole time.

These insights have led to the New View of 'human error.' In this view, errors are symptoms of trouble deeper inside a system. Errors are the other side of people pursuing success in an uncertain, resource-constrained world. The Old View, or the Bad Apple Theory, sees systems as basically safe and people as the major source of trouble. The New View, in contrast, understands that systems are not basically safe. It understands that safety needs to be created through practice, by people.

People Do Not Come To Work To Do A Bad Job

The psychological basis for the New View is the "local rationality principle." This is based on a lot of research in cognitive science.[3] It says that what people do makes sense to them at the time—given their goals, attentional focus and knowledge—otherwise they wouldn't be doing it. In other words: people do not come to work to do a bad job. Pilots do not check in for a flight in order

to die. Nurses do not sign in to go kill a patient (and if they do, it takes you into the realm of sabotage, criminality, terrorism which requires different explanations and interventions—not part of this book).

The local rationality principle is important. If people did things that seem, at first, inexplicable, it is not because they were doing inexplicable things. Jens Rasmussen, one of the original thinkers behind what is now the New View, suggested that it is because we are not positioning ourselves to understand why it made sense for them to do what they did. That burden is on us. We need to put ourselves in their shoes, see the world through their eyes. That is the whole idea behind a good 'human error' investigation. We are obliged to do so, because if these actions made sense for some people, it may well make sense to others too. And then the bad outcome might repeat itself.

If it made sense for people to do what they did, then it may make sense for others as well.

The point of a New View 'human error' investigation is not to say where people went wrong (that much is easy). The point is to understand why they thought they were doing things right; why it made sense to them at the time.

Table 1.2 Contrast between the Old View and New View of 'human error'

Old View	New View
Says what people failed to do	Tries to understand why people did what they did
Says what people should have done to prevent the outcome	Asks why it made sense for people to do what they did

New View investigations have the following characteristics:

- they are not about individual practitioners;
- they open a window on problems that all practitioners may be facing. Their "errors," after all, are symptoms of systemic problems that all practitioners may be exposed to;
- they are not a performance review;
- they are not about discipline or blame;
- they are a learning opportunity. They see failures as markers in the system's everyday behavior, an opportunity to learn about organizational, technological and operational features that create error potential.

The New View does not claim that people are perfect. Of course not. Goals are not always met, assessments are made on imperfect information, attention varies, decisions are made that have undesirable outcomes. These are all among the normal, expected "human factors" that play a role in how safety is made and broken.

The New View does not claim that people are perfect. But it keeps you from judging and blaming people for not being perfect.

But the New View *does* avoid judging people for that. It wants to go beyond saying what people should have noticed or done. It seeks to explain why. Why did they focus on that particular issue, for example? Why did it make sense for them? When you see a situation as closely as you can to the way in which people saw it themselves—unfolding from the inside out, rather than the outside in—you may begin to see how they were trying to make sense of their situation. That they were trying to make the best of ambiguous circumstances, of which they did not know the outcome. Had they known the outcome, as you do now, they probably would have done exactly as you believe you would have done. But the point of a 'human error' investigation is to understand why people's assessments and actions made sense at the time, given their context, and without knowledge of outcome, not to point out what they should have done instead. That is what the image of the "tunnel" or "pipe" tries to convey (see Figure 1.1.).

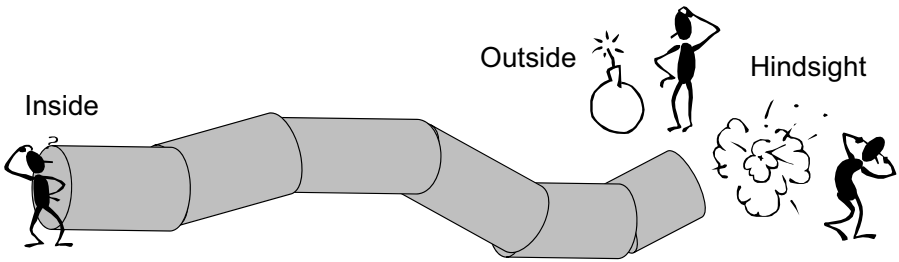


Figure 1.1 The "tunnel." Understanding 'human error' is about understanding the "inside" perspective—not the outside or hindsight one

But…What About The Idiots?

This question often comes up in conversations about 'human error' and the New View. When faced with people who do inexplicable things, stupid things,

the New View seems overly charitable. It seems too philanthropic—too loving of humanity, particularly when people do inexplicable things repeatedly and unrepentantly. "There needs to be some accountability," you might say. By which people often mean that we should be allowed to blame the truly blameworthy. To some, the New View seems to recognize only goodness and promise and possibility and human potential. How can that be maintained if there is overwhelming evidence of failure, of lack of potential, of people just not doing a good job?

Let's deal with this in three steps:

- First, we look at the claim that "Bad Apples actually do exist." Or, in other words, that some workplaces (perhaps yours) employ "idiots" who are beyond the charitable worker–rescue apologetics of the New View.
- We then look at this argument in a more neutral way, acknowledging that individual differences exist, but that accommodating those is a system issue, not an individual one.
- Finally, we take a brief look at the claim that "there has to be some accountability" to understand better what that may mean, and how such accountability might be achieved. This, indeed, is a brief look, as I have written an entire book on the subject, called *Just Culture*.[4]

"Bad apples clearly exist"

Bad apples exist. That seems to be a simple truth. And indeed a statistically supported one.

> *A study published in the British Medical Journal showed that a small number of doctors accounted for a very large number of complaints from patients. Look at the numbers and decide for yourself: 3 percent of doctors generated 49 percent of complaints, and 1 percent of doctors accounted for 25 percent of all complaints.*[5] *In addition, certain characteristics of the doctors in question, as well as past complaints, predicted future complaints. This seems to be overwhelming evidence for the existence of Bad Apples. Indeed, from this data the journal's editors conclude that some doctors are "repeat offenders" and that 'bad apples'—individuals who repeatedly display incompetent or grossly unprofessional behaviors—clearly exist." (Interestingly, the editors never say what they mean by "incompetent" or "grossly unprofessional;" this is left to the imagination of you, the reader.). Perhaps too little is done, they continued, about the doctors who practice, behave and communicate in ways that put patients at risk. If there have been repeated complaints and*

> *concerns about some doctors and not others, then that should raise flags, and we should be doing something about those individuals.*

This argument that Bad Apples exist—because the numbers about who is mostly involved in bad outcomes are there to show it—is almost a century old. The idea on which it is based, however, withered in the 1940s after having flourished for a decade or two. Let's look at that idea, and what happened to it. And then see whether the problems associated with it also make it at all possible to keep a "Bad Apple" argument alive today.

In 1913, Tolman and Kendall, two pioneers of the American safety movement, strongly recommended managers to be on the look-out for men who always hurt themselves and to take the hard decision to get rid of them. In the long run, that would be cheaper and less arduous than keeping them around the workplace. Precisely 100 years later, the *British Medical Journal* editors suggested that doctors with "recalcitrant behavior or continued legitimate complaints from patients will warrant restricted licenses or even removal from practice."[6] The idea is intuitively so appealing. The medical director of the Selby shoe company in England said in 1933 that:[7]

> *Anyone connected with an industrial clinic or hospital soon becomes acquainted with workers who make frequent visits because of accidents or illness. During the past seventeen years I have repeatedly remarked, "Here comes the fellow who is always getting hurt." At short intervals he reports with a cut, bruise, sprain or burn. Interspersed with his numerous minor accidents will be an occasional serious accident. Often you hear him remark, "I don't believe anybody has more bad luck than I have." He is the type of fellow who attracts the attention of safety engineers and safety committees. He and other such workers have been classed as "accident prone" … The accident prone workers are interesting people because they present so many problems and because the best authorities today admit they know so little about them (p. 35).*

Tolman and Kendall's recommendation, like that of the *British Medical Journal*, was based on data—at least at first sight. All the counting and tabulating of incident and injury statistics that industries had been doing in the first quarter of the twentieth century had started to yield a pattern. The public transport company in Boston, for example, found in the mid 1920s that 27 percent of its drivers on subways, street cars or buses were causing 55 percent of all accidents. Psychologically testing new employees, and eliminating those who tested badly on efficiency and power use, helped reduce their overall accident

rate—something that caught the attention of transportation operators around the world.

"Accident-prone" workers

Around 1925, independently from each other, British and German psychologists suggested that there were particularly "accident-prone" workers. The inclination to "accident" (used as a verb back then) was proportional to the number of accidents previously suffered, German psychologist Karl Marbe pointed out. Like the doctors referred to in the *British Medical Journal*, those who consistently had no accidents were significantly less likely to have them in the future as well. In England, Eric Farmer, an industrial psychologist at Cambridge, began devising tests to identify people who were likely to have accidents. Accident-proneness could be identified by carefully testing and screening employees—something for which psychological institutes across the continent, the UK and the US were developing ever cleverer simulations and contraptions. Testing workers and carefully vetting and selecting them, excluding the accident-prone ones from the jobs where they could do most harm, was thought to be a great investment in safety. It kept a growing machinery of *psychotechnik*, of testing and selection, busy well into the Second World War.

An additional suggestion by a prominent psychologist at the time was that workers should have a personal card on which accidents and errors would be noted—a trace for employers and others to base hiring and firing decisions on. This proposal was not followed concretely, but today similar employment records are used across many industries to guide managerial decision making. Like the complaints measured and recorded about individual doctors, this could turn into an index of the badness of the apple—how rotten it has become.

Dissent grew, however, toward the middle of the century. Others started to question the measurements and statistics on which the Bad Apple thesis was based. The method of percentages (such as those used by the Boston public transport system: that 27 percent of its drivers were causing 55 percent of all accidents) came under fire. It required, after all, that all drivers expected the same number of accidents (otherwise deviations from that norm could never be found). But can that ever be true? Some drivers will routinely be on routes through the busy center of the city that exposes them to vastly more collision risk. Others will spend their time mostly on quiet suburban streets. Some may drive at night a lot, others during the daytime.

Think in a similar vein about the doctors in the editorial of the *British Medical Journal*. Sweeping all doctors into one heap, and sorting the 1 percent who get

most complaints from all the others is like comparing apples and oranges. After all, what do those doctors do? What typical specialty and population does that 1 percent have? They may be involved in the kind of practice where patient harm or bad outcomes are more likely, simply because of the complicated or risky nature of what they do. A doctor practicing pediatric oncological surgery will have different outcome statistics from a general practitioner in a rural area, who will be different again from a flight surgeon attached to a squadron whose population consists essentially of carefully selected young individuals in very robust physical health. Some doctors may also have a larger share of a population which expects a particular treatment and communication style (which can then generate complaints because the doctor is not seen as "kindly" or as listening enough. That in itself has little to do with medical competence, or is really only a component of it).

Practitioners are not all exposed to the same kind and level of accident risk. This makes it impossible to compare their accident rates and say that some, because of personal characteristics, are more accident-prone than others.

Practitioners, in other words, are not exposed to the same kinds or amounts of accident risk. That depends on context (which is why *The Field Guide* invites you to carefully investigate and study context, the "tunnel" surrounding people). Many years ago, it was concluded that the evidence about "Bad Apples:"

> *…so far available does not enable one to make categorical statements in regard to accident-proneness, either one way or the other, and as long as we choose to deceive ourselves that they do, just so long will we stagnate in our abysmal ignorance of the real factors involved in the personal liability to accidents.*[8]

The Second World War was the death knell of the "accident-proneness" idea. Technological developments were so rapid, so pervasive, and so complex, that no amount of intervention toward the human worker alone could solve the emerging safety problems. Personal proneness as explanation for accident causation was taken over by other contributors. A worker's proneness to have an accident turned out to be much more a function of the tools and tasks he or she was given, and the situations they were put into, then it was the result of any personal characteristics. In addition, there was a growing concern that

a focus on accident-prone workers could get organizations and their managers off the hook. Simply blaming the worker left them scot-free, as it:

> *…may allow managements to escape their responsibilities in machine design, in selection and training of personnel in safe operating procedures, and in meticulous attention to the environment where high energy transfer can impinge on the human.*[9]

From 1925 onwards, significant organizational resources had been spent on measuring people and their performance, and tabulating the statistics, all in efforts to weed out accident-prone individuals. During and after the Second World War, this was increasingly seen as a waste. In fact, people were increasingly concerned that the focus on accident-prone people was hampering safety research; that it was getting in the way of more effective interventions. Safety conferences and professional meetings in the decades after the Second World War started to call attention more to managerial responsibilities and engineering solutions for safe work.[10] Indeed, they started asking *what* was responsible for safety and risk, not *who*.

Individual differences exist

If it is hard to prove that Bad Apples exist, isn't it still true that people are different? Of course it is. Individual differences exist. Not everybody is the same, and not everybody is suited for particular kinds of safety-critical work. That much is obvious. Selection psychology has been around for a hundred years or more—trying to understand such individual differences, map them and predict what they mean for individuals' performance across a range of tasks.

So what is the implication? When you observe a consistent, repeated "Bad Apple" in your operation (despite the difficulties with that argument, as outlined above), you may be looking at a mismatch between the task and the person. The question you need to ask yourself, and your organization, is where the responsibility for creating—and resolving—that mismatch lies:

- Is that a matter of personal responsibility for the individual involved? Is it up to them to not be mismatched to the work given to them? In certain cases you might think so. Take the doctors of the editorial in the *British Medical Journal*, for example, who are said to have communication styles and behavior patters that are ill-suited for safe patient care. It could be

argued that it is up to those doctors, as responsible professionals, to work to adjust such styles and patterns. That way, the mismatch can be reduced.

- Or is it a matter for the organization, the profession? Should *it* take responsibility for matching people to work as closely as possible? The less discretionary space people have in their work (that is, the less room they have to adjust their styles, their communication around it, their input into it, their attention to it), the more this responsibility would probably fall on the organization. It is then up to supervisors, human resources, to managers and others to help reduce the mismatch. And, of course, the mismatch says something about the organization's selection and recruitment as much as it says something about its people management. These may be things that are worth looking into as well.

A "Bad Apple" problem, to the extent that you can prove its existence, is a system problem and a system responsibility.

The New View does not avoid questions of personal competence. On the contrary. In aviation, for example, individual competence is not only considered to be hugely critical, but taken as a system responsibility. Individual competence is considered too important to leave its retaining, refreshing and checking to the individual (as in some medical specialties). Pilots are not simply entrusted by the system to stay competent and proficient on their own: the system actively helps them stay proficient and competent (not always in the right things, but there are regulatory and technological lags in all systems). Extensive systems of proficiency checking, training and retraining, as well as social monitoring and control, form some of the cornerstones of safe, competent practice in that world.

Blaming the individual for a mismatch short-circuits a number of things. It goes against the very principle of the New View, which is that "error" is not a cause of trouble but a symptom of trouble. Supposed evidence of a Bad Apple points to a host of things inside the organization or the profession. These things typically take the problem way beyond the individual Bad Apple, and when they are not addressed, they will keep driving the problem—through other situations or other people. Even the editors of the *British Medical Journal* have decided that a Bad Apple problem is a systems problem, and that addressing it is a systems responsibility: the time has come, they say, to design and evaluate *systems* that identify problematic individuals. So in healthcare, too, the recruitment and retaining of staff who turn out ineffective in their role is a systems issue—not one of defective personal accountability. A Bad Apple problem, seen this way, is not an individual problem or an individual responsibility.

"There Has To Be Some Accountability"

But, but…!

You may still protest: There has to be some accountability.

Sure. But what does that mean? The New View can be seen as an approach that seeks explanations for bad outcomes outside of individual people. It seems to "blame the system" rather than holding people accountable. I will return to this question at the end of the book—once all else has been said and done. But for now, think about this. If holding people accountable means getting them to take responsibility for their work, then the New View does not deny this at all. It is a misunderstanding to think that the New View eschews individual responsibility. Such individual responsibility, and the expectation of being held accountable for outcomes, is vital to much safety-critical work. It is often intricately connected to people's professional identity. The job would not be fun, would not be meaningful, would not be worth it, if it weren't for that responsibility and accountability. This is probably as true for doctors as it is for pilots, air traffic controllers, line men and street cleaners. This accountability forms the other side of professional autonomy and competence, to be seen to be good at what you do, and accepting the consequences when things do not go well. Such accountability gives people considerable pride, and it can make even routine operational work deeply meaningful.

But people do want to be held accountable fairly. This means not only that they want to be held accountable by those who really know the messy details of what it takes to get the job done—not by those (managers, investigators, judges) who only think they know. It is also unfair to make people responsible for things over which they have no or little authority.

The authority–responsibility mismatch

You cannot fairly ask somebody to be responsible for something he or she had no control over. It is impossible to hold somebody accountable for something over which that person had no authority.

> *"This is at the heart of the professional pilot's eternal conflict," writes Wilkinson in a comment to the November Oscar case. "Into one ear the airlines lecture, "Never break regulations. Never take a chance. Never ignore written procedures. Never compromise safety." Yet in the other they whisper, "Don't cost us time. Don't waste our money. Get your passengers to their destination—don't find reasons why you can't."*[11]

The responsibility–authority mismatch brings us back to the basic goal conflicts that drive most safety-critical and time-critical work. Such work consists of holding together a tapestry of multiple competing goals, of reconciling them as best as possible in real-time practice. Real work is full of responsibility–authority mismatches—where people have formal responsibility for the outcome of their work, but do not have full authority over the actions and decisions that take them to that outcome.

As a result, work involves efficiency–thoroughness trade-offs (ETTOs), as Erik Hollnagel calls them.[12] If an entire system is crying out for operators to be efficient, how can you then turn around after the occasional failure and all of a sudden demand that they should have been thorough all along instead? Of all the unreasonable things that we wreak upon one another in the wake of failure, says Erik Hollnagel, this is among the most unreasonable. It *does* lay down a good rule:

- holding people accountable is fine;
- but you need to be able to show that people had the authority to live up to the responsibility that you are now asking of them;
- if you can't do that, your calls for accountability have no merit, and you'd better start looking at yourself.

Valujet flight 592 crashed after takeoff 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."*[13]

But holding individuals accountable by prosecuting them 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, maintenance employees inhabited a world of boss-men and sudden firings, and that did not supply safety caps for expired oxygen generators. The airline may have been 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 now expected to put nonexistent caps. Also involved were canisters which were expired and expended, and others which were not expired but were expended. And then, of course, there was the simpler thing—a set of new replacement canisters, which were both unexpended and unexpired."[14]

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.

Accountability and the systems approach

A systems approach understands that each component or contributor in a system has specific responsibilities to help attain the system's overall goals.[15] These responsibilities cannot just be put on other people or other parts of the system. Surgeons have responsibility for the effectiveness and safety of surgery, but others have responsibility to ensure that required resources are available, all the way from junior assistants to operating room nurses, anesthetists, equipment and device manufacturers, to hospital administrators. A New Zealand surgeon, for example, was criminally prosecuted for a number of deaths to patients in his care. That he was forced to operate with help from medical students, because of a lack of available competent assistance, received scant attention.[16]

What is more, there is no evidence that a systems approach dilutes personal accountability. A growing body of research on second victims shows just *how much* responsibility practitioners take for things that go wrong, even

when they themselves can point to all the context that helped create the bad outcome.[17] They still take personal responsibility and feel guilty about it. In the cases where such second victimhood leads to grief, trauma, even suicide, there is no evidence for a lack of personal accountability. There is, however, evidence for severely underdeveloped systems of accountability. This is when we blame individuals (and, in a sense, blame the victim) without understanding the deeper context behind the event. This is a recipe for generating ever more first and second victims, and not learning much of value from the things that go wrong.

There is no evidence that a system approach dilutes personal accountability. In fact, second victims show just how much responsibility practitioners take for things that go wrong.

New models of accountability

Accountability doesn't have to be about blame and judging and pointing out mistakes and owning up to them under duress. You can think about ways to "hold people to account" without invoking all kinds of psychological and political defense mechanisms. As soon as you put people on the defensive, just imagine what happens to possibilities for learning from failure. They disappear. People will cover up, not tell you things, change or leave out inconvenient details.

The key is holding people accountable without invoking defense mechanisms.

Accountability can mean letting people tell their account, their story.

I was visiting the chief executive of a large organization to talk about safety when news came in about an incident that had just happened in their operation. A piece of heavy equipment, that should have been fastened, came loose and caused quite a bit of damage. As I was sitting there, the first reaction around the board table was "Who did this?! We must get our hands on this person and teach him a real lesson! We should turn him into an example for others! This is unacceptable!"

After the situation had calmed a bit, I suggested that if they really wanted other people to learn from this event, then it could be more profitable to talk to the person in question and ask him to write up his account of what happened and why. And then to publish this account widely throughout the

company. If the person would go along with this, then management should drop all further calls for accountability or retribution. It took some effort, but eventually they seemed to agree that this could be a more meaningful way forward.

What is the moral of this encounter?

- If you hold somebody accountable, that does not have to mean exposing that person to liability or punishment.
- You can hold people accountable by letting them tell their story, literally "giving their account."
- Storytelling is a powerful mechanism for others to learn vicariously from trouble.

Many sources point to the value of storytelling in preparing operators for complex, dynamic situations in which not everything can be anticipated. Stories contain valuable lessons about the kinds of trade-offs and sacrificing decisions that, after the fact, can be construed as controversial (people were efficient rather than thorough). Stories are easily remembered, scenario-based plots with actors, intentions, a narrative arc, and outcomes that in one way or another can be mapped onto current difficult situations and matched for possible ways out. Incident-reporting systems can capitalize on this possibility. In contrast, more incriminating forms of accountability actually retard this very quality. It robs from people the incentive to tell stories in the first place.

The move away from punishing good technicians for maintenance errors began about two decades ago as leaders began to understand the downside of disciplining to 'fix' errors—and the upside of instead conducting a thorough evaluation of the 'why' behind those errors. Even repetitive errors are usually the result of something other than a technician's negligence. A striking example of this occurred when, over a six-year period, 'hundreds of mechanics were cited for logbook violations. People working the aircraft on the gate were under pressure and they'd screw up the paperwork.' Violations meant suspensions or a fine. Then the airline wanted to print 50,000 new logbooks. Starting with the station that had most problems, it asked the mechanics to design the pages. They did. Another station made a few tweaks, and when the new logbooks were introduced, violations dropped to zero. The problem wasn't negligent mechanics, it was a poorly designed logbook.

> *As another example, airlines in the 80's and 90's were having problems with oil caps not being properly replaced on JT8D engines. Lots of good mechanics were temporarily suspended over this. Closer investigation revealed that the caps were so hot that mechanics could not properly get a hand on them to check whether they were actually sealed. A visual check was not sufficient, but supervisors did not believe mechnanics telling them about the problem. The bottom line: discipline without understanding the problem is ineffective.*[i]

To report or not to report

Depending on the safety level of the activity, a confidential reporting system is a key source of safety-related information. Confidential is not the same as anonymous. Anonymous means the reporter is known to nobody. Confidential means that the reporter is known, but only, for example, for safety people—*not* to people who could create career jeopardy (for example, supervisors, line managers). A confidential reporting scheme might help you get details about events which would otherwise elude you. Mike O'Leary and Nick Pidgeon showed how much the formal and confidential reports about the same event can diverge, and how little actionable, interesting information a formal report sent to the line manager might contain.

> *Here is an example of the difference: Formal report: "On final approach at 2,000 feet AGL the first stage of flaps was selected. Flaps failed with no flap movement. A decision was made to go around and hold while preparing the aircraft for a flapless approach. Flapless approach completed on runway XX."*

> *Part of the confidential report: "During the go-around, I was distracted by concern for proximity of high ground as the clearance was non-standard. 'Gear-up' was called. I was about to select it up when air traffic control called again. After the call, I continued the after-takeoff checks as if the gear was up. Neither of us realized it was still down for some five minutes."*

The confidential report touches on issues that never even made it into the formal one. These are issues that reveal additional exposure to risk in a situation like the one described, and how crews try to manage that risk. The effects of this mix of unexpected system failure, workload and distractions are never mentioned in the formal report, and thus would never make it into organizational consciousness.

---

i Baldwin, H. (2013). Probe, don't punish: Investigate the reasons behind errors. *Aviation Week and Space Technology*, December 31 2012–January 7 2013, p. 134.

With the confidential report in hand, organizations can think about investing in broader meaningful countermeasures (other than just fixing the flap problem on that one aircraft), such as influencing the design of go-around charts (at this particular airport and others), liaising with the airport about procedures, workload management, double-checking and crew coordination.

Accountability and the "just culture"

Many organizations struggle with how to create a "just culture." Some have bought a program off the shelf, but often find themselves without much justice, nor with meaningful accountability. Justice is hard. It is supposed to be. There are no silver bullets. But there are some directions forward.

The challenge is to create a culture of accountability that encourages learning. Every step toward accountability that your organization takes should serve that goal. Every step that doesn't serve that goal should be avoided.

Some organizations implement "just culture" as a program, based on categories of behavior:

- if worker behavior is judged to be white ("honest mistake"), s/he can stay;
- if it is grey ("risk-taking behavior"), s/he gets warned;
- if it is black ("intentional recklessness"), s/he gets sanctioned or fired.

Some organizations use culpability decision trees to help them decide. It seems so simple. It even seems fair. But who gets to say what is white, grey and black in your organization? Who has the power to draw that line?

Research on, and experience with these schemes show a couple of things:[18]

- People with more power in the organization tend to see the culture as more "just."
- Managers and supervisors can sometimes see a "just culture" program as a good way to get someone fired.
- There is little evidence that organizations which have implemented such schemes produce more learning. In fact, they sometimes encourage a climate of risk secrecy.
- There is also little evidence that practitioners avoid personal accountability, even when your organization has embraced a systems view. In fact, remember the findings on second victims mentioned above, which show just how much responsibility people take for things that go wrong on their watch.[19]

- Deciding whether behavior is an honest mistake or more culpable involves all kinds of value judgments (for example, about standards of care in the community, prudent persons, professional duties).[20] Somebody will have to make those judgments. Who will that be in your organization? In other words, there is no pre-existing line between the categories. There are only people who draw it. Categories into which we put human and social features are infinitely negotiable. What is white to someone at some point, can easily become grey to someone else, or even black.

If you truly want to create accountability and a "just culture" in your organization, forget buying it off the shelf. It won't work, independent of how much you pay for it. You need to realize that it is going to cost you in different ways than dollars. It is going to cost you in the cognitive and moral effort you need to put in. It is going to cost you when you look in the mirror and don't like what you see. Sure, you can try to create a "just culture" program based on categories. But sooner or later you will run into all the problems describe above.

If you truly want to create accountability and a "just culture" in your organization, forget buying it off the shelf. It won't work, independent of how much you pay for it. You need to realize that it is going to cost you in different ways than dollars.

Instead, think about creating justice in your responses to incidents or failures. Begin by addressing the points below. As you do so, remember that justice can never be imposed. It can only be bargained:

1. *Don't ask who is responsible, ask what is responsible.*
   Remember from the preface that, in the 1940s, human factors engineers and psychologists started asking what is responsible for errors, not who is responsible. Human factors showed that people's actions and assessments make sense once we understand critical features of the world in which they work. People's actions are systematically connected to features of their tools and tasks. Targeting those features (the what) is an action that contains all the potential for learning, change and improvement. Therefore, the first response to an incident or accident—by peers, managers and other stakeholders—should be to ask what is responsible, not who is responsible.
2. *Link knowledge of the messy details with the creation of justice.*
   One of the more frustrating experiences by practitioners involved in an incident is that those who judge them often do not really know what their work is like. They do not know the messy details, they lack technical

knowledge, misunderstand the subtleties of what it takes to get the job done despite the organization, the rules, the multiple constraints. Whether this is a supervisor, an inspector, the police, a judge, a jury—these are rarely "juries of peers." These groups do not have the same intimate knowledge of the work they are judging, and they may also have incentives to build a story that puts the practitioner at a disadvantage. So make sure you have people involved in the aftermath of an incident who know the messy details, and who have credibility in the eyes of other practitioners.

3. *Explore the potential for restorative justice.*
   Retributive justice focuses on the errors or violations of individuals. It suggests that if the error or violation (potentially) hurt someone, then the response should hurt as well. Others in the organization might have a desire to deny systemic causes, they might even fear being implicated in creating the conditions for the incident. Restorative justice, on the other hand, suggests that if the error or violation (potentially) hurt, then the response should heal. Restorative justice acknowledges the existence of multiple stories and points of view about how things could have gone wrong (and how they normally go right). Restorative justice takes the view that people do not come to work to do a bad job. Indeed, most people are willing to work constructively after a near miss has occurred. Restorative justice fosters dialogue between the actor and the surrounding community (for example, of colleagues), rather than a break in relationships through sanction and punishment.
4. *Go from backward to forward-looking accountability.*
   Backward-looking accountability means blaming people for past events. The idea of "holding someone accountable" is used for events that have already happened. It implies some sort of sanction, removal or dismissal. It is not clear what people hope to achieve with this sort of retrospective accountability, other than perhaps instilling a sense of anxiety and focus in others (*pour encourager les autres*). But this does not work: experience shows that it only motivates others to be more careful with reporting and disclosure. If, instead, we see somebody's act as a representation of an organizational, operational, technical, educational or political issue, then accountability can become forward-looking. The question becomes: what should we do about the problem and who should be accountable for implementing those changes and assessing whether they work? Forward-looking accountability is consistent with a new type of safety thinking. People are not a problem to control, but a solution to harness. Forward-looking accountability can help people focus on the work necessary for change and improvement, and connects organizational and community expectations to such work.

5. *Put second victim support in place.*
   Second victims are practitioners who have been involved in an incident that (potentially) hurt or killed someone else (for example, passengers, bystanders) and for which they feel personally responsible. Strong social and organizational support systems for second victims (psychological first aid, debriefings, follow-up), have proven critical to contain the negative consequences (particularly post-traumatic stress in all its forms). Implementing and maintaining support systems takes resources, but it is an investment not only in worker health and retention—it is an investment in justice and safety too. Justice can come from acknowledging that the practitioner is a victim too—a second victim. For some it can be empowering to be part of an investigation process. The opportunity to recount experiences first-hand can be healing—if these are taken seriously and do not expose the second victim to potential retribution or other forms of jeopardy. Such involvement of second victims is an important organizational investment in safety and learning. The resilience of second victims and the organization are intricately intertwined, after all. The lived experience of a second victim represents a rich trove of data for how safety is made and broken at the very heart of the organization. Those accounts can be integrated in how an individual and an organization handle their risk and safety.

Your organization's journey to a "just culture" will never be finished, even if you develop a program around the five questions above. Justice, after all, is one of those categories about which even reasonable people may disagree. What is just to one is unjust to another. But by following the steps above, you can help create a climate of honesty, of care, of fairness and of a willingness to learn. If you do that, justice may just come around by itself.

Of course, your organization cares about accountability. But it needs the kind of accountability that encourages learning. Accountability based on restorative justice can do that. It involves people telling their accounts, their stories. It involves expressing remorse for what happened and suggesting ways in which repetition might be prevented and relationships restored. Restorative justice does not nullify social obligations, it doesn't get people off the hook. It is not about "cheap grace." Instead, it sees accountability and learning as involving processes of disclosure, confession, apology, repentance and forgiveness. This also means that blame-free is not accountability-free. In fact, blame means less accountability: fewer accounts, less rich accounts. And less learning for your organization.

If you let your reactions to failure, and superficial demands for "accountability" get in the way of understanding and learning from the failure, you might never learn about the failure or its messy details in the first place. So learning to understand and control your own reactions to failure is critical to making progress toward the New View. Hence the next chapter.

Notes

1 Lee, T., Harrison, K. Assessing safety culture in nuclear power stations. *Safety Science* 2000;34(1):61–97.

2 Reason, J.T. *Human error*. New York: Cambridge University Press, 1990.

3 Woods, D.D., Dekker, S.W.A., Cook, R.I., Johannesen, L.J., Sarter, N.B. *Behind human error*. Aldershot, UK: Ashgate, 2010.

4 Dekker, S.W.A. *Just culture: Balancing safety and accountability* (Second Ed.). Farnham, UK: Ashgate, 2012.

5 Shojania, K.G., Dixon-Woods, M. 'Bad apples': Time to redefine as a type of systems problem? *BMJ Quality and Safety* 2013;22(7):528–31.

6 Ibid.

7 Burnham, J.C. *Accident prone: A history of technology, psychology and misfits of the machine age*. Chicago: The University of Chicago Press, 2009.

8 Arbous, A.G., Kerrich, J.E. Accident statistics and the concept of accident-proneness. *Biometrics* 1951;7:340–432.

9 Connolly, J. Accident proneness. *British Journal of Hospital Medicine* 1981;26(5):474.

10 Burnham, *op. cit*.

11 Wilkinson, S. The November Oscar incident. *Air & Space*, 1994: March, 80–87.

12 Hollnagel, E. *The ETTO Principle: Efficiency-Thoroughness Trade-Off. Why things that go right sometimes go wrong*. Aldershot, UK: Ashgate, 2009.

13 North, D.M. Let judicial system run its course in crash cases. *Aviation Week & Space Technology* 2000;152(20):66-67.

14 Langewiesche, W. *Inside the sky: A meditation on flight* (First ed.) New York: Pantheon Books, 1998.

15 Leveson, N.G. *Engineering a safer world: Systems thinking applied to safety*. Cambridge, MA: MIT Press, 2012.

16 Skegg, P.D.G. Criminal prosecutions of negligent health professionals: The New Zealand experience. *Medical Law Review* 1998;6(2):220–46.

17 Dekker, S.W.A. *Second victim: Error, guilt, trauma and resilience*. Boca Raton, FL: CRC Press/Taylor & Francis, 2013.

18 Dekker, S.W.A, Nyce, J.M. Just culture: "Evidence", power and algorithm. *Journal of Hospital Administration* 2013;2(3):73–8.

19 Dekker, *Second victim: Error, guilt, trauma and resilience*, *op. cit.*

20 Dekker, 2012, *op. cit.*

2 Containing Your Reactions to Failure

To understand failure, you first have to understand your own reactions to failure. Reactions to failure are typically:

- **Retrospective**. They arise from your ability to look back on a sequence of events.
- **Counterfactual**. They lay out what people could or should have done to avoid the outcome that you now know about.
- **Judgmental**. They judge people for not doing what you believe they should have done, or for not paying enough attention to what you now know is important.
- **Proximal**. They focus on those people closest in time and place to (preventing) the mishap.

A Navy submarine crashed into a Japanese fishing vessel near Hawaii, sinking it and killing nine Japanese men and boys. The submarine, on a tour to show civilians its capabilities, was demonstrating an "emergency blow"—a rapid resurfacing. Time had been running short and the crew, crowded in the submarine's control room with 16 visitors, conducted a hurried periscope check to scan the ocean surface. Critical sonar equipment onboard the submarine was inoperative at the time.

The more you react to failure, the less you will understand it.

The commander's superior, an admiral, expressed shock over the accident. He was puzzled, since the waters off Hawaii are among the easiest areas in the world to navigate. According to the admiral, the commander should not have felt any pressure to return on schedule. At one of the hearings after the accident, the admiral looked at the commander in the courtroom and said, "I'd like to go over there and punch him for not taking more time." As the admiral saw it, the commander alone was to blame for the accident—civilians onboard had nothing to do with it, and neither had inoperative sonar equipment.

Reactions to failure interfere with your understanding of failure. The more you react, the less you understand. When you say "how could they have been so stupid to…!" or ask "how could they not have noticed…?" you are reacting to failure. These reactions block you from seeing how it could have made sense, and how it could make sense again to others you are responsible for. They block you from exploring the second story—the deeper, more complex organizational story behind a 'human error'.

Retrospective

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 (where they were versus where they thought they were; what state their system was in versus what they thought it was in).
- Hindsight allows you to pinpoint what people missed and shouldn't have missed; what they didn't do but should have done.

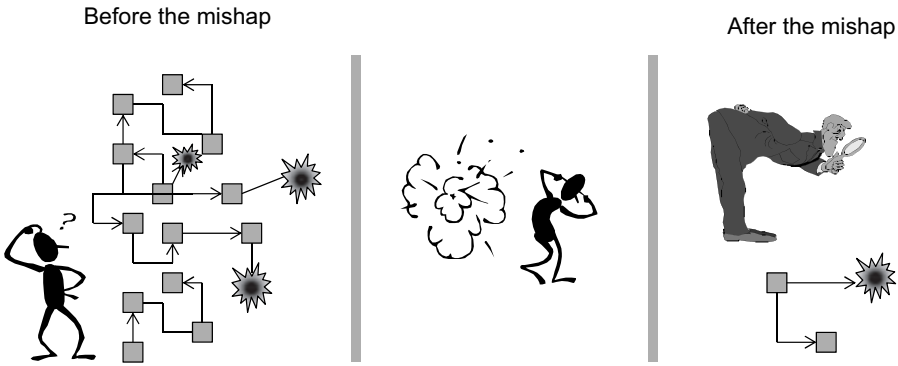


Figure 2.1 Hindsight changes how you look at past decisions. It turns past complexity into a simple, linear story; a binary decision to err or not to err (idea for image by Richard Cook)

Hindsight causes you to oversimplify history, compared to how people understood events at the time they were happening:

- You think that a sequence of events inevitably led to an outcome. You underestimate the uncertainty people faced: you easily forget how unlikely the actual outcome seemed at the time. Had you seen their situation from the inside, you'd likely understand that the outcome (that you now know about) was once a small probability; one among many other possible outcomes.
- You see a sequence of events as linear, leading nicely and uninterruptedly to the outcome you now know about. Had you seen the same situation from the inside, you would have recognized multiple possible pathways and many zigs and zags in them..
- You oversimplify causality because you reason backwards. When you can trace a sequence of events backwards (which is the opposite from how people experienced it at the time), you easily couple "effects" to preceding "causes" (and only those causes) without realizing that causal couplings are much more difficult to sort out when in the middle of things.

Hindsight gets you to oversimplify history. You will see events as simpler, more linear, and more predictable than they once were.

The Chairman of the investigation into the Clapham Junction railway accident in Britain realized this and made it explicit. He warned: "There is almost no human action or decision that cannot be made to look flawed and less sensible in the misleading light of hindsight. It is essential that the critic should keep himself constantly aware of that fact."[1]

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 misunderstood such a critical piece of information.

Hindsight biases you toward things that you now know were important (for example, "open descent mode"). As a result, you may assess people's decisions and actions mainly in the light of their failure to pick up these critical pieces of data. It artificially narrows your examination of the evidence and potentially misses alternative or wider explanations of people's behavior. The effect of knowing an outcome of a sequence of events is huge. It has an enormous impact on your ability to objectively look back on a piece of performance.

> *Consider the Greek mythological figure of Oedipus. He went about merrily, getting invited to make love to a woman named Jocasta. Which he did. It was only after the intimate encounter that he learned, from a messenger, that Jocasta was actually his mother. What do you think the difference is between Oedipus' memory of the brief affair—before and after he got the message? Knowing the outcome, he would no longer have been able to look back objectively on his behavior. In fact, he would probably go around asking himself how he could not have noticed, where he failed to double-check, what he missed, what he misjudged. He would, in other words, discover all kinds of "errors." Without hindsight, without the messenger, these "errors" would never have existed.*

If you look, for a moment, at the psychological research underlying all this, there are actually two ways in which your understanding of a past situation gets influenced:

- **The hindsight bias**. Finding out about an outcome increases the estimate we make about its likelihood. In other words, as a retrospective reviewer who knows the outcome of an event, you exaggerate your own ability to predict and prevent the outcome—while not even being aware of that bias.[2]
- **The outcome bias**. Once you know the outcome, it changes your evaluation of decisions that led up to it. If the outcome is bad, then you are not only more willing to judge the decisions, but also more likely to judge them more harshly.[3]

We typically assume that really bad consequences can only be the result of really bad causes. There are many studies demonstrating this. Faced with a bad outcome, or the potential for one, we assume that the acts leading up to it must have equally bad. This is the illusion of cause–consequence equivalence: we tend to believe in a fair world, where causes and effects are proportional.

But this is, indeed, an illusion in complex worlds. Even bad processes often lead to good outcomes. And good processes can lead to bad outcomes.

Processes may be "bad" in the retrospective sense that they departed from routines you now know to have been applicable. But this does not necessarily lead to failure. Given their variability and complexity, these worlds typically offer 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. This can actually serve as a marker of resilience, where people successfully juggle pressures for production with an ability to stay safe. The opposite goes too. Good processes, where people double-check and communicate (and even stick to procedures!), can lead to disastrous outcomes in unusual circumstances.

Bad process may still lead to good outcomes, and vice versa.

How can you avoid hindsight? Figure 2.2 shows two different perspectives on a 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 perspective from the inside of the tunnel**. This is the point of view of people in the unfolding situation. To them, the outcome was not known (or they would have done something else). They contributed to the sequence of events because of what they saw on the inside of the unfolding situation. To understand 'human error,' you have to take this perspective.

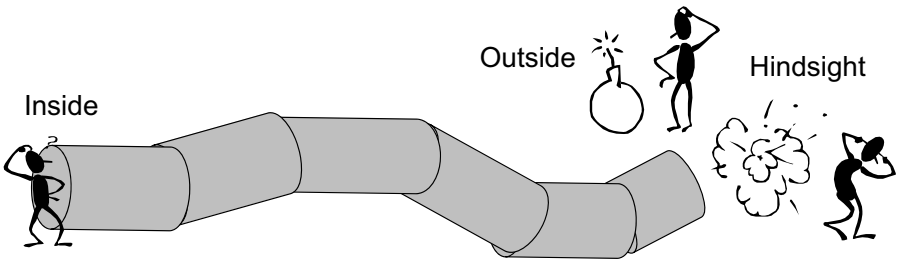


Figure 2.2 Different perspectives on a sequence of events: Looking from the outside and hindsight you have knowledge of the outcome and risks involved. From the inside, you may have neither

Further into the book, you will learn how to take this perspective, how to construct the situation as it may have looked to people inside of it.

Counterfactual

It is so easy to say what people should or shouldn't have done. You may even think that this explains the mishap; that you can understand 'human error' by saying what people should have done or did not do. But this is an illusion. When you say the following:

> "they shouldn't have…"
> "they could have…"
> "they didn't…"
> "they failed to…"
> "if only they had…!"

You are still reacting to failure. These are counterfactuals. They literally run "counter to the facts." You're talking about a reality that did *not* happen. Now you might think that this gets you closer to understanding 'human error.' It won't. Why waste time on laying out what did *not* happen? The point in understanding 'human error' is to find out why things happened the way they did. The point is not to be smart about the ways in which it could all have happened differently. In hindsight, anybody can say that.

What you believe should have happened does not explain other people's behavior. It just makes you look ignorant and arrogant.

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.*[4]

Counterfactuals say what could have happened if certain minute and often utopian conditions had been met. Counterfactual reasoning may thus be a fruitful exercise when recommending interventions against that exact failure

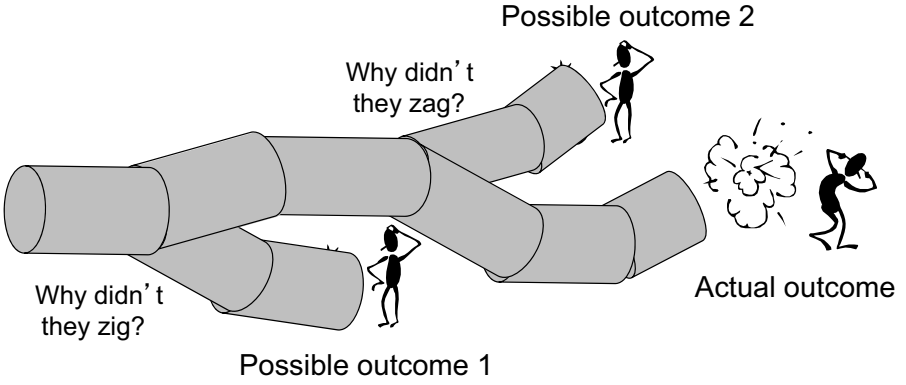


Figure. 2.3 Counterfactuals: Going back through a sequence, you wonder why people missed opportunities to direct events away from the eventual outcome. This, however, does not explain their performance

in the future. But when it comes to *explaining* behavior, counterfactuals do not contribute. Counterfactuals are not opportunities missed by the people you are investigating. Counterfactuals are just the products of your hindsight.

The effect of starting from the outcome and going backwards through a sequence of events is profound. You begin with the outcome failure. And to explain that failure, your first reflex is often to seek other failures. Where did people go wrong? What did they miss? But don't forget that you are walking backward—with knowledge of outcome, leisurely sampling the various choice moments you think people had. The forks in the road stand out so clearly to you as the various choice moments converge around you. You can stop and study them, you can take your time, and you can mentally play with the idea of doing one versus the other thing here or there. No pressure, no uncertainty. After all, you know the outcome. That's where you started. And that's why you are there.

> *The investigation into the grounding of a cruise ship found that, "had the officers regularly compared position information from the Global Positioning System and the Loran-C [another navigation device], they should not have missed the discrepant coordinates, particularly as the vessel progressed farther from its intended track."*[5] *So easy to say when going back in time from an outcome you already know about. But it explains nothing. As far as the officers were concerned, after all, the ship was ON its intended track, it was not progressing from anything at all. Our task is to understand why and how that could have been so—not smartly remarking that it wasn't.*

Indeed, imagine doing it the other way around—without knowing the outcome. You are inside the tunnel. And facing forward. And being pushed ahead by unfolding events. Time is ticking along. Now the forks are shrouded in uncertainty and the complexity of many possible options and demands; there are many more prongs and possible traps than you ever saw when you were traveling backward, while at the same time, you miss some others completely. Because all the time you are surrounded by time constraints and other pressures, and you do not have knowledge of the outcome. What people did and chose must have made sense to them given their goals, knowledge and attention at the time. Otherwise they, like you, would have done something else.

Judgmental

Explaining 'human error' often gets overridden by people getting upset about it, judging it.

> *I am looking at an* Aviation Weekly, *and its headline announces that "Computers continue to perplex pilots: Crash investigations again highlight prominence of 'human error' and mode confusion."*[6] *The writer seems upset that pilots of a Boeing 737 tried to get the autopilot engaged but did not succeed.*
>
> *After taking off in the night, the pilots of the 737 had not noticed that the autopilot had not switched on. At the same time, the captain got confused about which way the aircraft was turning. He rolled the wings of the aircraft the wrong way. Perhaps he was mixing up his old Russian artificial horizon display with the Western one he was looking at now. The first officer did not intervene. The jet crashed into the sea and everybody died.*
>
> *What happened? "The three-man crew of the 737-300 completely lost the bubble—and everybody died. The illogicality of what happened there has left everybody in disbelief." The* Weekly *found how the two pilots "tracked inexorably toward what would be a dumb-founding revelation." It found how the first officer "misunderstands" that you can't engage the autopilot if the wings aren't level and that this is dangerous if you're in a very black night over a featureless sea. But he tries anyway. The captain is flying and "evidently fails to register" what is going on, assuming "that the autopilot is engaged and has control." It does not, as the writer of the* Weekly *helpfully points out for us.*

> *With "less than a minute before impact," the captain "misreads" the attitude indicator, "confusion is evident," and an "unhelpful reassurance" from the copilot only "serves to further mislead the captain." "The pilot continues to roll the wrong way. The situation turns critical" and then becomes "irretrievable." The aircraft splashes into the sea.*

This writer stands on the sideline and scratches his head at the spiraling, escalating rush into lethal trouble that he's seen coming all along. The writer is counting down the seconds to impact, and then—

BOOM.

Told you so.

The problem about taking this position of retrospective outsider is that it does not allow you to explain anything. From that position, all you can do is judge people for not noticing what you find so important now (given your knowledge of outcome). From the position of retrospective outsider, it is possible only to condemn people for turning a manageable situation into an irretrievable one. You will never be able to make sense of the behavior of those people. Real understanding comes from putting yourself in the shoes of the people on the inside of the sequence of events; on the inside of the tunnel. Only there can you begin to see why it made sense for people to do what they did.

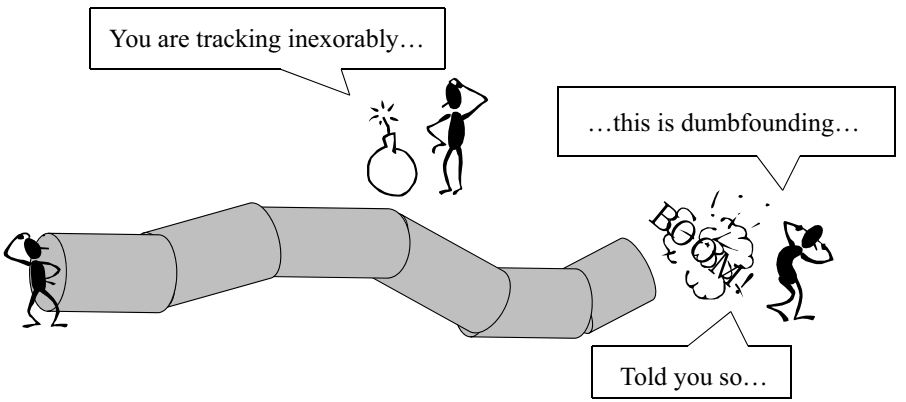


Figure. 2.4 The language you use gives away where you stand. Here you clearly take the position of retrospective, judgmental outsider

The very use of the word "failure" (for example: "the crew failed to recognize a mode change") indicates that you are still on the outside of the tunnel, looking back and looking down. The word failure implies an alternative pathway, a counterfactual one, one which the people in question did not take (for example, recognizing the mode change).

> *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.*
>
> *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 was available, and 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.*

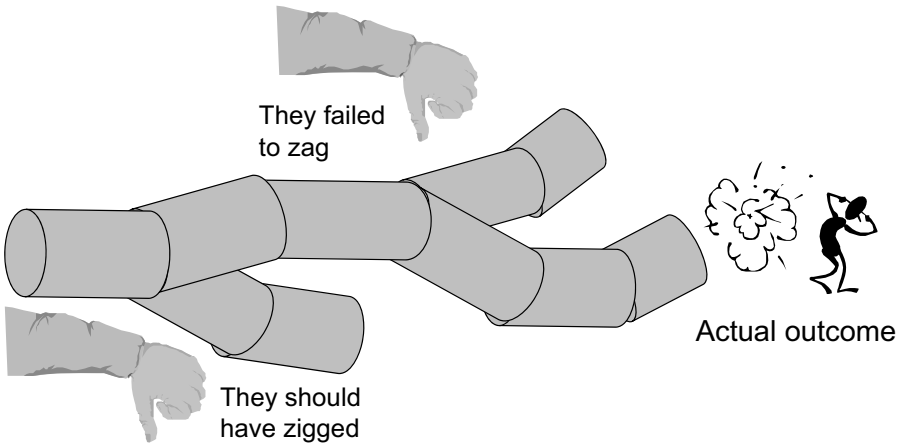


Figure 2.5 'Judgmental' means claiming that people should have done something they didn't, or failed to do something they should have. And then judging them for that. It does not explain their behavior.

The word "failure" is still popular in the probable cause statements of some investigative agencies. Take, for example, the first three probable causes of the windshear accident referred to earlier:

> *The board determines that the probable causes of the accident were: 1) the flightcrew's decision to continue an approach into severe convective activity that was conducive to a micro-burst; 2) the flightcrew's failure to recognize a windshear situation in a timely manner; 3) the flightcrew's failure to establish and maintain the proper airplane attitude and thrust setting necessary to escape the windshear.*

Saying what people failed to do has no role in understanding 'human error.' For instance, failing "to recognize a windshear situation" does not begin to explain how and why the crew interpreted their situation the way they did. And it does not help other crews avoid the same situation. Using this kind of language prevents you from understanding error. You simply keep occupying your judgmental perch, looking back down onto a sequence of events whose circumstances and outcome are now clear to you.

To understand error, take the view from the inside of the tunnel and stop saying what people failed to do or should have done.

So how can you avoid being judgmental? Consider some features of the report into the Swissair 111 accident.[7] Investigators were interested in the actions of the pilots of the large passenger jet after the crew noticed smoke in the cockpit. A diversion airport (Halifax) was in their vicinity, but they did not make an emergency descent, and never made it there. Instead, the pilots took time sizing up the situation, going through checklists, and making preparations for fuel dumping to reduce their landing weight. The developing fire caught up with them and rendered the aircraft uncontrollable. It crashed into the sea, killing everybody onboard. Now the interesting question is, what could the perspective of the crew have been? How could their actions have made sense? Let us look at the report.

> *When the pilots started their descent toward Halifax at 0115:36, they had assessed that they were faced with an air conditioning smoke anomaly that did not require an emergency descent. Based on their perception of the limited cues available, they took steps to prepare the aircraft for an expedited descent, but not an emergency descent and landing.*

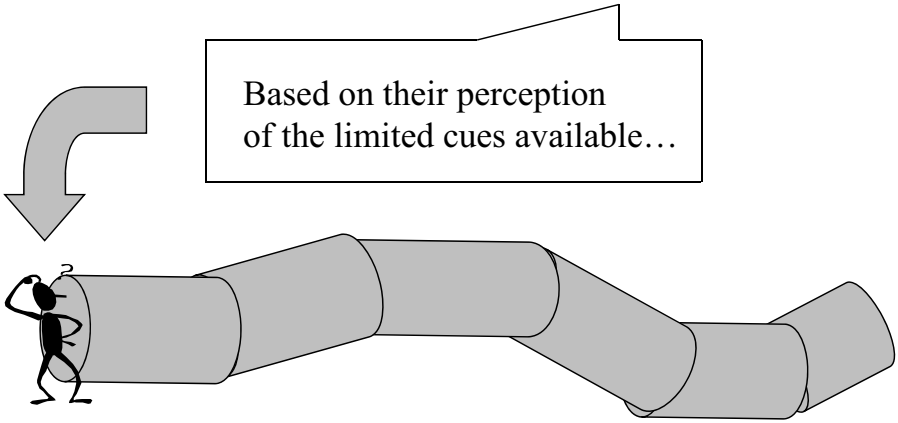


Figure 2.6 You can try to take the perspective of the people whose assessments and actions you are trying to understand

Here the report tries to speak from the crew's perspective; it acknowledges that there were only limited cues available and that the crew took action on the basis of its perception of those. It looks inside the tunnel.

Why did the pilots not rush to the diversion airport? Again, the writer tries to see the tunnel from the inside:

> *The pilots were unfamiliar with the Halifax International Airport and did not have the approach charts readily available. The back-course instrument landing approach to Runway 06 was not pre-programmed into their flight management system. The pilots knew that they would have to take additional time to familiarize themselves with, and set up for, the approach and landing. They were given the weather information by the crew of an overflying aircraft, but did not know the runway lengths or orientation. Having runway and instrument approach information available is normal practice and is important in carrying out a safe approach and landing, particularly at an unfamiliar airport at night. … The pilots also knew that the weight of the aircraft exceeded the maximum overweight landing limits for non-emergency conditions.*
>
> *In addition to these flight management circumstances, the pilots were aware that the meal service was underway, and that it would take some time to secure the cabin for a safe landing. Given the minimal threat from what they perceived to be air conditioning smoke, and the fact that there were no anomalies reported from the passenger cabin, they would likely have*

> *considered there to be a greater risk to the passengers and cabin crew if they were to conduct an emergency descent and landing without having prepared the cabin and positioned the aircraft for a stabilized approach and landing. It can be concluded that the pilots would have assessed the relative risks differently had they known that there was a fire in the aircraft.*

The report acknowledges that this crew too, would have acted differently if they had known the seriousness of the situation. The point is, they did not. How could it have made sense for the crew to attain and confirm their interpretation, and to consequently do what they did? How would it make sense to other crews too? These are critical questions, and the Swissair 111 report goes a long way in addressing them meaningfully.

Proximal

Reactions to failure focus firstly and predominantly on those people who were closest to producing or 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 from London, wanting to know how it was possible that 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.*

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 both supports and constrains activities at the sharp end (for example, the airline or hospital; equipment vendors and regulators).

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"). The blunt end shapes, creates, and sometimes even encourages opportunities for errors at the sharp end.

The Ladbroke Grove verdict of "driver error" lost credibility 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.

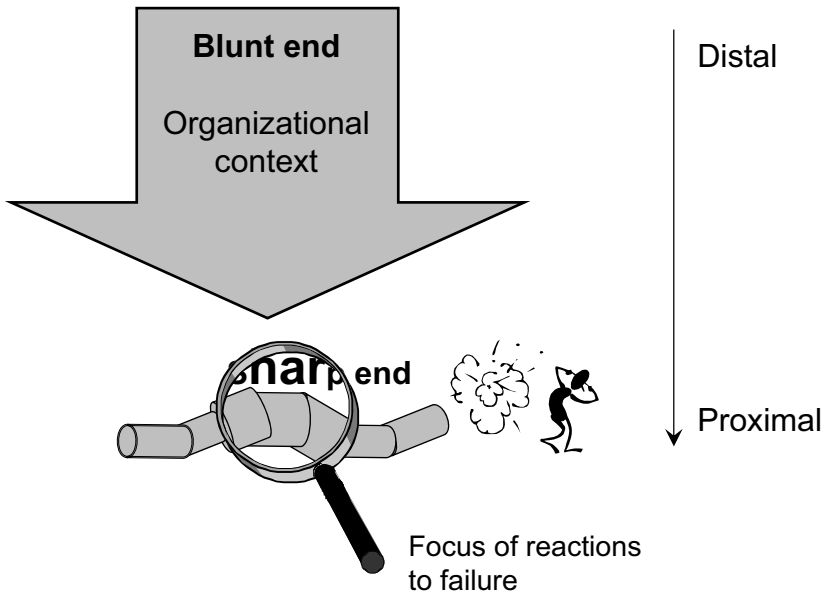


Figure 2.7 Failures can only be understood by looking at the whole system in which they took place. But in our reactions to failure, we often focus on the sharp end, where people were closest to (potentially preventing) the mishap

Faced with a bad, surprising event, your organization might want to re-tell the event, rather than change its beliefs about the system that made the event possible. Instead of modifying its views in the light of the event, it re-shapes the event until it fits a non-threatening view of itself. As far as organizational learning is concerned, the mishap might as well not have happened. The proximal nature of reactions to failure means that expensive organizational lessons can go unlearned.

> *Potential revelations about systemic vulnerabilities were deflected by pinning failure on one individual in the case of November Oscar.*[8] *November Oscar 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 he and the flight engineer were actually 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.*
>
> *November Oscar'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 dilemmas that confront crews all the time, not to mention the difficulties posed by a short approach with an old autopilot. 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—all of this was de-emphasized.*
>
> *The vilified captain was demoted to co-pilot status and ordered to pay a fine. He later committed suicide. The airline, however, had saved its public image by focusing on the proximal actions of a single individual who—the court showed—had behaved erratically and unreliably.*

Potentially disruptive (and expensive) lessons about the system as a whole can get transformed into isolated hiccups by a few uncharacteristically ill-

performing individuals. This transformation relieves the organization of any need to change views and beliefs, or associated investments and spending priorities. The system is safe, if only it weren't for a few unreliable humans in it.

> *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.*

Really understanding safety and risk not only begins with calling off the hunt for Bad Apples or their errors. It means seeing both the blunt end and sharp end and how they interact to shape practice and determine what is normal or expected throughout an organization. This means zooming out, away from looking just at the sharp end, and incorporating blunt end policies and priorities and design choices and how these help drive people's goals and practices at the sharp end.

Notes

1 Hidden, A. Clapham Junction accident investigation report. London: HMSO, 1989.

2 Fischhoff, B. Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance* 1975;1(3):288–99.

3 Baron, J., Hershey, J.C. Outcome bias in decision evaluation. *Journal of personality and social psychology* 1988;54(4):569–69. NTSB. Aircraft accident report: Flight into terrain during missed approach, USAir flight 1016, DC-9-31, N954VJ, Charlotte Douglas International Airport, Charlotte, North Carolina, 2 July 1994. Washington, DC: National Transportation Safety Board, 1995.

4 NTSB. Aircraft accident report: Flight into terrain during missed approach, USAir flight 1016, DC-9-31, N954VJ, Charlotte Douglas International Airport, Charlotte, North Carolina, 2 July1994. Washington, DC: National Transportation Safety Board, 1995.

5 NTSB. Grounding of the Panamanian passenger ship Royal Majesty on Rose and Crown shoal near Nantucket, Massachusetts, 10 June 1995. Washington, DC: National Transportation Safety Board, 1997.

6 Anon. Computers continue to perplex pilots: Crash investigations again highlight prominence of human error and mode confusion. *Air Safety Week*. Washington, DC: Flight Safety Foundation, 2005.

7 TSB. Aviation investigation report: In-flight fire leading to collision with water, Swissair Transport Limited, McDonnell Douglas MD-11 HB-IWF, Peggy's Cove, Nova Scotia 5 nm SW, 2 September 1998. Gatineau, QC: Transportation Safety Board of Canada, 2003.

8 Wilkinson, S. The November Oscar incident. *Air & Space*, 1994:March,80–87.

This page has been left blank intentionally

3 Doing a 'Human Error' Investigation

What if you are asked to do a human factors investigation for something that happened—for example, in your organization? This chapter takes you through steps that will lead to useful and verifiable results:

- getting human factors data;
- building a timeline;
- putting data in context;
- leaving a trace;
- constructing causes;
- making recommendations.

These steps will allow you to conduct an investigation that takes 'human error' out of the realm of soft data and mystery. You can gather hard data and build a set of traces that leave visible conclusions and useful recommendations.

Getting Human Factors Data

'Human error' is not just about humans. It is about how features of people's tools and tasks and working environment systematically influence human performance. So you need to gather data about all the features that are relevant to the event at hand. These might lie deeply buried in the organization surrounding people at the time—in its policies but also its unwritten preferences for on-time performance, for example. Let us look at two direct sources of human factors data that can help set you on a journey of discovery about these things:

- debriefings of participants;
- recordings of performance parameters.

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. 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. It quickly becomes impossible to separate actual events and cues that were observed from later inputs.
- Human memory tends to order and structure events more than they were; it makes events and stories more linear and plausible.

Gary Klein has spent many years refining methods of debriefing people after incidents: firefighters, pilots, nurses, and so forth. Insights from these methods are valuable to share here.[1]

The aim of a debriefing

Debriefings of mishap participants help construct the situation that surrounded people at the time and gets their view on that situation. Klein proposes the following debriefing order:

1. Have participants tell the story from their point of view, without presenting them with any replays or reminders that supposedly "refresh their memory" but would actually distort it.
2. Tell the story back to them as investigator. This is 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 a sequence of events.
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 to people and data that were actually observed by them.

At each juncture in the sequence of events (if that is how you want to structure this part of the accident story), 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?

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 colleague 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 any other sources of knowledge suggest what to do?
Goals	What were you trying to achieve? Were there multiple goals at the same time? Was there time pressure or other limitations on what you could do?
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 such a scripted set of questions, of course, as the relevance of questions depends on the event. Also, the questions can come across to participants as too conceptual to make any sense. You may need to reformulate them in the language of the domain.

Dealing with disagreements and inconsistencies in briefings

It is not uncommon that operators change their story, even if slightly, when they are debriefed on multiple occasions. Also, different participants who were caught up in the same events may come with a different take on things. How should you deal with this?

- Make the disagreements and inconsistencies, if any, explicit in your account of the event.
- If later statements from the same people contradict earlier ones, choose which version you want to rely on for your analysis and make explicit why.
- Most importantly, see disagreements and inconsistencies not as impairments of your investigation, but as additional human factors data for it. That people saw an unfolding situation differently can be crucial to your understanding of how that situation was managed. And that people might want you to rely on a particular interpretation can say a lot about the organization and its accountability processes.

Recordings, facts and analysis

Most 'human error' analysts wish they had more data. In some fields, they already have a lot (for example, cockpit voice recorders, intelligent vehicle monitoring systems, video cameras in some operating theaters). But what matters is our ability to make sense of these data. Such recordings, after all, represent only partial data traces: small, letterbox-sized windows onto assessments and actions that were part of a much richer picture. Data traces point beyond themselves, to a world that was unfolding around the people at the time, to tasks, goals, preoccupations and organizational influences.

There is not a finite amount of data that you could gather about a mishap and then think you have it all. Data keeps pointing to more questions, and more data. You will 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. 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. 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 be quite acceptable to present well-derived data as factual evidence.

Building a Timeline

Time is a powerful organizing principle, especially if you want to understand human activities in an event-driven domain. Event-driven means that the pace of activities is not (entirely) under control of the humans who operate the process. Things are happening in the process itself, or to it, that determine the tempo of, for example, people's situation assessment and decision making.

In event-driven domains, people's work ebbs and flows in sync with the processes they manage. The amount of work, the demands it poses, the pressures it puts on people—all of this varies over time. The kinds of work that people do (for example, troubleshooting, assessing, deciding, rechecking) also vary over time, in concert with demands posed by the process. If you want to begin to understand 'human error' (for example, why people seem to have missed things, or decided things that, in hindsight, do not seem to make sense) a good starting point is to build a timeline. This timeline, however, needs to be of sufficient resolution to reveal underlying processes that may be responsible for the "errors." In many attempts to understand 'human error,' timelines are of poor quality. Not because the data to build a good one weren't available, but because people did not realize the importance of a sufficiently detailed timeline for gaining insight into human performance issues.

> *I was helping with the investigation of a plane crash and was asked whether the pilots had been "rushed" in their approach to the airport. With hindsight, now knowing the landing checklist was completed below the altitude of 1,000 feet at which it "should have been," these guys must have been rushed, getting hot and high. I asked how many track miles the crew knew they had to go at various time fixes during the approach. To understand the most basic thing about workload, after all, it'd be good to plot time available against tasks they knew they still had to complete to get the jet ready for landing. The investigators came back and presented me with a little table which detailed the radial distance to the airport at various times during the approach.*

It was useless. Why? Imagine the jet on a downwind leg, passing the airfield. The radial distance to the field decreases as it nears the field, and the jet passes the field at a closest point of a couple of miles. Then it continues on, and the radial distance will increase again, up to, say, 12 miles. Then the jet might make, roughly, a 180-degree turn back onto final approach. During that turn, radial distance will essentially stay the same. Only on final approach will radial distance to the field count down in a way that is consistent with how track miles are gobbled up. My pedagogy must be lousy, as it took three attempts to persuade the investigators that radial distance was not what I needed. Granted, it was early in the investigation, and an accurate radar plot of exactly where the jet was in relation to the airport had not been recovered yet. Once I got the plot, I did the geometry myself.

Even then, it was not trivial. The track the jet ended up flying (and from which we could predict how much time there was to touch-down at any point during the approach) was not necessarily the track the crew knew they would fly and would have had represented in their flight management systems and on their moving map displays. Approach control gives corner shortcuts and leg extensions and speed restrictions to individual jets so as to merge approaching traffic as efficiently as possible. The track miles the crew believes they have to go at any moment in the approach, then, might become less, or more. It is generally hard to predict. This required us to match the clearances from air traffic control given during the approach, and see how this updated both the track and the crew's evolving understanding of how much time was left. Them being vectored onto final approach above the approach path for the runway threw an ultimate wrench into this understanding, as it suddenly and radically redefined "being rushed" or "hot and high." In all of this, it was always sobering to realize that the time between getting the approach clearance and them dying in the mud short of the runway was less than what it takes you to read this paragraph.

Remember at all times what you are trying to do. In order to understand other people's assessments and actions, you must try to attain the perspective of the people who were there at the time. Their decisions were based on what they saw on the inside of the tunnel—not on what you happen to know today.

Two issues before you go on

First, the data traces in this chapter may be impossible for you to construct, simply because you do not have access to such data (for example, voice

recordings of what operators said). However, other ways of getting time-dependent data may be available. You could have event logs, for example, that operators are required to keep. Or you may get machine read-outs of settings and processing that went on. These could allow you to infer what kinds of human activities were happening when (which things were set when, which in turn points to particular human assessments or decisions). Remember, the monitored process (for example, operation on a patient, journey of a ship, flight of an airplane, handling alarms in process control) evolves in parallel with human activities to manage that process, so one usually allows you to infer what was (or must have been) going on in the other.

While the kind of data you have access to can restrict the resolution of your analysis, that should not keep you from striving to get as much information out of your data trace as possible. For that, the sections below may still help.

The second problem is the beginning of your timeline. Despite its strengths as an analytic tool, a timeline imports all the difficulties and limitations of the sequence-of-events model (you will learn more about this when you get to the chapter on accident models). What should the beginning of a sequence of events be? There is inherent difficulty in deciding what counts as the beginning (especially the beginning—the end of a sequence of events often speaks for itself).

Beginning with the people who were closest in space and time to the eventual mishap is often the default choice. This may be fine, as long as you realize that the whole point of laying out their timeline is to understand why they did what they did. Beginning with what they did is only the first step for digging into the background that explains why. For that, you may have to go much higher or deeper, and much further back. Making clear where you start, and explaining this choice, is essential for a well-structured, credible 'human error' investigation.

Low-resolution communication timeline

Look at the following interaction between two people involved in managing an event-driven process. This could be any process and any two people managing it (for example, an anesthesiologist and her assistant discussing machine-assisted drug delivery to a patient during surgery; two pilots discussing settings during an approach into an airport; two operators of a power plant talking about intervening in their process; an officer on a ship's bridge discussing a course change with a subordinate).

The interaction is fictitious and short and deliberately generic, so that you don't get lost in the details of any particular domain at this stage. The point is

to see how human performance issues are, or are not, brought to the fore by the way you use time to structure your data. Suppose that the interaction below led up to a mishap later on, because the people involved miss-set whatever it is they had to set (drug delivery rate, minimum safe altitude, steam pressure level, ship heading). Now you need to find out how features of the interaction may have played a role in creating the conditions for that mishap. Here is the low-resolution timeline:

Time	Participant	Voice data
15:19:44	P1	We'll go down [to] one forty three
15:20:23	P1	You set it?
15:20:32	P1	Uh, you can put the steps in there too if you don't mind
15:20:36	P2	Yeah, I was planning **
15:20:37	P1	But you only need to put the steps in, ah, below the lowest safe
15:20:41	P2	Okay it's set

** = Unintelligible word

At first sight, not much seems wrong with the timeline above. The discussion between the two people is captured over time. Person 1 (P1) is talking more than person 2 (P2). P2 also seems to have a more leading role in managing the process, taking initiative and evidently directing P2 to do certain things in certain ways. P2 ends up making the setting and acknowledges as much (at 15:20:41).

Many people involved in understanding 'human error' could be content to stop here. After all, the trace here allows them to draw conclusions about, for example, the quality of the communication between P1 and P2. There is no closed loop, to mention one obvious issue: P2 says "Okay it's set," but does not close the loop on the target setting called for at 15:19:44, nor on the intermediate steps that apparently also could, or should, be set. You might even conclude that miscommunication is responsible for the mishap. But then you probably miss a bunch of other important things.

Another aspect of the low-resolution timeline is that it actually does not use time as organizing principle at all. Instead, it uses excerpts of communication as its organizing principle. The only occasions where you get a little window on the unfolding process is when somebody says something. And then this format hangs time (the starting times of those excerpts) on the side of it. Basically, this

timeline says what was said (and perhaps what was done). As to when it was said, this timeline provides a mere order of what came before what. A lot is lost when you represent your data that way.

Higher-resolution communication timeline

If you want to learn about human performance, you really have to take time as an organizing principle seriously. The reason is that phenomena such as task load, workload management, stress, fatigue, distractions, or problem escalation are essentially meaningless if it weren't for time.

Increasing the resolution of the timeline also reveals other things about, in this case, the coordination between the two persons. These things easily get lost in the low-resolution representation. Here is the same trace, at a higher resolution:

15:19:44	P1	We'll go down
15:19:45		
15:19:46		
15:19:47	P1	to one
15:19:48	P1	forty three
15:19:49		
15:19:50		
15:19:51		
15:19:52		
15:19:53		
15:19:54		
15:19:55		
15:19:56		
15:19:57		
15:19:58		
15:19:59		
15:20:00		

15:20:01		
15:20:02		
15:20:03		
15:20:04		
15:20:05		
15:20:06		
15:20:07		
15:20:08		
15:20:09		
15:20:10		
15:20:11		
15:20:12		
15:20:13		
15:20:14		
15:20:15		
15:20:16		
15:20:17		
15:20:18		
15:20:19		
15:20:20		
15:20:21		
15:20:22		
15:20:23	P1	You set it?
15:20:24		
15:20:25		
15:20:26		
15:20:27		
15:20:28		

15:20:29		
15:20:30		
15:20:31		
15:20:32	P1	Uh,
15:20:33		you can put the steps in
15:20:34		there too if you don't
15:20:35		mind
15:20:36	P2	Yeah, I was pla-
15:20:37	P1	But you only need to put the steps in
15:20:38	P2	-ning
15:20:39	P1	Ah
15:20:40	P1	Below the lowest safe
15:20:41	P2	OK it's set

From the higher-resolution timeline above, you can begin to see:

- **The ebb and flow** of work. Talk gets busier toward the end.
- **Silences**. For half a minute after the first suggestion by P1 about going down, both P1 and P2 are something that does not require speech (or that P1 and P2 *think* does not require speech). You might want to figure out what that is. Was P2 distracted or occupied with another task, as was P1? P1 may have been doing something else too, then gave P2 the assignment for a 143 setting, and then got occupied again with the other duties. So the period of silence may not mark a period of low workload (perhaps the contrary).
- **Overlapping speech**. P1 seems to be stepping on P2 (15:30:37). This can give you clues to look into hierarchical relationships, or the potential confusion presented by the device into which they need to make the entries.

Other aspects of the interaction and directed attention become visible too, and offer entry points for further examination:

- Why does P1 raise the issue again at 15:20:23? This could be surprise (why hasn't it been done yet?). P1 might believe that P2 needs coaching or monitoring in that task. As about 10 seconds go by, and P2 is (presumably)

making the setting, P1 could be supervising what P2 is doing (15:20:32) and thus not watching other things in the monitored process.
- This suggestion, in turn, might feel superfluous for P2, s/he was already "planning to" do it. Yet the coaching continues, in finer detail now. Until P2 breaks it off while P2 is still talking, by saying "Okay it's set," leaving the "it" unspecified. Rather than a communication loop not being fully closed, this excerpt from P2 could serve a message about roles and the dislike of how they are being played out. "Okay it's set" may mean "Leave me alone now."

Of course, these are all leaps of faith that need back-up with converging lines of evidence (either from talking to the people, if still possible, or their colleagues, or understanding more about the hierarchical relationships in this organization in general, the design of their devices and their training for them). You can throw even more science at this, by doing a formal conversation analysis.

Conversation analysis

Maurice Nevile, working with accident investigators, has led the way in applying a technique called conversation analysis to voice transcripts. This section is based on his work.[2] Conversation analysis uses an even finer-grained notation for systematically representing all kinds of aspects of talk (and non-talk activities). In comparison to the medium-resolution timeline, a timeline constructed using conversation analysis reveals even more about how things are said and done. It can make maximally visible how people themselves develop and understand their respective contributions to interaction and to the work required to operate their system. For example, a timeline using conversation analysis shows:

- The order of talk: how and when do participants switch roles as listener and speaker and how much of each role is played by the different participants.
- How much silence there is and when this happens.
- What overlap is there where different people talk simultaneously.
- Features of the manner of talk (stretching sounds, pitching up or down, slowing down talk or speeding it up, louder or softer talk).
- So-called tokens, such as "oh," "um" and "ah."

These features are not just interesting in themselves. They all mean something, based on elaborate theories of social interaction. They can point to the nature

of people's interaction in managing a process, and how that interaction may have contributed to, or detracted from, managing an unfolding situation safely.

This is what the exchange between our two people could look like using notation from conversation analysis:

(1.2)
P1 We'll go down (to) one (2.5) forty (.) three.
(34.3)
P1 You set it?
(8.1)
P1 Uh (0.9) you can put the (.) <steps in there too> °if you don't mi::nd°
(0.2)
P2 Yeah, I was [pla::°nning ()°]
P1 [But you only:.] need to put the steps in (0.8) ah
(1.1) below the lowest [safe,]
P2 [>Okay] it's s::et<
(1.0)

And this is what the codes and notations mean:

(1.2) pauses in seconds and tens of seconds
(.) micro-pause (shorter than two-tenths of a second)
>set< faster than surrounding talk
<set> slower than surrounding talk
set louder than surrounding talk
°set° quieter than surrounding talk
set, flat or slightly rising pitch at the end
set. falling pitch at the end
set? rising pitch at the end
se::t falling pitch within word
se::t rising pitch within word
set:. falling, then rising pitch
set:. rising, then falling pitch
[set] overlapping with other talk (also set in [ ])
() talk that could not be transcribed
(set) doubt about the talk transcribed
(set/sit) doubt about the talk/possible alternative

There are, of course, diverging opinions about what each of these things may mean for a particular utterance or the quality of people's interaction. But with

a high-resolution notation like the one above, you can begin to pick things out that the other timelines did not allow to the same extent. For instance:

- The pause in P1 suggests the setting could mean that P1 is figuring the setting out while uttering the suggestion to P2, or may even be looking it up somewhere while speaking.
- The emphasis on "set it" in P1's next utterance.
- Slowing down the "steps in there too" could point to an increasing hesitancy to interfere during that moment. Perhaps P1 is realizing that the suggestion is superfluous, as P2 is already doing so.
- An increasing hesitancy to interfere on part of P1 could be confirmed by the following quieter "if you don't mind." Also, the rising pitch in the word "mind" could make it more mildly suggestive, pleasantly inquisitive.
- The overlap between P2's increasingly silent, trailing-off "planning" and the loud start of P1's next suggestive sentence ("But you only…") as well as the rising/falling pitch in "only" could point to yet another reversion in how P1's experiences his/her role relative to P2. P2 may be seen to require more forceful, concrete guidance in making the setting.
- The subsequent overlap between the last part of P1's latest suggestion ("…the lowest safe") and P2's very quick and initially loud "Okay it's set" could mean that P2 is getting unhappy with the increasingly detailed prodding by P1. The falling pitch in "set" would confirm that P2 now wants the matter to be closed.

Taken individually, such features do not represent "errors" on the part of the people interacting here. But together, and in how they relate to each other, these features can be creating a context in which successful task performance is increasingly unlikely. This type of analysis gives you the opportunity to provide detailed, concrete evidence for a phenomenon such as "loss of effective crew resource management."

Connecting behavior and process

Now that we have one trace (the communication), we need to answer these questions:

- What was going on in the process at that time?
- What other tasks would people have plausibly been involved in simultaneously, if at all?

Understanding people's mindset begins with the unfolding situation in which the mind found itself. Here is how to go about that:

- Find out how relevant process parameters were changing over time, both as a result of human influences and of the process moving along.
- Find out how the values of these parameters were available to people—dials, displays, mode annunciations, alarms, warnings. Their availability does not mean people actually observed them: you have to make that distinction.

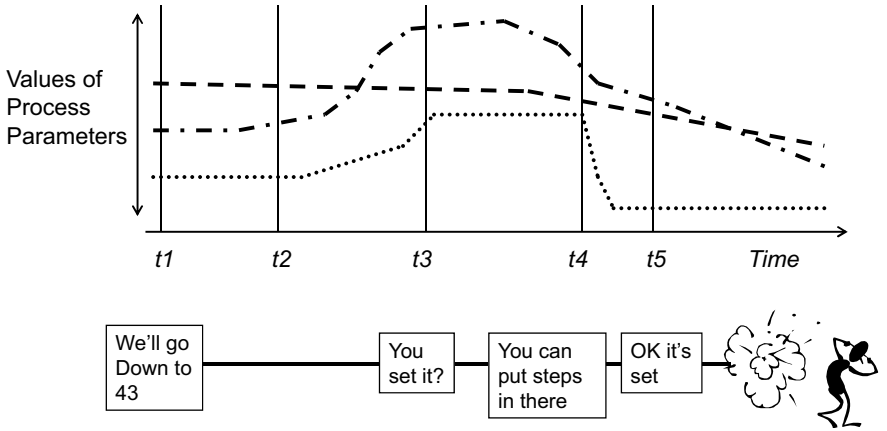


Figure 3.1 Connecting critical process parameters to the sequence of people's assessments and actions and other junctures

Indeed, you may now have laid out all relevant parameters, but what did people actually notice? What did they understand their situation to be? The answer lies in part in people's goals. People were in the situation to get a job done; to achieve particular aims. The goals that drive people's behavior at the time will tell you something about the kinds and numbers of tasks they would be trying to accomplish, and thereby the things they would, or would not, be looking for or at.

Finding what tasks people were working on need not be difficult:

- What is 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 operators.
- What was happening in the managed process? Starting from your record of parameters from the picture above, you can see how systems were set or inputs were made. These changes connect to the tasks people were carrying out.

- What were other people in the operating environment doing? People divide tasks among them in standard or otherwise predictable ways. What one operator was doing may give some hints about what the other operator was doing.

If you find that pictures speak more clearly than text, one like Figure 3.2 below gives an impression of the task load during the sequence of events.

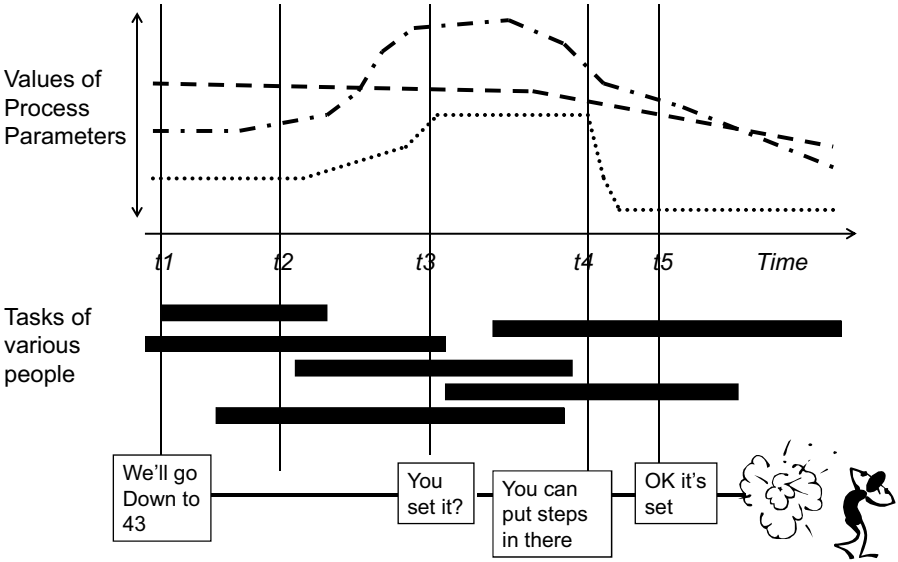


Figure 3.2 Laying out the various (overlapping) tasks that people were accomplishing during an unfolding situation

How do you identify "events" in your data?

So how do you identify the events that can serve as a thread through the presentation of your data? As shown in the timelines in the first half of this chapter, people's utterances may be strong markers for identifying events. But they may not be the only ones:

- other things may have happened in the process itself, without people triggering it, or commenting on it;
- people may have been doing things to the process without verbalizing;
- you can also infer people's interpretations from their subsequent actions in managing the process, rather than from what they say.

In a study on automation surprises, researchers were interested to find out if people noticed the loss of a particular signal in a critical phase of operations. The automation showed whether it received the signal or not, but this indication easily got hidden beneath other indications. It was easy to miss. In this case, researchers did not draw conclusions about people's awareness of the signal loss on the basis of what they said (or didn't say), but rather on the basis of their next actions. If people continued to operate the process as if nothing had happened, it was likely that they had not seen the loss of signal. Debriefings confirmed this.[3]

The "events" that were no events

Human decisions, actions and assessments can also be less obvious. For example, people seem to decide, in the face of evidence to the contrary, not to change their course of action. In hindsight, you may see missed opportunities to recover from their misunderstanding of the situation. These "decisions" to continue, these opportunities to revise, however, are events only in hindsight. The challenge for you becomes to understand how this was not an event to the people whose actions and assessments you are investigating (see, for example, *plan continuation* in Chapter 4). In other words, if you call something an "event" in your timeline, recognize that that is an analytic choice on your part—not necessarily a reality on the part of the people in the situation at the time.

Putting Data in Context

Understanding 'human error' means putting yourself in the shoes of the people whose behavior you are trying to understand.

- How did the world look to them at the time?
- How did the situation unfold around them; what cues did they get when?
- What goals were they likely pursuing at that time (not knowing the outcome you now know about)?

Remember, you want to understand why it made sense for these people to do what they did. That means that you need to put the data that you have gathered on what they did, back into their context. You need to put people's behavior back into the situation that produced and accompanied it. But this is not easy. Hindsight produces various ways in which behavioral data gets taken out of context:

- micro-matching data with a world you now know to be true and finding a mismatch;
- cherry-picking selected bits that prove a condition you have identified only in hindsight;
- presenting a shopping bag full of cues and indications that pointed to the real nature of the situation, and wondering how people could possibly have missed all that evidence.

Out of context I: Micro-matching

One of the most popular ways you can assess performance after the fact is to hold it up against a world you now know to be true. This can be:

- 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 that people's behavior falls short of.

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

Imposing procedures onto history

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 toward mountains because of a computer-database anomaly. The aircraft ended up crashing in 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.*[4]

Investigations can invest in organizational archeology to construct the regulatory or procedural framework in which operations took (or should have taken) place. In hindsight, you can easily expose inconsistencies between rules and actual behavior. 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 (Figure 3.3).

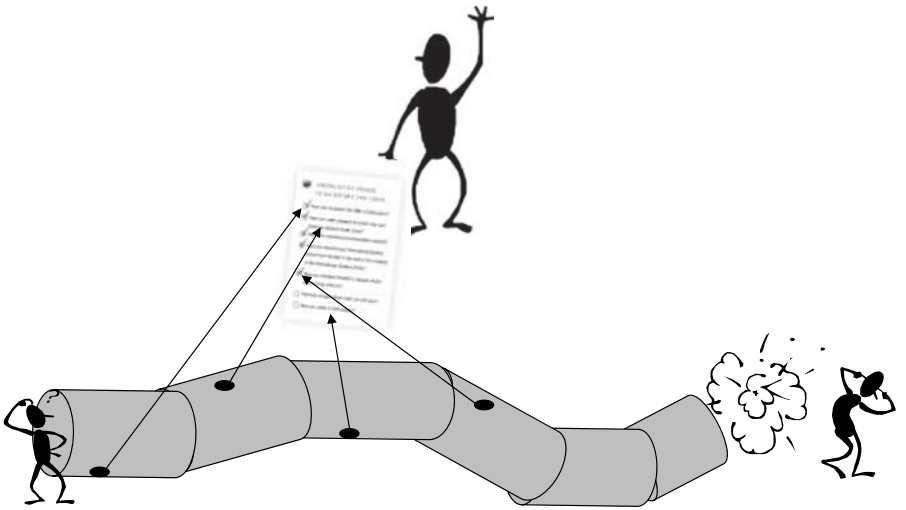


Figure 3.3 Micro-matching can mean that you take performance fragments from the stream of events and hold them up against rules or procedures that you deem applicable in hindsight. You don't explain anything by doing this

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 or the particular mishap. There is almost always a gap between how work is imagined (in rules) and how it is done (in practice)—mostly without bad consequences.

Imposing available data onto history

Another way to construct the world against which to evaluate individual performance fragments, is to turn to data in the situation that were not noticed but that, in hindsight, turned out to be critical.

> *Continue with the automated aircraft. 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."*[5]

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. There is a difference between:

- data availability: what can be shown to have been physically available somewhere in the situation;
- data observability: what would have been observable given the features of the interface and the multiple interleaving tasks, goals, interests, knowledge and even culture of the people looking at it.

The question, for understanding 'human error,' is not why people could have been so unmotivated or unwise not to pick up the things that you can decide were critical in hindsight. The question—and your job—is to find out what was important to them, and why.

Standards of good practice

Sometimes a controversial performance fragment knows no clear pre-ordained guidance but relies on local, situated judgment. Take, for example, a decision to accept a runway change, or to continue flying into adverse weather. For these cases there are always "standards of good practice," for example "good airmanship." You will always have the advantage here. You can point out that good practice was not what the people in question were doing, and hindsight helps you right along. The thing is, it does not explain the observed behavior. You have substituted your own world for the one that surrounded the people in question.

Out of context II: Cherry-picking

The second broad way in which you can take data out of context, in which you give them meaning from the outside, is by grouping and labeling behavior fragments that, in hindsight, appear to represent a common condition. For example:

> *"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 four-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 effort 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."*[6]

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."*

It is easy to pick through the evidence of a sequence and look for fragments that all seem to point to a common condition. This is called "cherry-picking"—selecting those bits that help an 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. Their similarity is entirely in the eye of you, the retrospective beholder.
- Much performance, much behavior, takes place in between the fragments that you select to build your case. These intermediary episodes contain changes and evolutions in perceptions and assessments that separate the fragments not only in time, but also in meaning.

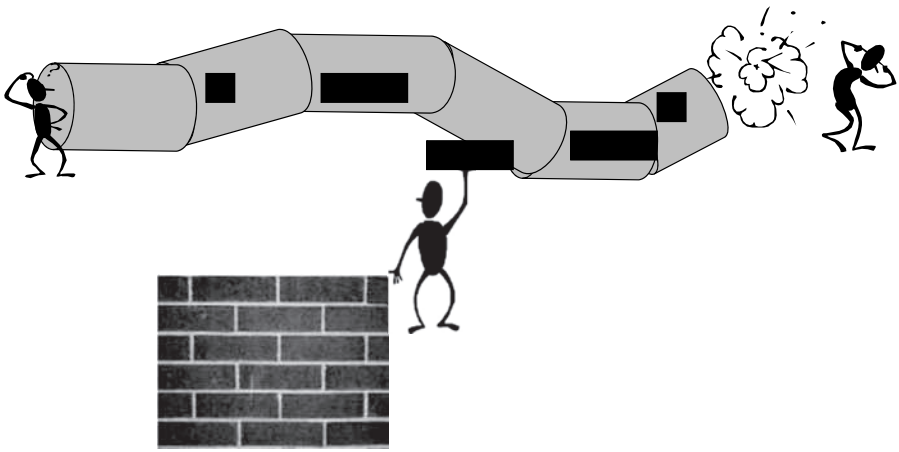


Figure 3.4 Cherry-picking means taking fragments from all over the record and constructing a story with them that exists only in your hindsight. In reality, those pieces may have had nothing to do with each other

Out of context III: The shopping bag

With the benefit of hindsight, it is easy to sweep together all the evidence that people should have seen. If they had, they would have recognized the situation for what we now know it turned out to be. But that doesn't mean the evidence presented itself that way to people at the time.

Airplanes sometimes run off the end of the runway after landing. This may happen because the landing speed is too high, or the weather is bad, with a lot of tailwind,

or the runway wet and slippery. I remember an accident where a combination of these things played a role, and I was asked to explain how a professional crew could have missed all the evidence that pointed to a deteriorating weather situation at the arrival airport. The wind had been shifting around, there was a thunderstorm moving about the area, the visibility was getting worse, and it was raining. It was as if my head was shoved into a shopping bag full of epiphanies, and I was asked, "Look, how could they not have seen all that evidence and concluded that attempting a landing was a bad idea?"

The question was misleading and driven entirely by hindsight. Only hindsight made my questioners able to chuck all the cues and indications about bad weather together in one bag. But this is not how these cues and indications revealed themselves to the pilots at the time! In fact, there was a lot of evidence, compelling, strong and early on, that the weather was going to be just fine. Cues and indications that the weather was deteriorating then came dripping in, one by one. They contradicted each other; they were not very compelling.

If you want to understand why the pilots did what they did, I replied, you have to reconstruct the situation as it unfolded around them. When did which cues come in? What did they likely mean given the context in which they appeared? What would that have meant for how the understanding of the pilots developed over time?

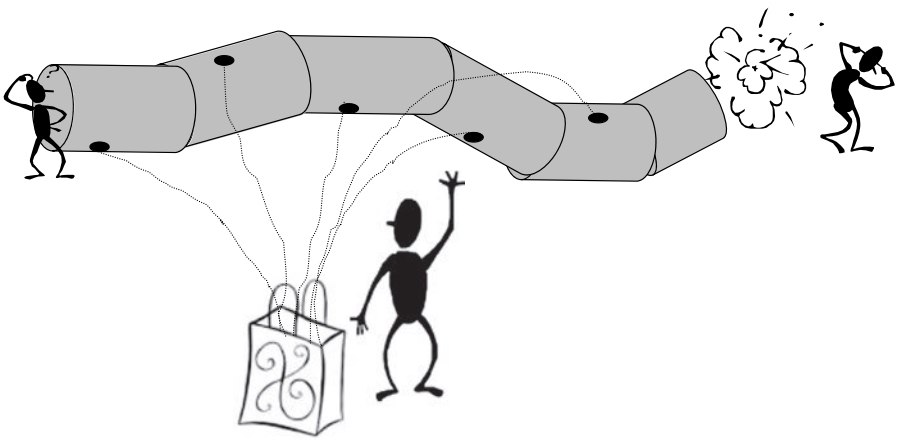


Figure 3.5 **It is easy to gather cues and indications from a sequence of events and lob them together as in a shopping bag. This is not, however, how people inside the unfolding situation saw those cues presented to them**

In order to understand 'human error,' and avoid micro-matching, cherry-picking or shopping-bagging:

- you have to put yourself in their shoes;
- imagine that you don't know the outcome;
- try to reconstruct which cues came when, which indications may have contradicted them;
- envisage what this unfolding trickle or flow of cues and indications could have meant to people, given their likely understanding of the situation at the time (remember, you are trying not to know the outcome);
- try to understand how their understanding of the situation was not static or complete, as yours in the shopping bag is, but rather incomplete, unfolding and uncertain. Your timeline should help a lot with this.

Of course, you can never be entirely sure that your recreation of their world matches how it must have looked to them. But you can take steps to avoid substituting your own outside/hindsight reality for the one that surrounded people at the time.

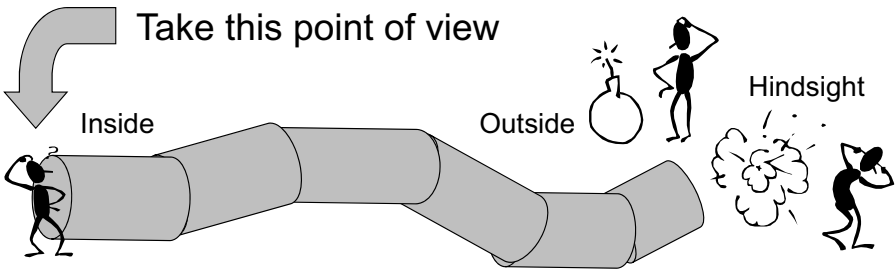


Figure 3.6 See the unfolding world from the point of view of people inside the situation—not from the outside or from hindsight

Leaving A Trace

> *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."*

Underspecified labels

'Human error' as explanation for accidents has become increasingly unsatisfying. So human factors has produced terms that try differently. But end up saying 'human error' all over again:

- "The crew lost situation awareness and effective crew resource management (CRM)" (which is why they crashed). 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.
- "It is essential in the battle against complacency that crews retain their situation awareness" (otherwise they keep missing those warning signals). 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.
- "Non-compliance with procedures is the single largest cause of 'human error' and failure" (so people should just follow the rules). Non-compliance is also a name for 'human error'—the failure to stick with standard procedures that would keep the job safe.

You end up hiding all kinds of interesting things by pasting a large label over your factual data. You can only hope it will serve as a meaningful explanation of what went wrong and why. But it won't. And if you really want to understand 'human error,' posting a large label is no substitute for doing hard analytic work.

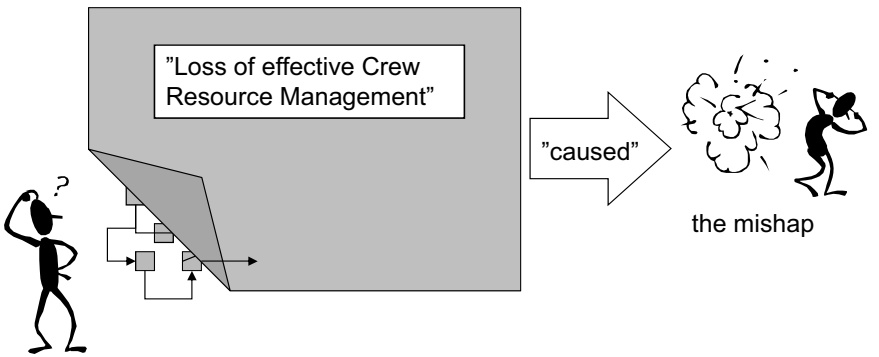


Figure 3.7 The interesting cognitive and coordinative dynamics take place beneath the large psychological label. The label itself explains nothing

Labels alone do not get you anywhere. Instead, you have to:

- specify very clearly what you mean by them. In formal terms, this is called "operationalization." Operationalization breaks up the phenomenon into smaller, measurable or observable bits;
- show that you can indeed observe or measure those bits in the sequence of events;
- demonstrate that this connects to the outcome of the sequence.

Operationalizing the label

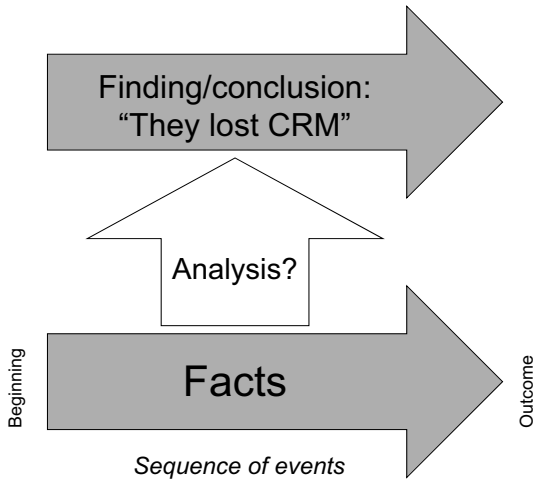


Figure 3.8 Don't make a leap of faith, from your facts to a big label that you think explains those facts. Leave an analytic trace that shows how you got to your conclusion

Figure 3.8 shows the relationship between the sequence of events (called, for convenience, "facts") here and the human factors conclusion you draw from it.[7]

Where do you find good operationalizations? This is what the human factors literature is for: the books and articles written on the basis of research done into human performance. Also, the next chapter will run you through a number of patters in which human performance breakdowns can happen. For now, let's look at two different operationalizations of "loss of effective CRM" as an example. Judith Orasanu at NASA has done research to find out what effective CRM is about.[8]

- shared understanding of the situation, the nature of the problem, the cause of the problem, the meaning of available cues, and what is

likely to happen in the future, with or without action by the team members;
- shared understanding of the goal or desired outcome;
- shared understanding of the solution strategy: what will be done, by whom, when, and why?

This is an operationalization. It breaks the label down into more specific components, for which you can actually seek evidence in your facts.

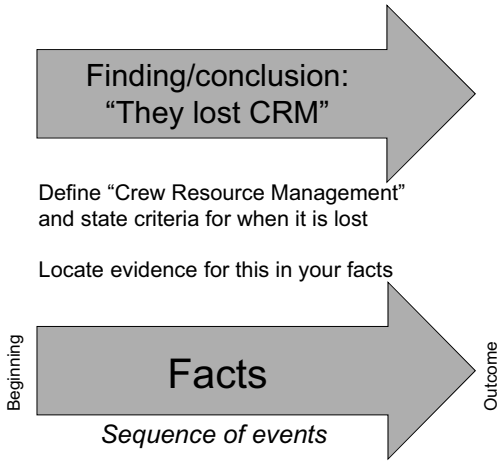


Figure 3.9 The way to bridge the gap between facts and conclusions (about those facts) is to find a definition or operationalization in the literature for the phenomenon and start looking in your facts for evidence of it

Such deconstruction, or breaking down a large concept into smaller, more specific ones, is critical for drawing credible conclusions. You can now, for example, go find evidence for a lack of common understanding of the cause of a problem. Evidence for this can be found in what people do or do not say to one another.

Other operationalizations are possible too (as they are for all human factors concepts, including situation awareness, workload, stress and more!). If you want, you can use multiple operationalizations and build converging lines of evidence. This will make your investigation stronger. In his work for the Australian Transportation Safety Bureau, for example, Maurice Nevile has operationalized loss of effective CRM as follows:[9]

- unequal turn-taking where one person does much more of the talking then others;

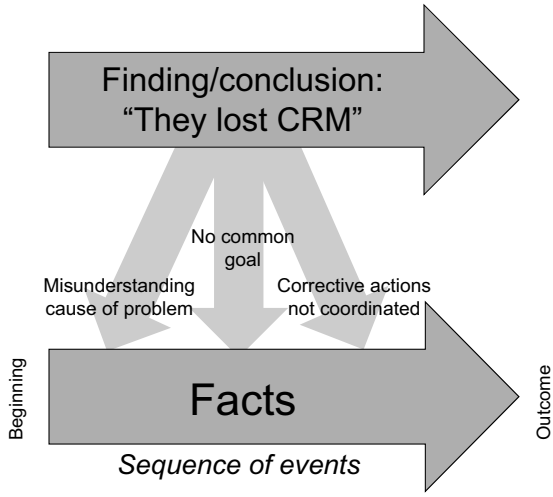


Figure 3.10 Leaving a trace. Using a definition for "loss of effective CRM" that lists misunderstanding the problem, no common goal and uncoordinated corrective actions, you can find evidence for that in your facts

- missing responses where responses are expected or standard, with one person regularly withholding talk, opting out, or replying only in a clipped fashion;
- overlapping talk where another person still has talk of substance to utter but is stepped on by someone else;
- repair of talk done by others. We often engage in repair of our own talk (to correct or clarify our own speech). But if other people repair our talk, this can point to problems in the interaction or hierarchy between them.

How does this connect to the outcome of your sequence? Taken together, these features can create what Maurice Nevile calls a context for error. Because of the nature of people's interaction, errors may become more likely, and their detection and recovery becomes less likely. Rather than just saying "a loss of effective crew resource management," the sort of analysis above allows you to put some meat on your argument. It allows you to leave an analytic trace, and to connect it to the outcome.

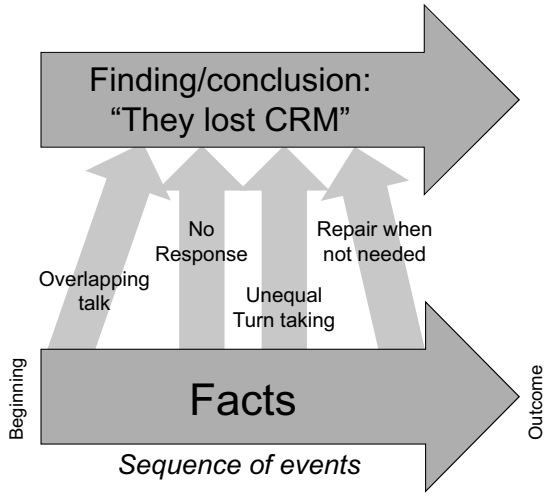


Figure 3.11 Leaving a trace. Overlapping talk, no response when one is expected, unequal turns at talking and offering repair of somebody else's talk when none is needed together could point to a "loss of effective CRM"

Constructing Causes

So how can we identify the causes of the mishap on the basis of the analysis we have done? Let's first deal with the supposed distinction between 'human error' and mechanical failure. The more you understand 'human error,' the less sense that actually makes. Then we turn to the construction of cause, rather than the finding of cause.

'Human error' or mechanical failure?

Was this mishap due to 'human error,' or did something else in the system break? You hear the question over and over again—in fact, it is often the first question people ask. But we have seen in Chapter 1 that 'human error' does not come out of the blue, and is not separate from the system in which it occurs. Error has its roots in the system surrounding it; connecting systematically to mechanical, programmed, paper-based, procedural, organizational and other aspects—to such an extent that the contributions from system and human begin to blur.

Passenger aircraft have "spoilers"—panels that come up from the wing on landing, to help brake the aircraft during its roll-out. Before landing, pilots

have to manually "arm" them by pulling a lever in the cockpit. Many aircraft have landed without the spoilers being armed, some cases even resulting in runway overruns. Each of these events gets classified as 'human error'—after all, the human pilots forgot something in a system that is functioning perfectly otherwise.

But deeper probing reveals a system that is not at all functioning perfectly. Spoilers typically have to be armed after the landing gear has come out and is safely locked into place. The reason is that landing gears have compression switches that tell the aircraft when it is on the ground. When the gear compresses, it means the aircraft has landed. And then the spoilers come out (if they are armed, that is). Gear compression, however, can also occur while the gear is coming out in flight, because of air pressure from the slip stream around a flying aircraft, especially if landing gears fold open into the wind. This could create a case where the aircraft thinks it is on the ground, when it is not. If the spoilers would already be armed at that time, they would come out too—not good while still airborne. To prevent this, aircraft carry procedures for the spoilers to be armed only when the gear is fully down and locked. It is safe to do so, because the gear is then orthogonal to the slipstream, with no more risk of compression.

But the older an aircraft gets, the longer a gear takes to come out and lock into place. The hydraulic system no longer works as well, for example. In some aircraft, it can take up to half a minute. By that time, the gear extension has begun to intrude into other cockpit tasks that need to happen—selecting wing flaps for landing; capturing and tracking the electronic glide slope toward the runway; and so forth. These are items that come after the "arm spoilers" item on a typical before-landing checklist. If the gear is still extending, while the world has already pushed the flight further down the checklist, not arming the spoilers is a slip that is easy to make.

Combine this with a system that, in many aircraft, never warns pilots that their spoilers are not armed; a spoiler handle that sits over to one side of the center cockpit console, obscured for one pilot by power levers, and whose difference between armed and not-armed may be all of one inch, and the question becomes: is this mechanical failure or 'human error?'

One pilot told me how he, after years of experience on a particular aircraft type, figured out that he could safely arm the spoilers four seconds after "gear down" was selected, since the critical time for potential gear compression was over by then. He had refined a practice whereby his hand would go from

the gear lever to the spoiler handle slowly enough to cover four seconds—but it would always travel there first. He thus bought enough time to devote to subsequent tasks such as selecting landing flaps and capturing the glide slope. This is how practitioners create safety: they invest in their understanding of how systems can break down, and then devise strategies that help forestall failure.

The deeper you dig, the more you will understand why people did what they did, based on the tools and tasks and environment that surrounded them. The further you push on into the territory where their errors came from, the more you will discover that the distinction between human and system failure does not hold up.

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 flight deck 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 flight deck 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.[10]

Cause is not something you find. Cause is something you construct. How you construct it, and from what evidence, depends on where you look, what

you look for, who you talk to, what you have seen before and likely on who you work for. I have even been on investigations where the causes that people picked were finely tuned to the people they knew were going to be on the board approving the report and its recommendations. Causes had more to do with political doability of what investigators wanted changed in the organization, than with the data in the sequence of events.

Table 3.1 Two different constructions of causes for the same accident

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

There is no "root" 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), a large number of contributory factors are necessary and only jointly sufficient.

How is it that a mishap gives you so many causes to choose from? Part of the story is the sheer complexity of the systems that we have put together, and that we have protected our systems so well against failure. A lot needs to go wrong for an incident or accident to occur. So you can really construct "causes" from everywhere. 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. That means that any policy or standard that gets you to identify root causes, or probable causes, is of necessity:

Cause is not something you find. Cause is something you construct.

- **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.
- **Oversimplified**. They highlight only a few hotspots in a long, twisted and highly interconnected, much finer-grained web that spreads far outside the causes you come up with.

If you find a root or primary or probable cause or contributory factor, it was your decision to distinguish some things in the dense causal mesh by those labels. So it may actually be better to think in terms of explanations than causes. Or perhaps even turn to interventions, recommendations, improvements or changes instead.

What is the cause of the accident? This question is just as bizarre as asking what *the* cause is of not having an accident.

> *In a break with the tradition of identifying "probable causes" in aviation crashes, 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 191 recommendations which pointed to the many "causes" or systemic issues underlying the 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.*[11]

Analyzing The Blunt End

If you look at some New View investigation reports (you will find pointers at the back of this book, toward the end of Chapter 8), you will see that they are all quite long. This is not surprising, of course. When you have learned how to pursue 'human error' as a consequence, rather than a cause, you quickly end up in the organizational blunt end. That is where you will find many factors, all necessary and only jointly sufficient, that contributed to the bad outcome.

Writing all of that up in prose is possible, of course. And many investigations, as well as revisionist books, do just that. But doing this (or doing only this) has some drawbacks:

- for readers, it is not always easy to distill the contributory factors at the organizational blunt end from a written account;
- for investigators, resource pressures (time, money) may mean that producing a full written account is not always possible.

A more efficient way to get readers to understand the important role played by the organizational blunt end could be to build a flowchart, through which they can trace the various contributions. When you do this, note the following:

- **Start at the blunt end**, not the sharp end. Many investigations do the opposite: they start at the sharp end—at the time and day of the accident. From there they trace back into the blunt end. The problem with that is that the focus remains on the sharp end; with the (apparently) unique, specific events that pushed the system over the edge into failure. If, instead, you start from the blunt end, you can get the reader to understand that the conditions that contributed to the accident are normal; that they were always there, embedded in how the organization did business.
- Make sure, through what you find, that you identify the **organization's model(s) of risk**, and how the organization thought it could control that risk or those risks. It is, after all, the organization's model of risk that made them invest in certain things (for example, automation or standardized procedures to control unreliable operators and 'human error') and ignore others (for example, how production pressures affected people's trade-offs at the sharp end).

An interesting example of an outdated model of risk can be found in the very common testing or examining of candidates for driving licenses. Three-point turns and handbrake-assisted hill starts hail from an era of cars without powersteering and with dodgy clutches. But a major risk for young drivers is losing control and wrapping themselves around a tree or other obstacle during a high-speed single accident. The many safety devices in cars (ABS, slip control, airbags) may have engendered a sense or promise of control where there really is little. Instead of doing three-point turns, a stronger investment in safety could be a course in slipping and skidding in cars with powersteering. Similar holdover practices exist in aviation, where pilots need to demonstrate competency on the kinds of engine failures (right after passing

the decision-speed or no-return speed on the takeoff runway) that hardly happen any longer given the reliability of modern jet engines. In some cases, pilots also need to show that they are capable of handflying approaches with the guidance of instruments that many airports no longer have, and which some airlines actually proscribe altogether. The model of risk, as expressed in what we expect practitioners to be competent in, are obsolete. Risks have been shifting as technology and experience has developed, but has the field moved along with it?

- Trace the difference between **work as imagined and work as actually done**. This is one way of teasing out the organization's model(s) of risk and risk control: they show, after all, how the organization thought work was done (according to rules, procedures and technical specifications, for example) versus how it was actually done under the influence of production pressures and local variations.
- Allow the reader to understand **what made the difference** on the day of the accident. If the pre-conditions at the blunt end were always present and handed down to the sharp end, then the reader will need an explanation for why things got out of hand that one time. Do not rely on 'human error' or "unsafe acts" to explain the difference—that is the cheap way out. Remember that people come to work to do a good job. The way they have learned to adapt around technical difficulties, resource constraints and efficiency pressures, however, may not be well-suited for all conditions in a complex, changing world.

Making Recommendations

The most popular recommendations to come out of 'human error' investigations (in industries ranging from construction to oil & gas to healthcare) are retraining people or rewriting a policy. This may get the organization off the hook or allow the safety department to "close out" their investigations. In fact, if you recommend design or organizational changes, you are likely to invite more opposition. This will cost money, after all, or suggest that some people higher up were not doing their jobs properly. You may then hear things like:

- "We already pay attention to that"
- "That's in the manual"
- "This is not our role, or not our problem"
- "This recommendation has no relevance to the mishap"
- "People are selected, trained and paid to deal with that."

But the sorts of low-end policy or behavioral recommendations that you can easily get away with organizationally, are never very useful. In order to know what to recommend instead, let's look at a couple of things here:

- what to recommend in terms of improvement depends on how safe that particular activity already is;
- there is a difference between explanatory and change factors;
- make your recommendations smart (specific, measurable, agreed, realistic and time-bound—see below).

Explanatory versus change factors

One reason why recommendations can fall by the wayside and not get implemented with any convincing effect is that they are fired off into an organization as limited, one-shot fixes. Many organizations—even those with mature safety departments and high investments in investigations—lack a coherent strategy on continuous improvement. Resources are directed primarily at finding out what went wrong in the past, rather than assuring that it will go right in the future. The focus is often on explanation, not change. And that misses the point of an investigation.

So let's make a difference between:[12]

- explanatory factors which explain the data from one particular sequence of events;
- change factors that are levers for improvement or prevention.

The thing that explains a particular instance of failure does not need to be the same thing that allows your managers to do something about its potential recurrence. Working up from explanation to recommendation can sometimes be very empowering if you are open-minded and creative about it.

> *In one mining operation, workers who maintained the big haul trucks were suffering from hand injuries. Each of these injuries was investigated and a report drawn up. It identified the typical explanatory factors: time pressure, people sticking their hands in places where they don't belong, lack of correct protective equipment, problems of light and visibility. These explanatory factors, however, did little to point the way to prevention or improvement. They suggested to people to stare harder, try more, follow the rules, be careful. None of that worked.*

The change factor came when someone identified that more than 80 percent of the injuries occurred during unscheduled maintenance along the dusty roads of the mine—not during scheduled maintenance in the shed. The company decided to invest aggressively in preventive maintenance (including a larger parts store), to ensure that as few trucks as possible would break down on the road. After it did this, hand injuries went down dramatically. The explanatory factors were not the same as the change factor.

Smart recommendations

If you want your recommendations to get organizational traction, you may want to consider running them by the following criteria (which constitute the SMART acronym):

- **Specific**. You want your recommendation to be clearly specified. Which parts of what organization do you want to do what and when?
- **Measurable**. You want your recommendation to somehow contain "measurable" criteria for success. This will help you and the organization see if it is indeed implemented, and monitor the effects once it is.
- **Agreed** (or at least agreeable). You want your recommendation to take into account legitimate organizational concerns about production or efficiency; otherwise you will not get any traction with those responsible for implementation and achievement.
- **Realistic**. While your recommendations may be difficult, and challenging to the typical worldviews of those who need to get to work with them, you need to keep them realistic—that is, doable for the person responsible for implementing them.
- **Time-bound**. You may want the implementation of your recommendation to have some kind of suggested expiration date, so decision makers can agree on deadlines for achievement.

How To Introduce Your Next 'Human Error' Investigation

If you are wondering how to possibly start your next 'human error' investigation, which arguments to use or how to word them, take a look at the example below. Parts of this may serve as a useful template (if properly referenced) for how to begin writing a New View investigation.

The Aim of This 'Human Error' Investigation

The aim of a 'human error' investigation is to try to understand why it made sense for people to do what they did—against the background of their physical and psychological work environment. The reason for this is simple. If what people did made sense to them at the time (even if it led to a bad outcome), then this may well make sense to others like them as well. If it does, and if we leave in place the conditions that make it so, then we will very likely repeat the bad event. The point of this investigation is not to assign blame or responsibility, but to learn; to learn and improve. It should, in that sense, not even be seen as an investigation, but as a learning review or a learning opportunity. And that is probably what we should call it as well. But we'll come back to that.

A Word on Hindsight

When confronted with a bad outcome, it is always tempting to ask why people didn't do something that, with hindsight and knowledge of outcome, would have made much more sense. This is a natural response. We as people use it to not only express our surprise and incomprehension and perhaps even indignation about what went on in a particular workplace, but also because we want to extract maximum learning from what we see and read and hear about. Questions or remarks such as "But why didn't they…?" or "If only they had…!" have a role to fulfill because they allow us to prepare ourselves and others for similar situations, and perhaps we may then do what we now wish others had done.

Such remarks or questions have no role to play, however, in explaining why people did what they did. Saying what people could or should have done does not explain the reasons behind what they in fact did. Such questions and remarks are, literally, counterfactual—counter the known facts of the accident. What people didn't do (but could or should have done) is not as pressing as finding out why they in fact did what they did, which is where a 'human error' investigation needs to direct its resources.

A standard response after a mishap is to 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. Pointing out, however,

that it should have been does not explain why it was perhaps not, or why it was interpreted differently back then. There is a difference between:

- data availability: what can be shown to have been physically available somewhere in the situation;
- data observability: what would have been observable given the features of the interface and the multiple interleaving tasks, goals, interests, knowledge of the people looking at it.

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

The most important thing to remember is that the people involved didn't know that the bad outcome was going to happen. Which we do now know. The people involved did not enjoy the hindsight we now have. If they had known, they would almost certainly have done something other than what they factually did, because we have to assume that it is not the outcome they desired. This is why counterfactuals, which talk about something the crew didn't factually do (for example: "if only they had…") are not very helpful for understanding why the crew did what they factually did. This also goes for reference to procedural guidance that would have applied during that time. While it is easy to point out where such procedural guidance may not have been followed, the real point is to find out why it would have made sense for people to not to precisely follow this guidance in the context of their situation and their unfolding understanding of the circumstances.

Knowing the outcome of a sequence of events can easily bias an investigation toward those data points that we now know were significant; that showed the real nature of the situation. But the world surrounding people at the time consisted of much more than just those few data points that we now know. After all, it is only now, with the outcome in our hands, that we know which ones were critical and which ones weren't. In what was likely a cognitively noisy and pressurized situation, with time constraints, multiple interleaving and overlapping activities and tasks, and a host of indications that needed attention, people had to accomplish a meaning integration that we, in hindsight, have no trouble accomplishing. But that is because we know the outcome. Without knowledge of the outcome, certain tasks and data points become glaringly obvious. But that is only with hindsight.

Knowledge of outcome tends to simplify a situation down to a choice to notice the critical data or not notice the critical data, to have the accident or not have the accident. That, of course, is not how the world looked to people at the time. This conversion into simple, linear causality is called the hindsight bias. It significantly distorts our ability to understand why it made sense for people to do what they did. The hindsight bias occurs because we know, and we can start from, the outcome. We can trace back from it, into the assessments and decisions that led up to it. Tracing back up the causal flow, it is easy to identify the junctures where people could have done something to avoid the outcome; where they could have zigged instead of zagged. Again, this makes us no wiser in explaining why the crew did what they actually did. The question is not what people should have done in order to avoid the outcome we now know (that is, given our understanding of the situation). The question is why people did what they did given their understanding of the situation (which did not include the outcome).

That is why the analysis offered here (as do many other 'human error' analyses) tries to reconstruct the unfolding situation from the point of view of the people involved as faithfully as possible. If we want to understand why people did what they did, we have to reconstruct the situation as it unfolded around them at the time, not as it looks to us now. If we can't comprehend what the crew did or why they did it, it is not because the crew's behavior was incomprehensible, but because we have taken the wrong perspective, the wrong point of view. People at the time, after all, acted on their understanding of the situation as it developed around them at the time. What was the pace, the tempo, the order in which cues and indications about the situation emerged and confronted the crew at the time? This is likely different from our understanding that we have gained over months of study after the fact. The aim of any such reconstruction is plausibility—knowing all the time that no amount of dedication, theory, or analysis is enough to ever entirely close the gap to full certainty about why people did what they did.

The Scope of This 'Human Error' Analysis

It has become custom in 'human error' analyses to spread out from the so-called "sharp end" (for example, the control room or operating theater) and up into the operation and organization and regulatory or industry environment that surrounded it. In other words, 'human error' analysis should not just focus on the proximal factors (those close in space and

time to the actual accident), but also on the distal factors that contributed and could again express themselves (but then in a different incident) if not addressed. The aim of a distal reach is to explore how aspects of training, documentation, operational or managerial decision making, regulatory influences or industry standards may all have influenced what people at the sharp end saw as making sense at the time. The value of doing this is as great as its difficulty. The further away in space–time we move from the accident, the more difficult it becomes to couple particular actions and assessments (by people far removed from that cockpit) to what went on at the sharp end on any particular occasion.

But the attention to distal factors, factors from further away in space and time, is justified. People at the sharp end do not work in a vacuum. Their performance is in great part the product, or the expression, of the environment in which they work, the environment from which they come, the environment which prepared them for the job. This includes, for example, how the organization or industry has trained and educated its practitioners, what it focuses on in proficiency checking and examining, how the industry has designed and maintained its equipment, how it has managed earlier incidents that bear a relationship to what happened on the mishap day, and so on. If it made sense for these particular people to do what they did, it may make sense to others as well who find themselves in a similar situation. Thus, attention to distal factors is necessary (where possible) also from a preventive perspective.

A Word on Time

There is something hugely artificial about any post-hoc analysis as conducted and the format in which it is delivered. It has taken months to excavate and study the data remnants from a few days, hours or perhaps even minutes of human performance. And it has taken scores of pages to represent the result of this study. In other words, we have had the luxury to consider and reconsider the assessments and decisions and actions of a crew, not only with the benefit of hindsight, but with the benefit of ample time and resources, able to draw on expertise from wide and far. This may contrast sharply with the few minutes the practitioners may have had to consider their assessments and decisions and actions in real time. They may have had about as much time for all that as it took you to read just this paragraph. This makes it difficult, but even more important, to do everything you can to put yourself in their shoes and to try to see the world from their point of view.

In that sense, this investigation, or learning opportunity, is as much about them as it is about you. If you still cannot understand why they did what they did, perhaps take a moment to stop looking at them. Start looking at yourself. Ask yourself what biases and reactions to failure you are bringing to this mishap that prevent you from understanding why it made sense for them to do what they did.

A Word on Investigating versus Learning

Perhaps this investigation should be called a learning review. That, after all, is what it is trying to do. And the people involved in it should not be called "investigators" but a "learning team" or "learning review team." "Investigation," after all, brings with it all kinds of negative connotations. Connotations like questioning people, interrogating them, finding out where they went wrong. "Investigation" suggests that the ultimate aim is to find out where things went wrong, to be able to offer the one official or definitive account of what happened. It isn't, or it shouldn't be. The ultimate aim is to learn and improve.

The questions that people involved in any incident should really be getting are, "What do you guys want to do, or need, to make this better?" Next time we conduct an "investigation," let's ask that question and see what we get. The results may be amazing. People will tell how they work in a complex world, and how they have learned to adapt around the many goal conflicts and resource constraints in it. Some of those ways may be getting a bit brittle, some may be really smart. But if we don't talk about it openly and honestly, we won't know.

Let us, above all, remember that nobody came to work to do a bad job, and that the ultimate aim of this activity is to learn and improve. Anything that gets in the way of doing that will further harm the organization, and the people in it.

Notes

1 Klein, G.A. *Sources of power: How people make decisions*. Cambridge, MA: MIT Press, 1998.

2 Nevile, M., Walker, M.B. *A context for error: Using conversation analysis to represent and analyse recorded voice data*. Canberra, ACT: Australian Transport Safety Bureau, 2005.

3 Sarter, N.B., Woods, D.D. Pilot interaction with cockpit automation II: An experimental study of pilots' model and awareness of the flight management system. *International Journal of Aviation Psychology* 1994;4(1):1–29.

4 Aeronautica Civil. Aircraft accident report: Controlled flight into terrain, American Airlines flight 965, Boeing 757–223, N651AA near Cali, Colombia, 20 December 1995. Bogota, Colombia: Aeronautica Civil de Colombia, 1996.

5 Ibid.

6 Ibid.

7 Xiao, Y., Vicente, K.J. A framework for epistemological analysis in empirical (laboratory and field) studies. *Human Factors* 2000;42(1):87–102.

8 Orasanu, J.M., Connolly, T. The reinvention of decision making. In: Klein, G.A., Orasanu, J.M., Calderwood, R., Zsambok, C.E., editors. *Decision making in action: Models and methods*. Norwood, NJ: Ablex, 1993:3–20.

9 Nevile, *op. cit*.

10 Aeronautica Civil, *op. cit*.

11 Moshansky, V.P. Final report of the commission of inquiry into the Air Ontario crash at Dryden, Ontario. Ottawa, Canada: Ministry of Supply and Services, 1992.

12 Stoop, J., Dekker, S.W.A. Are safety investigations proactive? *Safety Science* 2012;50:1422–30.

This page has been left blank intentionally

4 Explaining the Patterns of Breakdown

So, what went wrong? What patterns can you identify in your data? This chapter introduces some human performance issues that could guide you:

- cognitive fixation;
- plan continuation;
- fatigue;
- buggy or inert knowledge;
- new technology and computerization;
- procedural adaptations;

It then discusses the currently popular use of "loss of situation awareness" and "complacency" as supposed explanations for 'human error.' The advice of the field guide is very simple: don't use these things. It will explain why.

Of course, the results of a 'human error' analysis may remain infinitely negotiable and never become entirely certain. In a 'human error' analysis, the data we have available about what people did never allows us to entirely close the gap to *why* they did it. The data, in other words, always underdetermines the explanations a human factors analysis can offer. It is not surprising that human factors can often be seen as vague, speculative, indeterminate, or as merely a matter of opinion, and not in great need of the same level of expertise as a technical part of an investigation (after all, everybody has a mind, so we all should have some idea of how a mind works, right?). So others may always be able to draw different conclusions about human performance from the data. This is not to say that some explanations are not more plausible than others. Some explanations account for more of the data, or account for the data better and at a greater level of detail, or make better use of the theoretical, methodological and research bases for contemporary 'human error' research. The account offered in a New View 'human error' investigation should aim for a maximally plausible explanation of why people did what they did. It should use a number of analytical and theoretical perspectives to lay out the broadest possible basis for its conclusions about what went on in that workplace at the time.

Cognitive Fixation

A difficult issue in understanding 'human error' is how other people did not seem to see the situation for what it turned out to be, which you now know thanks to your hindsight. When people are in the middle of things, without knowledge of outcome, sensemaking is ongoing (see Figure 4.1). People's actions and assessments of what is going on are intertwined. By doing something, people generate more information about the world. This in turn helps them decide which actions to take next. The dynamics of sensemaking in unfolding situations, however, can create some interesting effects.

> *In 2001, the pilots of a twin-engine airliner flying from Toronto to Lisbon, with 293 passengers and 13 crew onboard, noticed a fuel imbalance. One wing held more fuel than the other, which was strange: the left and right engines should consume about as much fuel from their respective tanks. The pilots reviewed engine indications and found no anomalies. Fuel flow to both engines was the same. They could think of no in-flight event or condition responsible for a fuel loss from one wing tank. They even sent a*

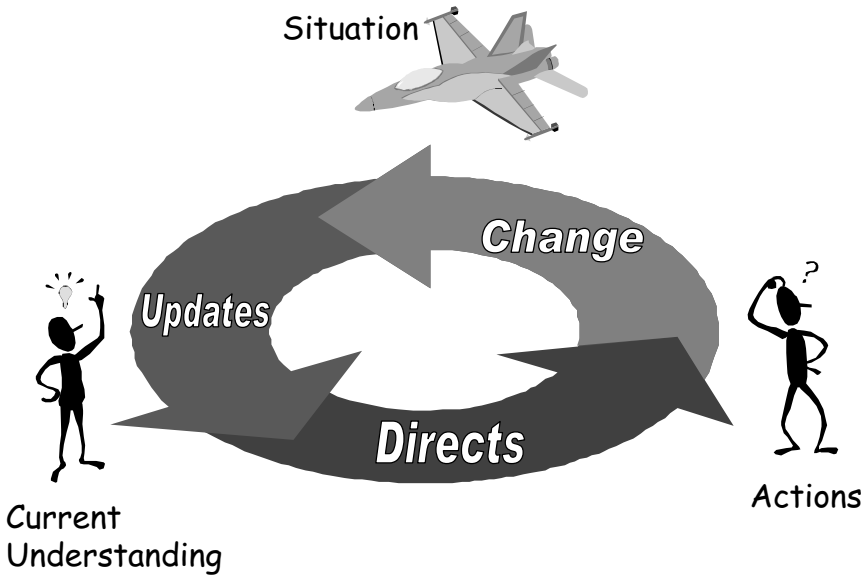


Figure 4.1 **We make assessments about the world, which update our current understanding. This directs our actions in the world, which change what the world looks like, which in turn updates our understanding, and so forth[1]**

crew member to visually check whether there was fuel leaking from the wing, but due to darkness none was seen. Handbooks and checklists did not help in resolving the mystery. In fact, the fuel quantity indications, calculations of fuel loss rates, as well as computer predictions of fuel remaining on arrival were truly "unbelievable." They made no sense. The crew concluded that they had an indication problem, but decided to divert to the Azores (Islands in the Atlantic Ocean) because of low fuel predictions in any case.

About 45 minutes after the first fuel imbalance caution, the right engine stopped. The diversion to the Azores now became a very good idea. The idea of an indication error collapsed: this was for real. Not much later, the left engine stopped as well (the pilots earlier had cross-fed fuel from one wing tank to the other). The pilots managed to glide the powerless airliner over a distance of 65 nautical miles to the Azores and land safely at Lajes Airport there. Later investigation turned up that a fuel hose in the right engine had come undone inflight because of rubbing against a non-approved part for that engine. Large quantities of fuel had been dumped overboard.

A case like the one above emphasizes the following about sensemaking, especially when people are faced with an unfamiliar or unexpected problem:

- These are situations without a well-formulated diagnosis of the problem. Handbooks and checklist are of little help in those cases.
- People have to make provisional assessments of what is going on based on partial and uncertain data.
- People's situation assessment and corrective actions are tightly interwoven. One constrains and informs the other.
- Taking action simplifies the diagnostic problem. It commits people to a particular interpretation.
- The side effect of taking action is that people then build an explanation that justifies their action. This explanation, as Karl Weick suggested, may persist and can get transformed into an assumption that is then taken for granted.

Characteristic of cognitive fixation is that the immediate problem-solving context biases people in some direction (for example, "this is an indication problem"). From uncertain, incomplete and contradictory data, people have to come up with a plausible explanation. A preliminary interpretation allows them to settle on an explanation that covers the data observed. But it can activate certain kinds of knowledge and trouble-shooting activities at the expense of others.

The crew of the airliner above did not believe there was a fuel leak. These factors supported their interpretation of the situation:

- *The combination of the suddenness and the magnitude of the indicated fuel loss were outside anything that could reasonably be explained by a real fuel loss.*
- *There had been earlier indications of an anomaly in oil cooling. This had created uncertainty about the nature of the problem (and whether it was fuel related at all).*
- *There was no warning or caution from the airliner's computer systems warning them of a severe problem.*
- *No other indication of an engine problem was discovered: fuel flow parameters were normal.*
- *Some information, such as the cabin crew confirming that there were no visible signs of a leak, confirmed their interpretation of an indication anomaly.*

In addition, the pilots had never before experienced a fuel leak or similar events. They had not been trained for this situation. They had neither the knowledge basis nor compelling data on which to build a plausible explanation of a real fuel leak.

Whether or not to abandon an initial interpretation is not about people's motivation. They are *very* interested in getting it right – in understanding what is going on. Because if they don't, they will be the first to arrive at the scene of the accident. Instead, it is about a cognitive balancing act. Imagine trying to understand and simultaneously manage a dynamic, uncertain situation:

- Should you change your explanation of what is going on with every new piece of data that comes in? This is called "thematic vagabonding," a jumping around from explanation to explanation, driven by the loudest or latest indication or alarm. No coherent picture of what is going on can emerge.
- Or should you keep your explanation stable despite newly emerging data that could suggest other plausible scenarios? Not revising your assessment (cognitive fixation) can lead to an obsolete understanding.

There is no right or wrong here. Only hindsight can show you whether people should have abandoned one explanation in favor of another, and when, or whether they should have settled on a stable interpretation instead of pursuing the latest clue. To understand human performance from the point of view

of people inside the situation, you need to acknowledge the existence and importance of this balancing act. Thematic vagabonding and getting cognitively locked up are opposite extremes, created by uncertain, dynamic situations in which we ask people to solve difficult, unclear problems.

Dynamic fault management

Another aspect of managing such problems is that people have to commit cognitive resources to solving them while maintaining process integrity. This is called dynamic fault management, and is typical for event-driven domains. People need to keep the aircraft flying (or the patient breathing) while figuring out what is going wrong. Not trouble-shooting or correcting may challenge the integrity of the entire process.

> *In the fuel leak case described above, the dual demands of maintaining process integrity while diagnosing the problem, increased crew workload. Their work included:*
>
> - *flying the aircraft, including navigating it to a new destination and coordinating with air traffic control and the airline;*
> - *diagnosing the fuel fault, which entailed looking up information in handbooks, requesting assistance from cabin crew, calling up the maintenance center, acquiring new data from computer systems and all the assessments and decisions these things triggered or called for;*
> - *managing the engine failures and preparing the aircraft for a possible ditching.*

Plan Continuation

What if the situation does change while people are working on a problem? This happens in dynamic domains a lot. While on approach to an airport, for example, weather conditions may deteriorate, but crews continue. While working on a patient, a problem occurs that seems to call for one intervention (for example, more stimulant intravenous drugs), which eventually becomes or aggravates the source of the problem.

Sticking with the original plan while a situation has actually changed and calls for a different plan is what Judith Orasanu calls "plan continuation."[2] This is an interesting line of research out of the NASA Ames Research Center. The NASA research takes as its starting point the psychology of decision making consistent with the last decades of research into it. Decision making in complex, dynamic settings, such as an approach, is not an activity that involves a weighty

comparison of options against pre-specified criteria. Rather, such decision making is "front-loaded:" this means that most human cognitive resources are spent on assessing the situation and then re-assessing it for its continued do-ability. In other words, decision making in a dynamic situation is hardly about making decisions, but rather about continually sizing up the situation. The "decision" is often simply the outcome, the automatic by-product of the situation assessment. This is what turns a decision on whether to continue with a plan into a constantly (re-)negotiable issue: even if the decision is not made on the basis of an assessment of the situation now, it can be pushed ahead and be made a few or more seconds later when new assessments of the situation have come in.

In plan continuation, early and strong cues suggest that sticking with the original plan is a good, and safe, idea. Only later, and weaker, cues suggest that abandoning the plan would be better. In hindsight, it is easy to forget to see the cues from the point of view of people at the time, and when and how strongly they appeared.

Even more important than the cognitive processes involved in decision-making, are the contextual factors that surround people at the time. The order in which cues about the developing situation come in, and their relative persuasiveness, are two key determinants for plan continuation. Conditions often deteriorate gradually and ambiguously, not precipitously and unequivocally. In such a gradual deterioration, there are almost always strong initial cues that suggest that the situation is under control and can be continued without increased risk. This sets people on the path to plan continuation. Weaker and later cues that suggest that another course of action could be safer then have a hard time dislodging the plan as it is being continued, and as evidenced from the situation as it has so far been handled.

In summary then, plan continuation means sticking to an original plan while the changing situation actually calls for a different plan. As with cognitive fixation, situational dynamics and the continuous emergence of incomplete, uncertain evidence plays a role:

- Early cues that suggest the initial plan is correct are usually very strong and unambiguous. This helps lock people into a continuation of their plan.
- Later cues that suggest the plan should be abandoned are typically fewer, more ambiguous and not as strong. Conditions may deteriorate gradually. These cues, even while people see them and acknowledge them, often do not succeed in pulling people into a different direction.

To understand this, make sure that your timeline shows which cues came in when. That way, you can begin to lay out what might have biased people into believing the plan was working, and what (later on) unconvincingly pointed in a different direction. Remember that they did not know the outcome, otherwise they would likely have changed their plan as well!

Why not make a go-around?

In hindsight, investigators in aviation often ask why the flight crew didn't make a go-around when they realized that they would not be able to complete the landing safely—or when all the actions to be completed before a certain height have not yet been accomplished (like being on the right speed, engine power and position, having the airplane configured with wheels and flaps where they need to be and with appropriate checklists completed). If all those actions are not completed before a certain height, then the approach is called "unstabilized." And, according to the book, unstabilized approaches should be broken off. The problem is, they often are not.

Continuing an approach against written guidance (particularly when events are considered in hindsight) is a common problem and may well have parallels in other operational worlds. The Flight Safety Foundation sponsored a study in the late 1990s to analyze the factors that play into approach and landing accidents, and concluded in part, predictably, that executing a missed approach is one of the best investments that pilots can make in safety to prevent approach and landing accidents.[3] *Such advice should not be confused with an explanation for why many crews do not do so. In fact, getting crews to execute go-arounds, particularly in cases of unstabilized approaches, remains one of the most vexing problems facing most chief pilots and managers of flight operations across the world. Characterizations such as "press-on-itis" do little to explain why crews press on; such words really only label a very difficult problem differently without offering any deeper understanding.*

Recommendations directed at flight crews are, as a result, difficult to implement. For example, one recommendation is that "flight crews should be fully aware of situations demanding a timely go-around." Even if crews can be shown to possess such knowledge in theory (for example, they are able to recite the criteria for a stabilized approach from the airline's Operations Manual), becoming aware that a timely go-around is demanded hinges precisely on a particular awareness of the situation itself. The data from continued approaches suggest that crews do not primarily interpret situations in terms of stabilized approach

criteria, but in terms of their ability to continue the approach. Soft and hard gates (for example, 1,000 feet, 500 feet), when set in context of the end of a flight at a busy, major airport on a scheduled flight, become norms against which to plan and negotiate the actual approach vis-à-vis the air traffic and weather situation, not iron-fisted stop rules for that approach.

Thus, there seems little mileage in just reminding crews of the situations that demand a go-around, as almost all crews are able to talk about such situations and offer advice about how to do the right thing—until they are players in a rapidly unfolding situation themselves. When on approach themselves, it is not primarily generic criteria that crews see against which they can make some cognitive calculus of the legality and wisdom of continuing their approach. When on an approach themselves, crews see a situation that still looks doable, a situation that looks like they can make it, a situation that may hold a useful lesson for their line-flying student in the right seat, a situation that suggests things will be all right before passing over the threshold. This is entirely consistent with decades of human factors research. Operational decisions are not based on a "rational" analysis of all parameters that are relevant to the decision. Instead, the decision, or rather a continual series of assessments of the situation, is focused on elements in the situation that allow the decision maker to distinguish between reasonable options. The psychology of decision making is such that a situation is not assessed in terms of all applicable criteria (certainly not quantitative ones), but in terms of the options the situation appears to present.

One promising countermeasure seems to be not to remind crews of the criteria for a stabilized approach, but to offer generic rewards for all the cases in which crews execute a missed approach. Chief pilots or managers of flight operations who offer a no-questions-asked reward (for example, a bottle of wine) to crews who make a go-around, generally report modest success in reducing the number of unstabilized approaches. It is crucial, in setting up such policies, to communicate to crews that breaking off an approach is not only entirely legitimate, but actually desired: getting pilots to buy into the idea that "each landing is a failed go-around." The handing out of such rewards should be advertised broadly to everybody else. Such encouragement is of course difficult to uphold in the face of production and economic pressures and incredibly easy to undermine by sending subliminal messages (or more overt ones) to crews that on-time performance or cost/fuel savings are important. The paying of bonuses for on-time performance, which has been, and still is, custom in some airlines, is an obvious way to increase the likelihood of unstabilized approaches that are not broken off.

Of course, any intervention, whether policy-based or otherwise, still has to deal with the difficult issue of getting crews to make an all-engine, full-power climb with a light airplane (because at the end of the flight most fuel will have been used up). This has shown to be a difficult maneuver, one that is in fact seldom practiced, and with which many crews are not very familiar.

Fatigue

Fatigue is a common condition, especially in work that cuts across normal waking periods or time zones. The effects of fatigue on safety-critical work are difficult to measure and quantify, as they are so confounded (they may well be the effect of something else, or of a combination of other factors). What research (and experience) do show, however, is that rest is the only really effective way of dealing with it—not coffee or energy drinks or sliding down the truck window or drumming on your knees. That said, fatigue itself can actually be difficult to pin down. Using people's self-reports or their judgments of colleagues is not very reliable. Fatigue actually impairs people's judgment about how fatigued they are and how it affects their performance.

In one accident where fatigue was said to play a role, investigators found how "the captain and the first officer had been continuously awake for at least 16 hours." Research indicates that the normal waking day is between 14 and 16 hours and that lapses in vigilance increase and become longer if the normal waking day is extended.

In addition, the Safety Board's 1994 study of flight crew-related major aviation accidents found that flight crews that had been awake for an average of about 13 hours made significantly more errors, especially procedural and tactical decision errors, than crews that had been awake for an average of about five hours … The "accident time was nearly two hours after the time that both pilots went to bed the night before the accident and the captain's routine bedtime…"

Research indicates that the ability to consider options decreases as people who are fatigued become fixated on a course of action … Also, automatic processes (such as radio calls and routine behavior) are affected less by fatigue than controlled processes (such as more complex behavior, responses to new situations and decision making). Further, fatigue deteriorates performance on work-paced tasks that are characterized by time pressure and task-dependent sequences of activities."[4]

As indicated in the example above, fatigue can be the result of a number of factors (or a combination), for example:

- workload intensity (or, conversely, inactivity) and sustained effort (performance deteriorates as a function of "time on task");
- physical and psychological exertion;
- sleep deprivation or other sleep disturbances;
- time of day effects;
- circadian desynchronization.

While performance effects of fatigue are often difficult to prove, some of the following effects have been linked to fatigue:

- Vigilance effects. Tasks requiring sustained attention or quick reaction times are particularly vulnerable.
- Cognitive slowing. Slower responses on all kinds of cognitive tasks, such as reasoning, tracking, arithmetic, generating options and making decisions.
- Memory effects. Build-up of memory compromised by attention deficits (and possibly lapsing).
- Lapsing. Also called "micro-sleeps" (from ten seconds in duration), where people do not respond to external stimuli.

A "New View" on fatigue would suggest that promising solutions lie in scheduling and rostering, and in other conditions surrounding people's work.

Buggy and Inert Knowledge

The application of knowledge, or using it in context, is not a straightforward activity. In order to apply knowledge to manage situations, people need to:

- possess the knowledge;
- have that knowledge organized in a way that makes it usable for the situation at hand;
- activate the relevant knowledge in context.

Ask yourself whether the right knowledge was there, or whether it was erroneous or incomplete. People may have been trained in ways that leave out

important bits and pieces. The result is buggy knowledge, or a buggy mental model (with gaps, holes and bugs in it).

> *There was an account of pilots trying to take off in snowy weather who relied on the jet blast of the aircraft in front of them to clear the latest snow off the runway. This may be a case of pressure to takeoff without waiting for plows to clear the snow, but also in part a case of buggy knowledge about how effective that clearing is. The question is not how pilots can be so silly to rely on this, but how they have been led to believe (built up a mental model) that this is actually effective.*

Just like the presence of knowledge, its organization is also in part the result of how the material is taught or acquired.

> *Feltovich has investigated how knowledge can be wrongly organized, especially in medicine, leading to misconceptions and misapplications. One example is that students have learned to see highly interconnected processes as independent from one another, or to treat dynamic processes as static, or to treat multiple processes as the same thing, since that is how they were taught. For example, changes in cardiac output (the rate of blood flow, which is the change of position of volume/minute), are often seen as though they were changes in blood volume. This would lead a student to believe that increases in cardiac output could propagate increases of blood volume, and consequently blood pressure, when in fact increases in blood flow decreases pressure in the veins.*
>
> *Another example where knowledge organization can mismatch application is in Problem-Based Learning (PBL). Students might see one instance of a problem they are confronted with in training as canonical for other instances they will encounter subsequently. This is overgeneralization: treating different problems as similar.*[5]

Even when people can be shown to possess the knowledge necessary for solving a problem (in a classroom where they are dealing with a textbook problem), but that same knowledge won't "come to mind" when needed in the real world; it remains inert. If material is learned in neat chunks and static ways (books, most computer-based training) but needs to be applied in dynamic situations that call for novel and intricate combinations of those knowledge chunks, then inert knowledge is a risk. In other words, when you suspect inert knowledge, look for mismatches between how knowledge is acquired and how it is (to be) applied.

New Technology and Computerization

Human work in safety-critical domains has almost without exception become work with technology. This means that human–technology interaction is an increasingly dominant source of error. 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 toward breakdown.

- New technology can lead to an increase in operational demands by allowing the system to be driven faster; harder; longer; more precisely or minutely. 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. They have to "finish the design" in practice.
- New technology shifts the ways in which systems break down, asking people to acquire more knowledge and skills, to remember new facts, settings or entries. This adds new vulnerabilities and can open new routes to breakdown.

In general, here are some of the things that may happen:

- **Mode error**. The user thought the computer was in one mode, and did the 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 apparently meaningless repetition, such coordination may not happen consistently.
- **Workload**. 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 busy programming computers when other things were equally deserving of their attention.

- **Data overload**. People may be forced to sort through large amounts of data produced by their computers, and may be 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. This requires a lot of cognitive work by the human.

How to notice that something is not happening

In a number of automation-related accidents, the machine kept on doing something for longer than it was supposed to (or at least longer than the human operators had intended it to). This may be a climb in an airplane that gradually loses energy and stalls at a high altitude, the continuation on a heading in a cruise ship that takes it miles off course and into an obstacle, or the supply of 100 percent oxygen to a patient when it was no longer necessary. In all these cases, the question is how people are to notice a non-event. How can they see that something is not happening? So instead of people not noticing changes, these accidents were about people not noticing an absence of change.

Studies on the monitoring of dynamic processes have shown that it is very difficult for people (as well as machines) to notice non-events. Things that do not happen are not meaningful or informative phenomena in the monitoring of dynamic processes. Something that is not happening is not a good trigger for human (or even machine) intervention. These studies show that non-events, that is, the lack of change over time (or even the lack of change in the derivative, i.e. lack of change in a change) are associated with difficulty for practitioners to detect meaningful phenomena in their monitored process. Evolutionarily, change is information; lack of change has little information or adaptive value. The sensory and neural systems of most organisms, including mammals (which includes humans) are highly attuned to pick up and respond to change, while reserving cognitive resources for other tasks when there is no noticeable change in the environment.

Studies also show that requiring practitioners to increase their "mode awareness" by calling out what mode their automated system is in is of limited value. Such call-outs typically get shed when workload is higher or when other tasks are taking up the practitioners' attention. Also, many systems do not support mode awareness very well either, because they might just show a label of mode status rather than give the practitioners a sense of the system's current and intended behavior.

- **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. People can have a hard time discovering that automation is not behaving according to their intentions. There is little evidence that people are able to pick the mismatch up from the displays or indications that are available—again because they are often status-oriented and tiny. People will discover that the automation has been doing something different when they first notice strange or unexpected process behavior. Serious consequences may have already ensued by that time.[6]

Speed tapes: How a cockpit knows its speed

Today, in automated cockpits, the awareness of airspeed, built up through preparatory work in the cockpit leading up to the approach, may not be supported in the same way as it was in older-generation cockpits. The calculation and representation of airspeeds, the cockpit work and crew cooperation associated with setting and keeping those speeds, as well as the manipulation of artifacts associated with airspeed, has undergone a dramatic change since the introduction of flight management systems and large computer displays to show instrument readings into airliner cockpits. All ways of representing information about a dynamic process have advantages and disadvantages. They all affect the cognitive, collaborative and manual work that needs to go on in and among human operators in their own ways—both good and bad. New ways of representing and manipulating information may take certain traps and error opportunities away, but they will inevitably open new possible pathways to failure.

In 1995, ethnographer and pilot Ed Hutchins published a seminal paper called "How a cockpit remembers its speeds"[7] which traced in great detail

the cognitive and collaborative work that goes on in non-flight management system and non-primary flight display-equipped airliners to look up, compute, cross-check, remember and then monitor speeds associated with different flap settings on an approach. This cockpit work begins after initiation of the descent and what it involves:

For example, many older cockpits use the table in the operations manual and a hard plastic landing data card on which the arrival weather conditions, go-around thrust settings, landing gross weight and landing speeds are indicated with a grease pencil. Still others use the table in the operations manual and write the speeds on a piece of paper (flight paperwork, printout of destination weather and so forth). [*Some use*] *a booklet of speed cards. The booklet contains a page for each weight interval (usually in 2,000 pound increments) with the appropriate speeds permanently printed on the card.*

The preparation of landing data consists of the following steps:

1. *Determine the gross weight of the airplane and select the appropriate card in the speed card booklet.*
2. *Post the selected speed card [or other written representation of speeds] in a prominent position in the cockpit.*
3. *Set the speed bugs on both airspeed indicator (ASI) instruments to match the speeds shown on the speed card.*

In preparing such a cockpit for the approach, in other words, there are activities such as looking up, cross-comparing (with a/c gross weight), writing down, positioning a physical artifact (speed booklet or other piece of paper with speeds on it). There will also very likely be some discussion, as well as the physical manipulation of speed bugs around the airspeed indicator and the visual and verbal cross-checking of those settings on the right and left airspeed indicator while or once they are completed. This rich variety of active work is very likely able to help the construction of a strong memory trace about what speeds will be appropriate for the coming approach. Not only is memory of speeds externalized to some extent (for example, in the speed booklet and bugs); the physical manipulation and active verbal and visual engagement of these external artifacts forces crews to consciously consider each speed step down to approach speed—both in advance of, and during airframe configuration changes (where speeds may be called out and cross-checked once more). This all contributes strongly to a memory trace in pilots' own minds. What marks such cognitive work, however, is not people simply recalling numbers. Instead, significant functions in the cockpit are

achieved by people working with and interpreting material symbols: the memory of which speeds to fly is not just in pilots' heads, it is distributed across the entire cockpit, built over time and represented in a number of physical artifacts with which the crews interact manually, visually and verbally.

This intricate cognitive architecture is altered with a change in technology. The introduction of the flight management system and large computer displays (primary flight display) has made almost all of these earlier crew manipulations of artifacts unnecessary. The flight management system provides all the speeds based on aircraft gross weight, and sets the bugs automatically on the speed tape of the primary flight display. The reference speed (or Vref) for the desired final flap setting appears automatically in the FMS (based on aircraft gross weight), and gets cross-checked verbally and visually only once as part of the descent checklist, long before actually achieving and having to keep that speed. And, while "set speed" is called out under certain conditions during configuration changes, there is no active verbal engagement with the actual speed value for that particular configuration.

The new design evidently takes away a host of error traps inherent in the previous arrangement (for example, where possibilities included selecting the wrong gross weight, selecting the wrong page in the booklet or writing errors on the piece of paper, setting bugs incorrectly along the airspeed indicator). But, again, there is no neutral ground in design. The new design changes the work associated with figuring out and remembering approach speeds and reduces the cognitive depth of crews' engagement with the task. There is much less physical manipulation and transferal of speed figures across different representations, less cross-checking and less verbal coordination that involves mentioning the various figures. It is likely that this has consequences for crews' appreciation of configuration/speed relationships, as well as for the build-up of memory traces about speeds in their own minds.

Monitoring round dials versus tapes

One of the strong advantages of the round airspeed indicator turned out to be an unintended by-product of its design. It, together with the bug settings, allows crews to monitor in terms of speed spaces.[8] *Rather than matching the pointer of the dial with a particular bug (let alone a particular airspeed number), the visual matching task becomes one in which the pointer needs to be kept inside a particular space (like a pie slice on the round dial) for*

any particular configuration—a "speed space." Over time, crews will build up a memory of which speed spaces (or regions or pie slices on the dial) roughly match particular airframe configurations (even with adjustments for high or low gross weights on landing) and will construct expectations of the immediate visual appearance of the ASI for particular phases of the approach. Pilots learn to see regions of the airspeed indicator scale as having meanings that are not expressed in numbers. Even a furtive glance will reveal immediately whether the pointer is in the right space or not.

There are other features of a round speed dial that support the visual matching task (and thereby speed monitoring). One is that the endpoints are visible simultaneously, providing the floor and ceiling and thereby the entire context for possible travel of the airspeed pointer. The pointer sitting in, or traveling into or out of a particular region, relative to the visible endpoints of the scale, carries immediate meaning about configurations and configurability that needs no further staring or inquiry. Also, the velocity with which the pointer travels into or out of a region (and into the next one) is an immediate visual representation of a speed trend—a representation that requires no technological enhancement to be meaningful to a human observer.

The speed tape, in contrast, carries no such inherent meaning. This does not mean that crews do not gradually get used to a speed tape as opposed to a round dial, because experience shows that they do. But on a speed tape, the entire notion of what is up and what is down is in fact negotiable, as is the question of what should go up or down during a speed increase or decrease. Endpoints are not visible during an approach, which means there is little context available in which to place movements of the tape. Speed spaces or regions do not exist in any discriminable way, because, to the extent that there are any, they look like a linear portion of a linear tape, not a pie slice of a unique size, uniquely angled along a circle in a unique direction.

Noticing change on a speed tape is harder than on a round dial. The difference between large changes and small changes are easy to note on a round dial. In contrast, the difference between large and small changes almost gets lost in the linear format of a speed tape. This, in part, has made it necessary to artificially engineer a visual trend back into the display (with a trend arrow), so as to make changes in speed (and the relative size of the change) salient. The trend vector on a speed tape, however, is a representation of a number, and thus technically not a representation of a trend in the way that a pointer moving across a round dial is. The trend vector, after all, represents a numerical prediction: what the airspeed will be at a fixed number of seconds into the

future (based on the assumption of conditions remaining the same). Another issue is known as display compatibility. To achieve compatibility between the conceptual analog quantity of airspeed in the world and the representation of airspeed in the cockpit, the display of speed should be analog too (and not digital, not just a number). That means that, for purposes of display compatibility, an airspeed scale should be fixed. Moving a pointer across that fixed scale brings out the conceptual analog quantity of airspeed, both in the movement and the position of the pointer. This became the principle of the moving part. The moving part in a compatible representation of airspeed is the pointer, not the scale. Yet in a speed tape, this principle is violated: it is the scale that moves and the pointer that remains fixed. One advantage of a speed tape versus a round dial is its more parsimonious use of cockpit real estate, but for that gain, there are costs in the way it supports the cognitive and visual work of speed monitoring.

More technology, higher performance expectations

New technology is also not neutral when it comes to the performance expectations put on the system. Larry Hirschorn talks about a law of systems development, which is that every system always operates at its capacity.[9] 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 first 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. Not only that, almost without exception, technology did not meet the goal of unencumbering the military personnel operating the equipment. Weapon and support systems often required exceptional human expertise, commitment and endurance. The Gulf War showed that there is a natural synergy between tactics, technology and human factors: effective leaders will exploit every new advance to the limit.

Procedural Adaptations

Why don't people follow the rules? Applying procedures is not simple rule-following. In most complex, dynamic systems, procedures are not sufficiently sensitive to the many subtle variations in context. This is why we need people to apply them. Situations may (and often do) occur where multiple procedures

need to be applied at once, because multiple things are happening at once. But items in these various procedures may contradict one another.

> *There was one case, for example, where a flight crew noticed both smoke and vibrations in their cockpit, and actions to combat either seemed to contradict each other. There was no procedure that told them how to deal with the combination of symptoms. Adaptation was necessary to deal with the situation.*

Applying procedures can thus be a delicate balance between:

- Adapting procedures in the face of either unanticipated complexity or a vulnerability to making other errors. But people may not be entirely successful at adapting the procedure, at least not all of the time. They might then be blamed for not following procedures; for improvising.
- Sticking to procedures and discovering that adapting them would perhaps have been better. People might then be blamed for not being flexible in the face of a need to be so.

As with human error, there are different ways to look at the nature and application of procedures (see Table 4.1). Labels such as "procedural violation" or "non-compliance" are obviously Old View. They are counterproductive and judgmental labels—a form of saying 'human error' all over again, without explaining anything.

Table 4.1 Old and New View interpretations of procedural adaptations

Model 1 (Old View)	**Model 2 (New View)**
Procedures are the best thought-out, safest way to carry out a task	Procedures are resources for action (next to other resources)
Procedure-following is IF-THEN, rule-based behavior	Applying procedures successfully is a substantive, skillful cognitive activity
Safety results from people following procedures	Procedures cannot guarantee safety. Safety comes from people being skillful at judging when and how they apply
Safety improvements come from organizations telling people to follow procedures and enforcing this	Safety improvements come from organizations monitoring and understanding the gap between procedures and practice

Forgetting to arm the ground spoilers (which help braking and dump the wing lift) on a passenger aircraft before landing is a classical error trap. People actually create safety by adapting procedures in a locally pragmatic way. Recall the pilot from Chapter 3, who had figured out that he could safely arm the spoilers four seconds after "gear down" was selected, because the critical time for potential gear compression was over by then. He had refined a practice whereby his hand would go from the gear lever to the spoiler handle slowly enough to cover four seconds—but it would always travel there first. He then had bought himself enough time to devote to subsequent tasks such as selecting landing flaps and capturing the glide slope. This obviously "violates" the original procedure, but the "violation" is actually an investment in safety, the creation of a strategy that help forestall failure.

Of course, Airline Operations Manuals prescribe the standardized techniques that should be used in conducting an approach to landing. In particular, such written guidance prescribes the procedural flows to be used, and checklists to be accomplished, and when. It is possible to write such guidance only if a number of assumptions are made about the environment in which the guidance is to be followed. The most important of these assumptions have been captured by Loukopoulos and colleagues,[10] *and include:*

Assumption 1—*The environment is linear. This assumption says that tasks always follow a prescribed order in a fixed sequence. The successive order in which tasks are to be accomplished is prescribed by operations manuals. The idea is that each item can be completely finished before the next item is taken on, or that a whole checklist can be finished before a new task can be taken on (for example, flying the aircraft down from 1,000ft). That way, tasks and checklist can be done in the prescribed order, and the possibility of changing, or having to change, the order of tasks is mentioned in such guidance only as an afterthought (for example, "minor variations in order are acceptable"). There is no guidance on how to manage such variations or on how to ensure that an entire checklist is indeed accomplished, or that the delayed execution of a checklist does not intrude into concurrent or subsequent tasks. The assumed linearity implies seriality: only one activity is performed at the time, there are no parallel task demands. This of course contradicts the real nature of flying an approach: monitoring various things in the cockpit and coordination with air traffic control always accompanies the execution of checklists in parallel.*

Assumption 2—*The environment is predictable. This assumption says that tasks and events can all be exactly anticipated, both in nature and timing.*

Variations in pacing, timing and sequencing are hardly discussed in available written guidance. Also, the assumption implies that when pilots need to complete a certain task, then all information for completing that task is available unambiguously and in its entirety. There is no mentioning of the need to deploy prospective memory (the memory for accomplishing things yet to do, in the (near) future). A monitoring pilot (as opposed to the flying pilot), for example, has to assume that the other pilot is ready to respond to a checklist when such a checklist is pulled out.

Assumption 3—*The environment is controllable. Another important assumption in the construction of checklists and procedures is that pilots are in control of the environment in which they fly. That is, pilots are assumed to be in control over the timing, pacing and manner of execution of their tasks. This would also mean that sufficient time is available to actually complete the tasks, that pilots can devote their full attention to the task at hand, and that they can execute the tasks in the manner they had anticipated or planned. If other tasks intrude in, for example, checklist reading, or checklist reading intrudes in other tasks, there is no written guidance available on how priority should be given (or not) to the intervening tasks.*

Of course, only very few of the distractions, changes or interruptions that occur in the real world are possible to anticipate in any meaningful way or in any detail in written procedures (such as the operations manual). And not all eventualities and complexities can be accommodated in written guidance. The most effective way to write up tasks to be done is in the sequence in which they normally (or desirably, ideally) are conducted. But that leaves two major problems. First, procedures developed from the perspective of an ideal operating environment (based on the three assumptions above) tend to be brittle, or fragile. That is, their execution, order or completeness easily crumbles under the pressure of real task pacing, concurrent task demands or other complexities or surprises of real-world operations. Second, although pilots do learn to manage concurrent task demands, particularly through experience, they (as do other professional groups) tend to underestimate their own increased vulnerability to error in those circumstances.

Understanding real work on an approach to a busy airport requires understanding of the contrast between the ideal, assumed operating environment and how it fits procedures and checklists, and how these same procedures and checklists need to be interleaved, spliced, paused or paced in real-world operations. In many real-world situations, pilots cannot delay or defer one task long enough to complete another one fully. This means that multiple tasks need to be interleaved

> *(for example, monitoring instrumentation and automation when accomplishing a checklist). Interleaving unpracticed tasks, particularly those that involve novel or unexpected aspects leads often to errors of omission. It is interesting to note that much of the checklist research is focused on what happens if checklist-reading and execution is interrupted, and how various aspects of checklist and procedural design can, as much as possible, insulate checklist behavior from such interruptions and the omissions they engender. Less attention is paid in this research to a case where the execution of the checklist (procedurally demanded as a precondition for landing), and itself complete and uninterrupted, intrudes into other tasks, such as instrument monitoring.*

What if people don't follow the procedures? The difference between Old and New View offers you two ways of looking at the mismatch between procedures and practice:

- As non-compliant behavior, or "violations." People who violate the procedures put themselves above the law. Violations are a sign of deviance. "Violation" is a judgmental label that you impose from the outside. From that position you can see a mismatch between what people do locally and what they are supposed to do, according to your understanding of rules governing that portion of work.
- As compliant behavior. Even though actual performance may mismatch written guidance, people's behavior is typically in compliance with a complex of norms, both written and implicit. Getting the job done may be a mark of expertise against which real performance is scored, by juniors, peers and superiors. This may be more important (even to management) than sticking with all applicable written rules.

While a mismatch between procedures and practice almost always exists, it can grow over time, increasing the gap between how the system was designed (or imagined) and how it actually works. This is called drift:[11],[12] a slow, incremental departure from ideas about how to operate a system. This is what lies behind it:

- Rules that are overdesigned (written for tightly coupled situations, for the worst case) do not match actual work most of the time. In real work, there is slack: time to recover, opportunity to reschedule and get the job done better or more smartly. This mismatch creates an inherently unstable situation that lays the basis for drift.
- Emphasis on local efficiency or cost effectiveness pushes operational people to achieve or prioritize one goal or a limited set of goals (for example, customer service, punctuality, capacity utilization). Such goals

are typically easily measurable (for example, customer satisfaction, on-time performance), whereas it is much more difficult to measure how much is borrowed from safety.

- Past success is taken as guarantee of future safety. Each operational success achieved at incremental distances from the formal, original rules can establish a new norm. From here a subsequent departure is once again only a small incremental step. From the outside, such fine-tuning constitutes incremental experimentation in uncontrolled settings. On the inside, incremental nonconformity is an adaptive response to scarce resources, multiple goals and often competition.
- Departures from the routine become routine. Seen from the inside of people's own work, violations become compliant behavior. They are compliant with the emerging, local ways to accommodate multiple goals important to the organization (maximizing capacity utilization but doing so safely; meeting technical requirements, but also deadlines).

Breaking the rules to get more recruits
Some say cheating needed to fill ranks

New York Times, *May 4 2005*

It was late September when the 21-year old man, fresh from a psychiatric ward, showed up at a US Army recruiting station. The two recruiters there quickly signed him up. Another recruiter said the incident hardly surprised him. He has been bending or breaking enlistment rules for months, he said, hiding police records and medical histories of potential recruits. His commanders have encouraged such deception, he said, because they know there is no other way to meet the Army's recruitment quotas.

"The problem is that no one wants to join," the recruiter said. "We have to play fast and loose with the rules just to get by." Others spoke of concealing mental health histories and police records. They described falsified documents, wallet-size cheat sheets slipped to applicants before the military's aptitude test, and commanding officers who look the other way. And they voiced doubts about the quality of troops destined for combat duty.

Two years ago, a policy was ended that nearly always dismissed serious offenders from recruiting. The Army's commander of recruiting explained how "his shift in thinking was that if an individual was accused of doctoring a high-school diploma, it used to be an open-and-shut case. But now he looks at the person's value to the command first."

Recruiting has always been a difficult job, but the temptation to cut corners is particularly strong today, as deployments in Iraq and Afghanistan have created a desperate need for new soldiers, and as the Army has fallen short of its recruitment goals in recent months. Says one expert: "The more pressure you put on recruiters, the more likely you'll be to find people seeking ways to beat the system." Over the past months, the Army has relaxed its requirements on age and education, a move that may have led recruiters to go easier on applicants.

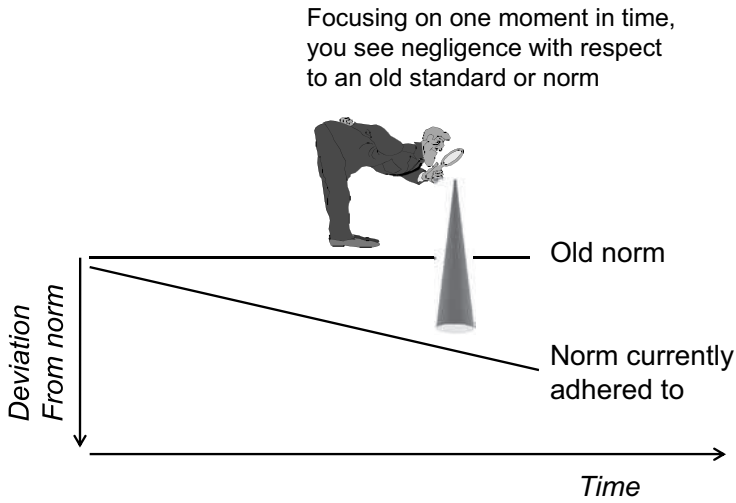


Figure 4.2 At any one moment, behavior that does not live up to some standard may look like complacency or negligence. But deviance may have become the new norm across an entire operation or organization

So when you discover a gap between procedures and practice:

- Recognize that it is often compliance that explains people's behavior: compliance with norms that evolved over time—not deviance. What people were doing was reasonable in the eyes of those on the inside the situation, given the pressures and priorities operating on them and others doing the same work every day.
- Find out what organizational history or pressures exist behind these routine departures from the routine; what other goals help shape the new norms for what is acceptable risk and behavior.

- Understand that the rewards of departures from the routine are probably immediate and tangible: happy customers, happy bosses or commanders, money made, and so forth. The potential risks (how much did people borrow from safety to achieve those goals?) are unclear, unquantifiable or even unknown.
- Realize that continued absence of adverse consequences may confirm people in their beliefs (in their eyes justified!) that their behavior was safe, while also achieving other important system goals.

A major driver behind routine divergence from written guidance is the need to pursue multiple goals simultaneously. Multiple goals mean goal conflicts. In most work, contradictory goals are the rule, not the exception. Any human factors investigation that does not take goal conflicts seriously, does not take human work seriously.

Don't Use "Loss Of Situation Awareness"

> *At a founding conference on situation awareness in the mid-nineties, aviation safety veteran Charles Billings wondered aloud whether the use of "situation awareness" was necessary to explain what causes people to see, or miss, or remember seeing, something. "It's a construct!" he said in a keynote at a foundational conference on situation awareness. "Constructs cannot cause anything!"*[13]

> *Charlie Billings passed away in 2010. By then, "loss of situation awareness" (that is, loss of a construct) had become the favored cause for an epidemic of human performance-related mishaps in aviation and in other settings. Many reports produced by the Safety Boards across the world contain references to a "loss of situation awareness." A meta-analysis of over 300 civil aviation incident reports conducted by human factors researchers indicated that that "loss of situation awareness" causes more incidents when the captain is at the controls than when the first officer is, and that the pilot flying is more likely to lose situation awareness than the pilot not flying.*[14]

Why is it a bad idea to use "loss of situation awareness" when trying to explain human performance difficulties? In short, it is a bad idea because it is simply a new way of saying 'human error.' Here is what it does:

- "Loss of situation awareness" simply says that you now know more than the people who were in the tunnel. It is not an explanation at all.

Instead, it puts you in the position of retrospective outsider—outside the tunnel, judging people for what they did not see but should have seen (which you know because you stand outside, in hindsight).

- "Loss of situation awareness" is based on seventeenth-century ideas about the workings of consciousness, when people believed that the mind is a mirror of the world. Developments in cognitive science have long since overtaken this notion.
- "Loss of situation awareness," because of the seemingly scientific ring to it (as opposed to 'human error') is now being used by others (media, the judiciary) to condemn operators for their failures.

Let's look at these three reasons for not using "loss of situation awareness" in more detail.

"Loss of situation awareness" puts you outside, in hindsight

Hindsight easily slips into our thinking of human performance—even under more modern labels. "Loss of situation awareness" is no more than the difference between what you know now, and what other people knew back then. And then you call it *their* loss.

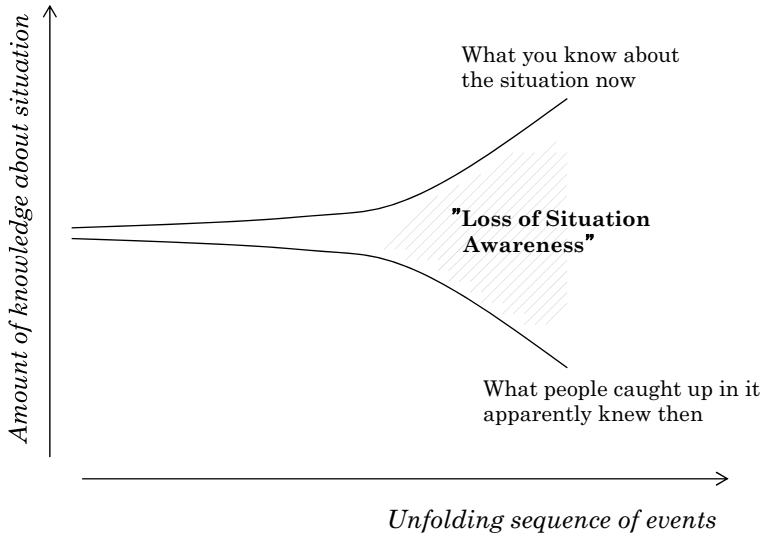


Figure 4.3 "Loss of situation awareness" as the difference between what you now know and what other people knew back then

"Loss of situation awareness" is the difference between what you know now, and what other people knew back then. And then you call it *their* loss.

What this really shows, however, is your ignorance about the situation other people faced. You are not making an effort to go into the tunnel and understand why it made sense for them to look at what they were looking at. That *you now* know what was important means little and it explains nothing. In fact, it really interferes with your understanding of why it made sense for people to do what they did.

"Loss of situation awareness" is based on seventeenth-century thinking

"Loss of situation awareness" is actually based on seventeenth-century ideas about the workings of consciousness.[15] What philosophers and scientists at the time believed was that the mind was like a mirror: a mirror of the world outside. Knowledge of the world is based on correspondence. What is in the mind, or in awareness, needs to correspond to the world. If it doesn't correspond, then that knowledge is imperfect. And the human needs to do more work to make the correspondence perfect.

Such seventeenth-century ideas have made it possible for today's researchers and investigators to claim that they did know what was in the world. And then they can show that the operators did not. In their own words, when claiming that there was a "loss of situation awareness" in experiments or investigations,

> *… there is a "ground truth" against which its accuracy can be assessed (e.g., the objective state of the world or the objective unfolding of events that are predicted).*[16]

This is a pretty arrogant position to take, of course. And a meaningless one. You basically say to an operator after the fact that *you now* know the real state of the world, that you know the ground truth (that is, your correspondence was perfect), but that she or he evidently did not. "Loss of situation awareness" is a judgment you make about the correspondence to your ground truth and their understanding. And if that correspondence is not perfect, then you call it their "loss of situation awareness" or their inaccurate situation awareness.

The biggest problem in this is obviously that *you* get to decide what is the ground truth. Indeed, what *is* the ground truth? The data that you now know were important? But you only know that with knowledge of outcome,

in hindsight. If the operators had known the outcome, they too would have considered that data important. The point is, they did not know the outcome. If they had, they would have done something different. Your job is not to point out (by whatever labels) that you are now smarter than operators were back then. To judge them for never rising to your ground truth. Your job is to go down into the tunnel with them, and to understand their awareness. To understand what made sense to them at the time (without knowing the outcome!) and to explain to yourself and others why.

Developments in cognitive and other sciences have since shown the correspondence idea of knowledge, and this notion of "ground truth" to be unsustainable. Here is the really short version: No person in the world is so privileged as to have access to a "ground truth" against which all other people's understanding can be proven wrong or inaccurate. The way each individual looks at the world is determined by where that individual stands. And that goes for all of us:

- everybody's perspective is unique, determined by the position from which you take it;
- no two people can be in the *exact* same position at the same time, so no two perspectives can ever be the same;
- nobody's perspective on the world can be reduced to someone else's;
- everybody's perspective is subjective;
- nobody has an "objective" perspective, against which other people's "subjective" perspectives can be compared.

As an investigator, in hindsight, and with knowledge of outcome, you might believe that you have a ground truth. But your view is subjective and unique (just like the one of the people whose performance you are examining), and just as constrained (though by other things) as the next person's. Anything else would assume that you can take a view from *nowhere*—an objective view that is entirely free from any influencing background, knowledge or position in the world. There is no such viewpoint. You have to stand somewhere! You have to view things from somewhere. And given that you have to, you might as well try your best to stand on the inside of the tunnel rather than on the outside, in hindsight.

"Loss of situation awareness" causes accidents and puts people in jail

Charlie Billings' remark that situation awareness is a construct, and that constructs don't cause anything, must have gone unheard. Or unheeded. The

popularity of the construct has meant that "loss of situation awareness" now causes plenty of accidents.

> *I learned recently of a criminal court case against an operator who, in the words of the prosecution (the Crown in this case) had "lost situation awareness" and had therefore been criminally negligent in causing an accident that killed two people. The operator was sentenced to years in jail. In another case, the coroner who investigated a friendly fire incident which killed three British soldiers in Afghanistan in 2007, rendered the verdict that the crew of an American fighter jet had lost "situational awareness" and were looking at the wrong village when they dropped the bomb.*[17]

This is no longer just a supposed explanatory use of "loss of situation awareness." Situation awareness, in these cases, represents a duty of care, the professional commitment expected of operators whose actions can influence the lives of others. When coroners or prosecutors (or researchers) demonstrate a "loss of situation awareness" (which is very easy):

- it represents an absence of a duty of care; a breach of the relationship of trust with patients, passengers, colleagues, collateral;
- it represents a failure to live up to a professional commitment;
- it represents a possibly prosecutable crime.

This takes situation awareness beyond the wildest dreams of those who introduced it to the human factors lexicon. Other people can always show that there was more in the world than there was in the mind, because in hindsight anybody can show that. And they can then call that difference the operators "loss of situation awareness." Research might even aid such bastardization of human factors. Consider the very first sentence of a recent article that attempts to introduce one particular model of situation awareness to anesthesia:[18] "Accurate situation awareness (SA) of medical staff is integral for providing optimal performance during the treatment of patients." Just imagine the following exchange that may show up in medical liability, medical indemnity or even criminal negligence cases:

> *Q. Wouldn't you agree, doctor, that accurate situation awareness by medical staff like yourself is integral for providing optimal performance during the treatment of patients? This is what the leading journal in your specialty claims. See, here it says so [counsel points to exhibit].*

A. Uh, I'd have to agree.

Q. Would you say, doctor, that your performance in this case, in which your patient died as a result of the care you provided, was optimal?

A. Uh, we all hoped for a different outcome.

Q. Were you, or were you not aware of the situation that this particular drug X, when used in combination with Y and Z, had produced problems for this patient 18 years before, when she was living in another State?

A. I was not aware of that at the time, no.

Q. Yet you agreed that accurate situation awareness is integral for providing optimal performance during the treatment of patients?

A. … [silence]

Q. No further questions.

I would like to see those in human factors who champion the construct, help defend the operator who is accused (implicitly or explicitly) of losing situation awareness. I do not know whether they, or anybody, can. The most fundamental problem is that situation awareness locks you into that hopelessly old-fashioned seventeenth-century thinking where there is a world and a mind, and the mind is merely the (imperfect) mirror of the world. If we urge people to be less complacent, to try a little harder, then that mirror can become a little less imperfect. The inverse is true too. If people turn out to have an imperfect mental mirror of the world (a "loss of situation awareness"), we know that because the outcome of their actions was bad—and in hindsight we can easily point to the exact few critical elements that were missing from their mental picture. We, or others, can then blame their deficient motivation (their complacency, their violation of the duty of care, their breach of the professional relationship) for this imperfection.

"Loss of situation awareness," like any characterization of 'human error,' raises the difficult question of a norm, or standard (recall the notion of 'ground truth' discussed above). And you might recall this issue from the Preface. But what is the norm, or the standard relative to which the behavior is "erroneous?" What is the awareness norm we put up relative to which someone else's awareness was "lost?" We can only define situation awareness in relation to some target world, or norm, that the operator should have been

aware of. But we cannot use hindsight knowledge of the total situation as that norm, because the operators in question did not have that knowledge. Sure, we can compare our outside/hindsight knowledge of the situation with operators' inside knowledge. Of course we are going to know more, or at least different things, about what turned out to be critical. But that is not a useful norm. It just makes us both arrogant *and* ignorant:

- arrogant because we think we know better (and because of hindsight and outcome knowledge, we indeed just might, but that doesn't mean anything);
- ignorant because we haven't done anything to understand people's unfolding perspective inside the tunnel.

Don't Use "Complacency"

The same problems apply to complacency.[19] Complacency is a huge term, often used to supposedly explain people's lack of attention to something, their gradual sensitization to risk, their non-compliance, their lack of mistrust, their laziness, their "fat-dumb-and-happiness," their lack of chronic unease. Something that explains so much usually explains very little. Which is true for complacency too. Here are the reasons to be suspicious of it when you see it used, and to avoid its use if you can:

- To prove complacency, you first have to show what an optimal sensitivity, attention or commitment would be. Otherwise we cannot judge people to be complacent. They can only be complacent relative to that optimum benchmark. So you have to define that first.
- The problem is, even when using optimal strategies in a complex, dynamic world, people will miss things. This happens because there is always more to pay attention to, to look at, to consider, then there is time or are cognitive resources to do so.[i] To point out, in hindsight, that people missed certain things, only shows how much *you* are in

i When we speak of complacency, human factors researchers Moray and Inagaki point out, we are concerned with monitoring, with sampling visual displays or other representations or indications of process behavior, particularly in automated systems. "The notion of complacency arises from the suggestion that, particularly when the automation being monitored is highly reliable, operators may not merely trust it, but trust it too much, so that they fail to sample (monitor) the variables often enough." Complacency refers to an incorrect strategy that leads to sub-optimal monitoring. Important signals may be missed because of operator complacency, because they have too great a trust in their systems doing the right thing. Such trust tends to grow in the operation of increasingly reliable systems.

the position of retrospective outsider. They were looking at all kinds of things in ways that have developed to be optimal over time. That you don't understand that shows that you have not done enough to put yourself in their shoes, to understand how the world was unfolding inside the tunnel, without knowledge of outcome.

- Accusing others of complacency really throws the charge back onto yourself. Using 'complacency' is an investigative or managerial cop-out. Guess who

---

To be complacent, Moray and Inagaki explain, an operator must be shown to sample a process indicator less often than is optimal, given the dynamics of the source. But how often and when should a display be visually sampled in order that monitoring is "optimal?" Studies that claim to show the existence of complacency are defective, because none have ever rigorously defined optimal behavior in supervisory monitoring. The existence of complacency cannot be proved unless optimal behavior is specified as a benchmark. If signals are missed, then it cannot be claimed that missed signals imply, or are caused by, complacency. More insidiously, the supposed "benchmark" against which actual monitoring behavior is sampled is often defined in hindsight, with knowledge of outcome. Of course with hindsight, it is really easy to show what an optimal sampling rate would have been, because a bad outcome easily shows the sub-optimal. If only people had seen *this* piece of data *then* in the unfolding process, the outcome could have been avoided. This of course means nothing: it is a mere retrospective judgment using a folk model, and not an explanation of anything. As Moray and Inagaki explain:

This is not just a matter of semantics. To claim that an operator missed a signal because of complacency is to have recourse to the classic tactic which all claim to eschew, namely, put the blame on the operator and say that 'human error' was the cause of the problem. A term like "complacency" implies, in everyday language, a character trait that the operator has acquired and which could be changed. To say that a person's "complacency" was the cause of an error is not the same as saying, for example, that a fault was due to limited working memory, a characteristic of the nervous system that is not under voluntary control (p. 362).

Even if optimal monitoring strategies are achievable, would it guarantee that no signals will be missed? Moray and Inagaki demonstrate how even optimal monitoring will often (or almost always) lead to critical signals being missed. In real-world systems, some sources of information are associated with high hazard (for example, airspeed) and optimal monitoring strategies are typically developed relative to that. But they are not the only sources that need to be monitored. There are always trade-offs between sampling different information sources. These are driven by, as signal detection theory explains, payoffs and probabilities. And as signal detection theory can also explain (see the chapter on automation): there is no such thing as an optimal sampling rate to make sure no signals are missed. If that optimum is developed relative to one channel, or one process, or one kind of signal, then immediate sacrifices are made with respect to all others. As Moray and Inagaki dryly observe: "The only way to guarantee to detect all signals is devote attention entirely and continuously to the one process on which the critical signals are expected to appear." This is indeed what human factors researchers Parasuraman et al. (1993) instructed their operators to do in one experiment, and complacency all but disappeared. But, as Moray and Inagaki point out, in real worlds this is rather impractical.

> is being 'complacent' when you use the term in your investigation…! It is saying 'human error' all over again, a 'human error' (complacency) which was the cause of trouble. Remember, the New View asks you to look at 'human error' as a consequence, an effect, a symptom. If you notice behavior you would like to call "complacent," ask yourself what that is the effect, symptom or result of. Dig deeper. And don't say that such behavior is the result of people not paying enough attention, or because they lost situation awareness. Because then you will just go around in circles.

The risk of getting into circular arguments with novel categories such as "loss of situation awareness" or "complacency" is pretty big. Don't let yourself get sucked into these (as John Flach warned over two decades ago):

Using complacency is an investigative cop-out. Guess who is being 'complacent' when you use the term in your investigation!

- Why did you lose situation awareness?
- Because you were complacent.
- How do we know you were complacent?
- Because you lost situation awareness.

Remember, someone else's "loss of situation awareness" is merely the difference between what they knew then versus what you know now. If you turn that into a judgment about what those others also should have known, you easily get into all kinds of trouble. Because why didn't they know? Well, they didn't because they were complacent. Or were they complacent because they didn't know? It doesn't matter, there will surely be a way to deem them guilty anyway.

Avoid the use of these novel ways of saying 'human error' altogether, and you can stay out of trouble. It will also be better for those whose safety you are supposed to manage, or for those whose assessments and actions you are investigating.

Notes

1 Neisser, U. *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W. H. Freeman, 1976.

2 Orasanu, J.M., Martin, L., Davison, J. Cognitive and contextual factors in aviation accidents: Decision errors. In: Salas, E., Klein, G.A., editors. *Applications of naturalistic decision making*. Mahwah, NJ: Lawrence Erlbaum Associates, 1996.

3 Khatwa, R., Helmreich, R.L. Analysis of critical factors during approach and landing in accidents and normal flight. *Flight Safety Digest*. Washington DC: Flight Safety Foundation, 1998:256.

4 NTSB. Runway overrun during landing American Airlines Flight 1420, McDonnell Douglas MD-82, N215AA, 1 June 1999. Washington, DC: National Transportation Safety Board, 2001.

5 Feltovich, P.J., Spiro, R.J., Coulson, R. The nature of conceptual understanding in biomedicine: The deep structure of complex ideas and the development of misconceptions. In: Evans, D., Patel, V., editors. *Cognitive science in medicine: Biomedical modeling*, pp. 19–51. Cambridge, MA: MIT Press, 1989.

6 Sarter, N.B., Woods, D.D., Billings, C. Automation Surprises. In: Salvendy G, editor. *Handbook of human factors/ergonomics,* pp. 1926–1943. New York: Wiley, 1997.

7 Hutchins, E.L. How a cockpit remembers its speeds. *Cognitive Science* 1995;19(3):265–88.

8 Ibid.

9 Cordesman, A.H., Wagner, A.R. *The lessons of modern war, Vol. 4: The Gulf War*. Boulder, CO: Westview Press, 1996.

10 Loukopoulos, L.D., Dismukes, K., Barshi, I. *The multitasking myth: Handling complexity in real-world operations*. Farnham, UK; Burlington, VT: Ashgate, 2009.

11 Dekker, S.W.A. *Drift into failure: From hunting broken components to understanding complex systems*. Farnham, UK: Ashgate, 2011.

12 Snook, S.A. *Friendly fire: The accidental shootdown of US Black Hawks over Northern Iraq*. Princeton, NJ: Princeton University Press, 2000.

13 Billings, C.E. Situation awareness measurement and analysis: A commentary. In: Garland, D.J., Endsley, M.R., editors. *Experimental analysis and measurement of situation awareness*. Daytona Beach, FL: Embry-Riddle Aeronautical University Press, 1996:1–5.

14 Jentsch, F., Barnett, J., Bowers, C.A., Salas, E. Who is flying this plane anyway? What mishaps tell us about crew member role assignment and air crew situation awareness. *Human Factors* 1999;41(1):1–14.

15 Dekker, S.W.A. On the epistemology and ethics of communicating a Cartesian consciousness. *Safety Science* 2013;56(7):96–9.

16 Parasuraman, R., Sheridan, T.B., Wickens, C.D. Situation awareness, mental workload and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making* 2008;2(2):140–60.

17 Bruxelles Sd. Coroner criticises US as he gives 'friendly fire' inquest verdict. *The Times,* 28 April, 2010:1.

18 Schulz, C.M., Endsley, M.R., Kochs, E.F., Gelb, A.W., Wagner, K.J. Situation awareness in anesthesia: Concept and research. *Anesthesiology* 2013;118(3):729–42.

19 Moray, N., Inagaki, T. Attention and complacency. *Theoretical Issues in Ergonomics Science* 2000;1(4):354–65.

5 Understanding Your Accident Model

The kind of safety work you do depends on what you think is the biggest source of risk. That in turn depends on what your accident model is. Accident models offer ideas about what risk is and how you can best contain it. For the last century or more, a lot of scientific literature has appeared about this.[1] This chapter takes you through four accident models (or, perhaps more correctly, families of models). The next chapter considers the role of the safety department. This role, too, is determined in part by your organization's ideas about its sources of risk and the nature of accidents.

Chain of Events

The Chain-of-events model stems from the 1930s. It says that a linear series of errors, failures and violations is necessary to push a system over the edge into breakdown. It is like a line of dominoes, one falling against the other until the last one is down.

> *The idea of the chain of events comes from H.W. Heinrich, whose research into injury causation showed that 88 percent of accidents were caused by workers themselves (this was based on supervisor reports about those accidents, so no wonder there). Heinrich concluded that, "The occurrence of an injury invariably results from a completed sequence of factors—the last one of these being the accident itself. The accident in turn is invariably caused or permitted by the unsafe act of a person and/or a mechanical or physical hazard."*
>
> *Heinrich was the first to articulate the idea of worker "unsafe acts." Note that "unsafe acts" by front-line workers still play a crucial role in the Swiss Cheese model that many use today. The thinking on which it is based is getting on in age—soon a century.*

Table 5.1 These four different accident models provide different explanations of risk, what it is and how it can best be contained

Accident Model	Risk Defined as	Major Threat to Safety	Manage Safety by
Chain of events	Weakest link in the chain	Unreliable humans	Getting rid of weakest link
Barriers (Swiss Cheese)	Accident trajectory not stopped	Weak defenses	Plugging holes
Systems theory	A control problem	Complexity and goal conflicts	Making goals and erosion visible
Drift	Gradual acceptance of lower margins	Being successful	Staying chronically uneasy

Please stop using the triangle

Heinrich, by the way, is also to many the father behind the so-called "triangle:" the idea that there is a proportional relationship between major accidents/fatalities, injuries and incidents, and minor events. For example, in some industries, there are believed to be 600 minor incidents for each fatality. For those who have looked at his work in more detail, you will know that he was working with numbers from insurance and actuarial science—not safety or accident prevention. All the same, the triangle promises the following:

- There are common causes to incidents and accidents (this idea is maintained in chain-of-events, defenses-in-depth and Swiss Cheese models).
- Serious injuries, accidents and fatalities can be avoided by reducing or avoiding minor incidents and safety events.

There are significant empirical problems with either of these. More recent research has shown that the safer your industry becomes, the less applicable the common-cause hypothesis becomes. As you will see in later chapters in *The Field Guide*, incidents in very safe systems are caused by radically different things than accidents or fatalities. The kinds of incidents you have, and that you might be spending energy on avoiding and suppressing, are lousy predictors of the accidents you will suffer. Data from the construction

industry, for example, shows that while frequency rates of minor injuries and incidents may have been declining, the absolute number of fatalities and life-altering injuries has remained steady for years. Also, the fixed ratios between various kinds and levels of incident or accident turn out to be (in large part because of the reason above) a false promise. The ratios, if they can be shown unambiguously at all, vary wildly between places, professions, jobs and industries. And they also vary considerably for different kinds of injuries and accidents.

A continued belief in the triangle can probably make your industry *less* safe. It rocks you to sleep with the lullaby that the risk of major accidents or fatalities is under control as long as you don't show minor injuries, events or incidents.

The promise of fixed ratios and common cause has a harmful effect on our thinking about risk and safety: the false idea that we can control the risk of a major disaster by counting, recording, tabulating and supressing the small stuff.

> *Take the Macondo accident (also known as Deepwater Horizon). Just before the accident, managers were celebrating six years of injury-free performance. According to the triangle, that should have delivered some 3,600 years of accident-free performance (by which time oil in the Gulf of Mexico has probably run dry). The very next day, however, the accident caused the death of 11 people and the biggest oil spill in the history of humanity.*

A continued belief in the triangle can probably make your industry *less* safe. It rocks you to sleep with the lullaby that the risk of major accidents or fatalities is under control as long as you don't show minor injuries, events or incidents. It isn't. Later chapters in *The Field Guide* will give you a deeper and better understanding of this.

Let's get back to the chain of events. Some believe a chain-of-events model is still a good way to describe how things go wrong. The Threat-and-Error Model (or TEM), for example, does this by suggesting that one error or violation typically follows the next, leading to an unsafe aircraft state.[2]

Chain-of-events models have some analytic advantages:

- it is easy to plot the linear trajectory to an outcome and you can make it look convincing;
- you can make suggestions about where to cut the chain, or about which links to take out.

Most people believe that unreliable humans are the weakest link in the chain. This link can be strengthened by more accountability, good professional standards, education and training. And if none of that works, the link could be removed by taking the human out of the loop with automation. The analytic sacrifices the chain-of-events model makes, however, are pretty substantial. For example:

- You have to choose a beginning of the chain. This is an arbitrary choice. Who gets to say, and on the basis of what?
- By putting some events in the chain, you necessarily exclude lots of others that may have impacted the sequence. Who are you to say what goes in and what is left out?
- What if events in the chain happened simultaneously? A linear chain cannot deal with that, as one thing can only follow another and not happen at the same time.
- Causes and effects in the chain are proportional. All the dominoes, after all, are similar in size. This means that small causes cannot lead to big effects.
- Taking one link out of the chain will likely do very little to remove the failure potential (see the example below) because, in complex systems, chains are not as neatly linear as the chain idea proposes.

In an explosive decompression incident, a DC-10 airliner lost part of its passenger floor, and thus all the control cables that ran through it, when a cargo door opened in flight in 1972. The pilot had actually trained himself to fly the plane using only the engines because he had been concerned about exactly such a decompression-caused collapse of the floor. He was able to land the plane safely. Afterwards, he recommended that all DC-10 pilots be trained in engine-only flying techniques. The Safety Board, the regulator and a subcontractor to McDonnell Douglas (the manufacturer of the plane) all recommended changes to the design of the aircraft. McDonnell Douglas, however, attributed the incident to 'human error:' the baggage handler responsible for closing the cargo compartment door had not done his job properly. It was shown to be very difficult to see whether the door had actually latched, but a 'human error' here represented a convenient, and cheap, link in the chain to stop at.

Two years later, another DC-10 crashed near Paris, killing all 346 people onboard. The cargo door had come unlatched again. This again led to an explosive decompression. This time, however, it made the passenger floor

collapse onto the control cables. The jet could not be controlled or steered to a safe landing. It plowed into a forest at very high speed.

Removing one link in the chain, or putting in one barrier to halt the chain, does not acknowledge the complexity of cause in typical safety-critical systems.

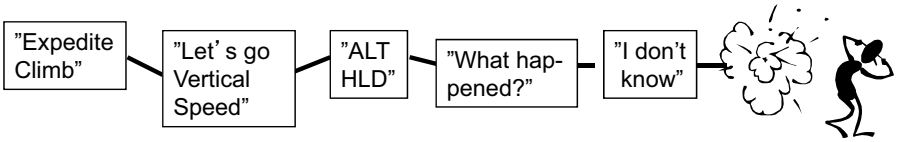


Figure 5.1 Laying out a chain of events, including people's assessments and actions and changes in the process itself

In the sequence-of-events model, a countermeasure could be to no longer use the vertical speed mode when asked to expedite a climb. In other words, make a rule, write a procedure. This puts a barrier between the request to expedite a climb and the use of the vertical speed mode. This is represented in Figure 5.2.

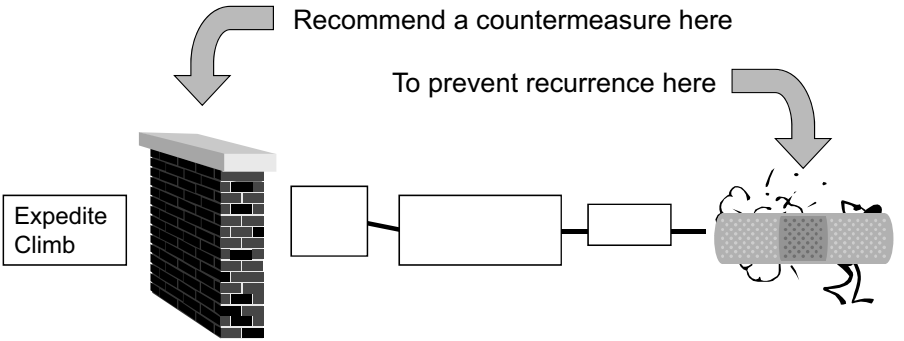


Figure 5.2 We may believe that blocking a known pathway to failure somewhere along the way will prevent all similar mishaps

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. This is why the "Findings as to risk" as used in the Swissair 111 MD-11 report are so powerful. Investigators there highlight factors through which the entire industry is exposed to problems that played a role in their accident.

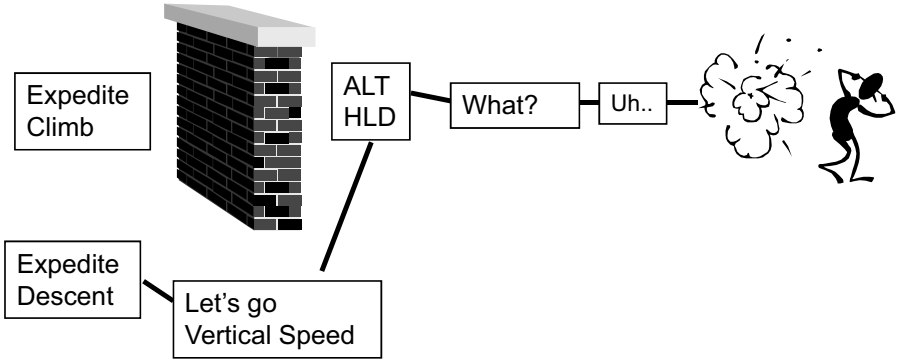


Figure 5.3 Without understanding and addressing the deeper and more subtle vulnerabilities that surround failure, we leave opportunities for recurrence open

Barriers

Barrier models propose that our safety-critical activities are generally well protected against risk. We have both hard and soft defenses (or barriers) in place, consisting of, for example, blast walls, or procedures. A lot needs to go wrong for a failure to happen in these systems.[3] The last line of defense is often thought to be the human operator at the sharp end. The barrier model is inspired by industries where containing (dangerous) energy is the major safety issue, for example process control, oil & gas, nuclear power generation. Large-scale industrial accidents during the late 1970s and 1980s inspired the creation of the model, which sees accidents as an effect of the combination of:

- active errors or "unsafe acts," committed by those on the sharp end of a system; and
- latent errors, or conditions buried inside the organization that lie dormant for a long time but can be triggered in a particular set of circumstances.

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 latent failures that combined with more active problems on a fateful night. For example, instrumentation in the process control room was inadequate: its design had 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 either did not exist or came up short. 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 times 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.

The view that accidents result from long-standing deficiencies is not new, but its popularization with the barrier model in the 1980s and 1990s turned people's attention to upstream factors, away from just frontline operator "errors." The aim is to find out how those "errors" too are a systematic product of managerial actions and organizational conditions. It encourages us to probe organizational contributions to failure and to see 'human error' at the sharp end not as a cause, but an effect. But it still presents serious problems:

- The role given to "unsafe acts" is important in this model—people need to do something wrong or unsafe at the sharp end for all the bad energy to breach the final barriers. This seems to once again invoke 'human error' as final and necessary cause, and the human as one of the weakest links. Fortunately, a member of a prominent accident investigation board told me recently they have abandoned the use of "unsafe acts" altogether, precisely to prevent such misconceptions about the sources of risk.
- Also, the search for latent failures (for the holes or weaknesses in defense layers) can quickly become pointless, as everything can be construed as a possible "resident pathogen." This means that while the model is helpful for finding and categorizing accident precursors or contributors in the rubble *after* a mishap, it is more difficult to make meaningful predictions with it.

I recall talking to the safety manager of an air traffic control organization. He had commissioned a study into "latent pathogens" in his organization, in order to get a better estimate of the fragility (or accident liability) of his company. This should help people estimate or predict how close an organization had come to breakdown, and direct countermeasures to places where they could have most effect.

When the safety manager's team crossed the mark of 837 latent pathogens, it gave up and stopped counting. There was no end in sight. With a bit of fantasy, everything could potentially be a latent pathogen. What exactly was a latent pathogen? Where in practice should you look for one (or rather, not look for one)? Was a latent pathogen a hole in a layer of defense? A lack of a layer of defense? A potential interaction of one layer of defense with another?

The barrier model is, in many ways, still a chain-of-events model.[4] One thing causes the next, in a linear sequence, until all barriers have been breached. A safer system, then, is a stronger one—with better defenses.

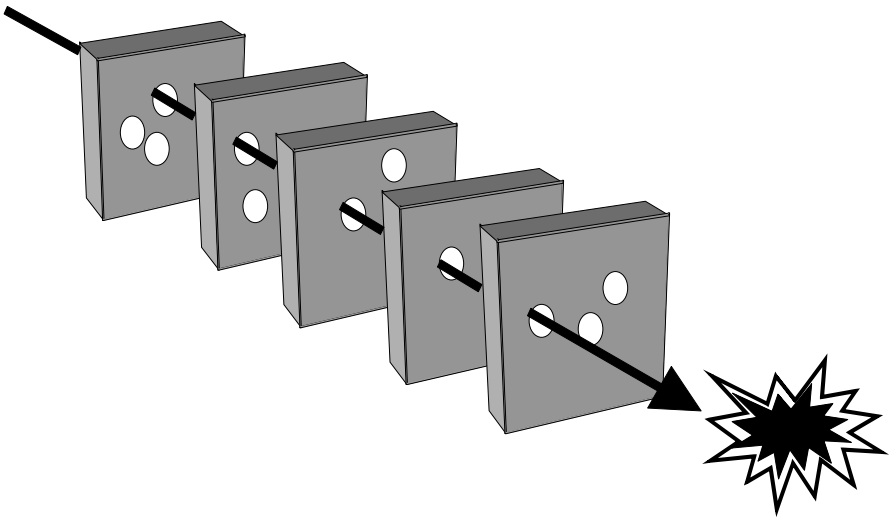


Figure 5.4 The "Swiss Cheese" analogy. Latent and active failures are represented as holes in the layers of defense. These need to line up for an accident to happen (after Reason, 1990)

A barrier model can faithfully explain the last few minutes (or seconds, or perhaps hours, depending on the time constants in your domain) before an accident. One reason for this is that it might make sense, in many domains, to see risk during those final minutes or seconds as energy that is not contained (a drug with a high therapeutic index, for example, or two aircraft that come too close to each other in the sky).

The "physical cause" of the loss of the Space Shuttle Columbia in February 2003 was "a breach in the Thermal Protection System on the leading edge

of the left wing. The breach was initiated by a piece of insulating foam that separated from the left bipod ramp of the External Tank and struck the wing in the vicinity of the lower half of Reinforced Carbon-Carbon panel 8 at 81.9 seconds after launch. During re-entry, this breach in the Thermal Protection System allowed superheated air to penetrate the leading-edge insulation and progressively melt the aluminum structure of the left wing, resulting in a weakening of the structure until increasing aerodynamic forces caused loss of control, failure of the wing, and breakup of the Orbiter."[5]

The barrier, or defenses-in-depth, model, is still a chain-of-events model too. One breach needs to precede the next. Also, the final link in the chain is an "unsafe act" by a frontline operator. A 'human error' in other words.

What is much harder for a barrier model to do is to explain the social, organizational and bureaucratic context that gave rise to weak defenses. It is no longer useful to talk about risk as energy-to-be contained. What about the historical trajectory in an organization along which the nature and extent of risk was renegotiated? Where lots of engineers and analysts were doing their normal daily work, and, in the process, "allowed" acceptable risk to become redefined? This is typical in stories ranging from the Challenger Space Shuttle to Enron to the collapse of sub-prime mortgage lending and Lehman Brothers. That is where models that consider "drift" come in.

Also, in some cases, the assumption that there is a linear, proportional relationship between the layers of defense can be dangerous. It can create new or additional risks:

On one hospital ward, a medication administration double-checking routine was put in place to avoid wrong medications or wrong doses from going into patients. It became policy that two nurses should check the calculations before the drug was administered. This was inspired by the barrier model—more barriers are better then fewer: it makes for a safer system. It turned out, however, that this ward actually began having more medication adverse events. The reason was that the supposed added barrier (the extra nurse who needed to re-calculate) was not independent from the first one. Nurses know each other. They know each other's work. They have seen each other for an entire shift, they may have administered this very medication to this very patient before. This is called the "fallacy of social redundancy:" barriers consisting of people who know each other (and each other's work) are not independent at all. They interact in ways to actually erode both. So applying

one particular accident model at the expense of others can actually serve to increase risk, rather than reduce or contain it.

It is something that systems theory has discovered as well, though in a different way. Let us turn to that now.

Systems Theory

The focus on defenses as the best investment in safety has led in many systems to greater complexity. More defenses mean more engineered systems, more procedures, more paperwork, more people. Many of these have to interact in their work to ensure safety. Aerospace engineers in the 1950s already discovered this, and decided that increased defenses and extra parts could actually be dangerous. As the designers of the Mig-29 (an awesome fighter aircraft) said: "The safest part is the one we could leave off."

- The more complex a system, the more difficult it becomes to control.[6] More complexity typically leads to a system that is difficult to see through—it is opaque, full of unexpected interactions and interconnections. These can play up in ways that human beings cannot keep up with.
- The more complex a system, the more difficult it becomes for people to even know whether they still have adequate control or not.[18] With lots of parts and interdependencies responsible for ensuring safety, it may not be easy to see how some are not living up to expectations, how some may have eroded.

There is nothing inherently complicated about putting a coffee maker aboard a large airplane, but a simple part in a complexly interactive system can have such consequences that Murphy is almost vindicated. On this flight during a very cold winter, passengers had been told there would be no coffee because the drinking water was frozen. Then the flight engineer noticed he could not control cabin pressure. It was later found that frozen drinking water cracked the water tank, heat from ducts to the tail section then melted the ice in the tank, and because of the crack in the tank, and the pressure in it, the newly melted water near the heat source sprayed out. It landed on an outflow valve in the pressurizing system, which allows excess cabin pressure to vent to the outside. The water, which had gone from water to ice and then back to water, now turned to ice because the outside of the air valve is in contact with—50 degrees air outside the plane at 31,000 feet. Ice on the

> *valve built up pressure in the valve and caused it to leak, and the leak made it difficult to … maintain proper cabin pressure.*[7]

The complexity of many systems, and of the technology they employ, can also mean that one kind of safety needs to be considered against another. All designs, as is said, are compromises. Here is an example of a goal trade-off that results from the nature of safety in different contexts:

> *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. But 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 less risk along the way than a single structure filled with rocket fuel.*[8]

Systems models focus on the whole, not the parts (like the accident models above do). The interesting properties of systems (the ones that give rise to system accidents) can only be studied and understood when you treat them in their entirety. System models build on two fundamental ideas:[9]

- **Emergence**: Safety is an emergent property that arises when system components and processes interact with each other and their environment. Safety can be determined only by seeing how parts or processes work together in a larger system;
- **Control** imposes constraints on the degrees of freedom (for example, through procedures, design requirements) of components, so as to control their interaction. Such control is based on prior (and possibly false) ideas about how components and processes interact. Control can be imperfect and even erode over time.

System accidents result not from component failures, but from an erosion of control of safety-related constraints on the development, design and operation of the system. People and organizations may have thought that their control

was adequate—based on their knowledge, goals and focus of attention at the time. But systems are not static designs; they are dynamic processes, continually adapting to achieve goals in a changing environment. What was adequate control before may now have eroded constraints and pushed the system closer to the edge, operating with smaller margins. Such erosion typically happens because these systems are trying to pursue multiple (conflicting) goals at the same time.

> *A commuter train derailed near Osaka, Japan in 2005, crashing into an apartment building. The accident killed at least 56 passengers and injured 440. The train had been running 90 seconds late, and was suspected of doing twice the speed limit for that section of the line, when five of its seven cars jumped the track.*
>
> *Investigators focused on the inexperienced 23-year-old driver, who had been reprimanded for overshooting a platform once before during his 11 months on the job, and may have been speeding to make up for the delay. The Japanese Prime Minister called on officials to "respond firmly in order to prevent future accidents."*
>
> *Trains in Japan, however, are known for running with such precision that riders traveling on unfamiliar lines can plot complex itineraries, secure in the knowledge that they will not miss connections because of delays. Accuracy and speed are key ingredients in running a successful commuter system in dense urban areas. The driver may have felt he was expressing the preferences and expectations of his entire society.*[i]

Production pressure and goal conflicts are the essence of most operational systems. Though safety is a (stated) priority, these systems do not exist to be safe. They exist to provide a service or product, to achieve economic gain, to maximize capacity utilization. But still they have to be safe. One starting point, then, for understanding a driver behind routine deviations, is to look deeper into these goal interactions, these basic incompatibilities in what people need to strive for in their work. If you want to understand 'human error,' you need to find out how people themselves view these conflicts from inside their operational reality, and how this contrasts with other views of the same activities (for example, management, regulator, public).

> *NASA's "Faster, Better, Cheaper" organizational philosophy in the late 1990s epitomized how multiple, contradictory goals are simultaneously present*

i *International Herald Tribune*, 26 April 2005.

and active in complex systems. The loss of the Mars Climate Orbiter and the Mars Polar Lander in 1999 were ascribed in large part to the irreconcilability of the three goals (faster and better and cheaper), which drove down the cost of launches, made for shorter, aggressive mission schedules, eroded personnel skills and peer interaction, limited time, reduced the workforce, and lowered the level of checks and balances normally found. People argued that NASA should pick any two from the three goals. Faster and cheaper would not mean better. Better and cheaper would mean slower. Faster and better would be more expensive. Such reduction, however, obscures the actual reality facing operational personnel in safety-critical settings. These people are there to pursue all three goals simultaneously—to make it faster, better and cheaper.

To explain failure, system models do not need a component to break or a human to err. In fact, they do not have to rely on anything "going wrong," or anything being out of the ordinary. Even though an accident happened, nothing really might have gone wrong—in the sense that nothing that happened was out of the ordinary. Accidents, in other words, are typically the by-product of the normal functioning of the system, not the result of something breaking down or failing inside of that system. In that sense:

- There is not much difference (if at all) between studying a successful or a failed system, since the organizational processes that lead to both are very similar.
- System models, in other words, are hardly "accident models." They are models of a system's or organization's normal functioning. It is this normal functioning that contains the seeds both of success and failure.

A study of flight dispatchers illustrates a basic dilemma. Would bad weather hit a major hub airport or not? What should the dispatchers do with all the airplanes en route? Safety (by making aircraft divert widely around the weather) would be a pursuit that "tolerates a false alarm but deplores a miss." In other words, if safety is the major goal, then making all the airplanes divert even if the weather would not end up at the hub (a false alarm) is much better than not making them divert and sending them headlong into bad weather (a miss). Efficiency, on the other hand, severely discourages the false alarm, whereas it can actually deal with a miss.[10]

System models do not rely on linear cause–effect relationships to explain how factors interact or relate to one another. This means that they can stay closer

to the complexity and goal conflicts behind system success and failure. It also means that they, as models, are more complex.

> *Some organizations pass on their goal conflicts to individual practitioners quite openly. An aviation publication commented on one of those operators (a new airline called Excel, flying from England to holiday destinations): "As part of its punctuality drive, Excel has introduced a bonus scheme to give employees a bonus should they reach the agreed target for the year. The aim of this is to focus everyone's attention on keeping the aircraft on schedule."*[ii]

Such plain acknowledgement of goal priorities, however, is not common. Most important goal conflicts are never made so explicit, arising rather from multiple irreconcilable directives from different levels and sources, from subtle and tacit pressures, from management or customer reactions to particular trade-offs.

> *For example, the operating manual of a large international airline opens by stating that "(1) our flights shall be safe; (2) our flights shall be punctual; (3) our customers will find value for money." In other words, the flights have to be better, faster and cheaper. All at the same time.*

The management of goal conflicts under uncertainty gets pushed down into local operating units—control rooms, cockpits, operating theaters and the like. There the conflicts are to be negotiated and resolved in the form of thousands of little and larger daily decisions and trade-offs. These are no longer decisions and trade-offs made by the organization, but by individual operators or crews. What they accept as risky or normal will shift over time:

- as a result of pressures and expectations put on them by the organization;
- and as a result of continued success, even under those pressures.

This takes us to another family of models about risk and accidents: Drift into failure.

Drift

> *In 1966, a portion of a coal mine tip (unusable material) near Aberfan, South Wales, slid down into the village and engulfed its school. It killed 144 people, including 116 children. Investigating events leading up to the*

ii *Airliner World*, 2001, p. 79.

disaster, Barry Turner found that this period was characterized by events that went unnoticed or were disregarded. Basically, they were at odds with taken-for-granted beliefs about hazards and norms for their avoidance. For the first time in safety science, Barry Turner turned not to the managerial and administrative processes of the organization. In these, he believed, we could find the origins of the discrepancy between growing risk and the sustained belief that it was under control.

This gave birth in the 1970s to "man-made disaster theory." It was really the first to see accidents as the result of a drift into failure, and to focus on the organizational blunt end to explain that drift.[11] This theory was a call to understand accidents not as sudden phenomena where energy was not contained, but as phenomena over time, where people and whole organizations subtly changed their idea of what was risky in the first place.

Telling yourself to look for holes and failures makes you forget that, to the people working there, their organization is not typically a brittle collection of porous layers, full of people committing failures on a daily basis. To them, it is a place where normal people come to do normal work. If all you look for is the holes, the abnormal, the failed, the broken, you will have difficulty understanding why things are normal to them; why what they do makes sense to them and their colleagues and superiors (and often even regulators). To understand that, the incremental nature of drift is important. While the end result of drift may seem a very large departure from originally accepted routines, each step along the way is typically only a small increment (or decrement)

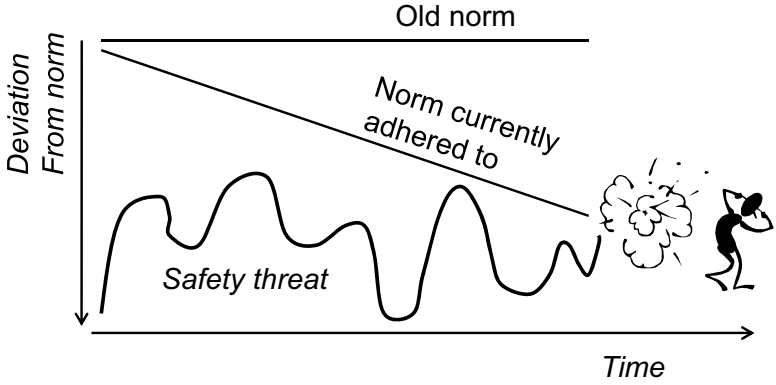


Figure 5.5 Murphy's law is wrong. What can go wrong usually goes right, and then we draw the wrong conclusion: that it will go right again and again, even if we borrow a little more from our safety margins

from what was accepted previously. As such, the small departure from the previous routine does not normally concern people that much, especially if adverse consequences to not occur. Doing what you do today (which could go wrong but did not) does not mean you will get away with it tomorrow.

We will look at safety culture in more detail in later chapters. But for now, if we say that a safety culture is a culture that allows the boss to hear bad news, then the hard problem here is to decide what is bad news.

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 the expected norm, for example, even if we 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. From the inside, drift may become invisible.

A safety culture is a culture that allows the boss to hear bad news.

Diane Vaughan called this process of drift "normalization of deviance." A group's construction of risk can persist even in the face of continued (and worsening) signals of potential danger. This can go on until something goes wrong, which (as Turner would have predicted) reveals the gap between the presence of risk and how it

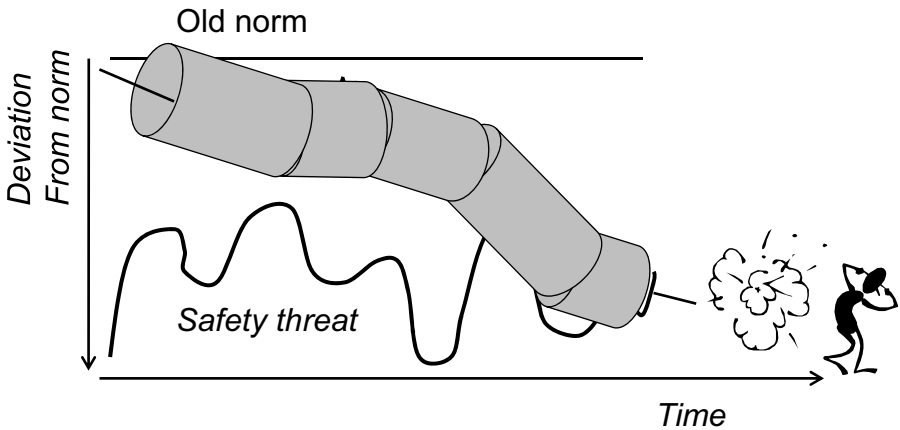


Figure 5.6 Drift into failure is hard to see inside of the tunnel. Each step away from the norm is only small and offers success on other important measures (for example, time, cost, efficiency, customer satisfaction)

was believed to be under control. Small departures from an earlier established norm are often not worth remarking or reporting on. Such incrementalism contributes to normalization. It helps the organization learn the wrong thing:[12]

> *Experience generates information that enables people to fine-tune their work: fine-tuning compensates for discovered problems and dangers, removes redundancy, eliminates unnecessary expense, and expands capacities. Experience often enables people to operate a socio-technical system for much lower cost or to obtain much greater output than the initial design assumed (p. 333).*

This is why high-reliability organizations (HRO) deal with risk by remaining chronically uneasy. Otherwise they may:

- be overconfident in past results. In a dynamic, complex system, past results are not a good guide for future safety;
- suppress minority viewpoints, even though these could show more clearly where risk brews among the messy details of daily practice;
- give of priority to acute performance expectations or production pressures.

To guard against these drift-inducing impulses, HRO suggests you stay curious, open-minded, complexly sensitized, inviting of doubt and ambivalent toward the past.[13] People in HROs are described, ideally, as skeptical, wary and suspicious of quiet periods. Continued success, after all:[14]

> *…breeds confidence and fantasy. When an organization succeeds, its managers usually attribute success to themselves or at least to their organization, rather than to luck. The organization's members grow more confident of their own abilities, of their manager's skills, and of their organization's existing programs and procedures. They trust the procedures to keep them apprised of developing problems, in the belief that these procedures focus on the most important events and ignore the least significant ones.*

Weick and colleagues echoed this two decades later:[15]

> *Success narrows perceptions, changes attitudes, reinforces a single way of doing business, breeds overconfidence in the adequacy of current practices, and reduces the acceptance of opposing points of view.*

Chronic unease, and its basis in your understanding of actual practice, requires leadership involvement. It includes managerial and supervisory visibility at the sharp end and a real interest in what goes on there.

Resilience and Safety I versus Safety II

If the major risk is success—that is, things going right—instead of failures, then why do we spend the majority of our safety resources on investigating what goes wrong? This is exactly the question Erik Hollnagel raises. Managing safety on the basis of incidents is only one way—and in a sense a very limited way. It focuses, after all, on the few occasional times when things go (almost) wrong, rather than on the many times that things go right. It is a reactive, lagging kind of safety management that might turn into firefighting instead of a proactive, coordinated improvement of overall organizational performance. This is why Erik Hollnagel has made the distinction between:[16]

- **Safety I**. Safety is the absence of negative events. A system is safe if there are no incidents or accidents. The purpose of safety management is to ensure that as little as possible goes wrong. The focus is on negative events and reducing their severity and number. This often translates into trying to reduce the variability and diversity of people's behavior—to constrain them and get them to adhere to standards.
- **Safety II**. Safety is the presence of positive capacities, capabilities and competencies that make things go right. This is resilience: the ability of a system to adjust its functioning before, during or after changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions. This needs to translate into using the variability and diversity of people's behavior—to be able to respond to a range of circumstances, both known and new.

In Safety II, the focus of safety management is to ensure that as many things as possible go right. This requires safety people as well as operational managers to understand how the system succeeds under varying conditions. Humans are not a source of weakness or unreliability, but a source of flexibility and resilience. An organization committed to Safety II will probably not want to stop learning from the few things that go wrong. But it will also want to understand the messy details of how people normally get their jobs done—despite the goal conflicts, design flaws, communication difficulties and more. In an already safe system, after all, this is where the risk for bigger failure brews: not in the occasional incident, but in people's daily, normal work with all the little hiccups and adjustments that do not get formally reported. As Erik Hollnagel suggested, if we want to understand the latter, we have to focus on frequent events,

Is safety making sure those few things don't go wrong, or that as many things as possible go right?

not necessarily severe ones. What presents difficulty on a daily basis, the often-encountered workarounds and frustrations? Such things might indeed be better predictors of system safety and risk than your formally reported incidents.[17] So when you do your next safe work observations, for example, do not walk around telling people how they are supposed to work. Try to understand why they work the way they do, and why it is, or seems, normal for them at the time.

Notes

1 Woods, D.D., Dekker, S.W.A., Cook, R.I., Johannesen, L.J., Sarter, N.B. *Behind human error*. Aldershot, UK: Ashgate, 2010.

2 Jensen, R.S., editor. Models of threat, error and response in flight operations. Tenth International Symposium on Aviation Psychology; 1999; Columbus, OH. The Ohio State University.

3 Reason, J.T. *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate, 1997.

4 Hollnagel, E. *Barriers and accident prevention*. Aldershot, UK: Ashgate, 2004.

5 CAIB. Report Volume 1, August 2003. Washington, DC: Columbia Accident Investigation Board, 2003.

6 Perrow, C. *Normal accidents: Living with high-risk technologies*. New York: Basic Books, 1984.

7 Ibid.

8 Vaughan, D. *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. Chicago: University of Chicago Press, 1996.

9 Leveson, N.G. *Engineering a safer world: Systems thinking applied to safety*. Cambridge, MA: MIT Press, 2012.

10 Smith, K. Incompatible goals, uncertain information and conflicting incentives: The dispatch dilemma. *Human Factors and Aerospace Safety* 2001;1:361–80.

11 Turner, B.A. *Man-made disasters*. London: Wykeham Publications, 1978.

12 Starbuck, W.H., Milliken, F.J. Challenger: Fine-tuning the odds until something breaks. *The Journal of Management Studies* 1988;25(4):319–41.

13 Weick, K.E. The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly* 1993;38(4):628–52.

14 Turner, *op. cit.*

15 Weick, K.E., Sutcliffe, K.M. *Managing the unexpected: Resilient performance in an age of uncertainty* (Second Ed.) San Francisco: Jossey-Bass, 2007.

16 Hollnagel, E. *Safety I and Safety II: The past and future of safety management*. Farnham, UK: Ashgate, 2014.

17 Dekker, S.W.A. *Drift into failure: From hunting broken components to understanding complex systems*. Farnham, UK: Ashgate, 2011.

This page has been left blank intentionally

6 Creating an Effective Safety Department

Given the different models about risk and accidents, what should your safety department do, and how and where should it be organized? The answers tend to differ across industries. Until not so long ago, for example, many hospitals would have no meaningful safety function. Even the airline I flew for had a safety department of exactly one person. Now, in many ways that could be a good thing: it prevents the bureaucratization of safety, or the sense that safety is the problem of the safety department and that everybody else can just focus on operations. But it can also have downsides: a lack of safety intelligence in the organization, or an inability to respond adequately to incidents, for instance. So let's look at the other side. I know a construction company that has hundreds of safety people globally. Oil majors may have even more. Many people in those safety departments, however, may be kept busy with record-keeping and managing various liabilities and accountabilities, and may have little sense of what it takes to do an operational job safely. Indeed, the correlation between the number of safety people and these organizations' and industries' safety records seems pretty loose.

Staff or Line?

Safety is made and broken on and in the line. So perhaps that is where responsibility for it should lie as well—with the managers, supervisors and workers who get to work at or near the sharp end, who are in daily contact with the safety-critical processes and the hazardous technologies. Yet many industries have turned large parts of their organizations' safety work over to a staff function. This has advantages, for sure. There is some distance from the hubbub of daily pressures, some neutrality, some more resources that can be devoted to analysis and reflection. But there are risks too. Safety as a staff function can get pushed into:[1]

- being a mere arms-length tabulator of largely irrelevant or unusable data. For example, you may find yourself gathering numbers and producing bar charts about events or incidents every month that nobody seems to use;

- compiling a trail of paperwork whose only function is to show compliance with regulations or safety targets set by the organization;
- being a cheerleader for past safety records and re-mouthing how everything was better before (for example, before deregulation, or before the new management);
- being a cost center whose only visible role is to occasionally slow down production;
- getting excluded from organizational activities and decisions that affect the trade-off across production and safety goals;
- delivering systemic safety recommendations while line management actually wants to put responsibility on the front-line people involved.

Granted, some organizations may *like* their safety departments to be in these roles. It means that the safety people are not going to ask difficult questions. They will not interfere with management work. They can conveniently be kept out of line decisions about which they may otherwise have bothersome opinions on.

I remember working with one organization whose safety department was regarded solely as an intelligence gathering operation. Their role was to provide management with safety information (targets for how much safety information was even set up), and that was it: a mere bottom-up supplier. At the same time, management typically came down hard on operators who had been involved in incidents, often without waiting for more extensive investigation in what had gone wrong or why.

While the safety department looked on from the sideline, management went about squelching the very sources from which it owed its safety intelligence. With punitive measures meted out after each incident, operators became less inclined to report problems or events. This deprived the safety department of the data they needed to fulfill their targets.

So while management held up its one hand to receive safety data, it slapped operators around with the other. Nobody was apt to report about safety issues. The rather powerless, passive role given to the safety department was part of the problem, and the solution was to make it a much more active contributor to top-down management work—helping craft responses to incidents, debriefing people involved in them, and participating in other management decisions that affected trade-offs between safety and efficiency.

Especially when production pressures increase, or when economic margins are small and losses just around the corner, safety departments may need to speak up even more. Yet it is also precisely during these times that the organization might be less inclined to listen.

Escaping the passive safety-supplier trap

So how can you escape this role? Suppose that your organization or industry has located its safety work firmly in a staff function. Here is what you should expect in that case:

- Significant and independent resources (both human and monetary). These resources must be independent of production or financial fluctuations for two reasons. First, safety monitoring and vigilance may actually have to go up when economic cycles go down. So you may need the resources most when the organization can least afford them. Second, a safety department may contain part-time operators. These have to be shielded from production ebbs and flows, because safety concerns often accelerate when production pressures go up (production pressures that could simultaneously rob the safety department of its people).
- A constructive involvement in management activities and decisions that affect trade-offs between safety and efficiency, as well as involvement in managerial actions in the wake of incidents. If you remain on the sideline as a supposed "impartial" department, you will see the basis for your work be whittled away. You must be ready to do battle for your data and the safety concerns they feed and represent. This often requires direct access to relevant organizational decision-making levels without having to pass through various levels of non-practice-oriented middle management.
- An end to weekly or monthly or quarterly targets. Safety is not something that the safety department produces, so targets make little sense. They also favor quantitative representations (bar charts) of supposed safety issues, rather than qualitative intelligence on what is going on—when it is going on (not because the week or month is up). If your organization has incentives associated with meeting targets like these, you will quickly find that operational areas in the organization will adjust their reporting to help meet those targets. They do not want to look bad in the eyes of their superiors or peers, after all.
- A continued grounding in operational reality. Having only full-time safety people can make your department less effective, as they can lose

their idea (or never had one) of what it is to operate at the sharp end. They probably have no sensitivity to what the shifting sources of risk and resilience out there may be. If you do involve practitioners, make sure you do not let one professional group (for example, doctors, plant operator and pilots—versus nurses, maintenance technicians and flight attendants) dominate your department, as this will skew your safety intelligence and ways of dealing with it.

- That said, just being a practitioner (or having once been one) does not in itself qualify people to be members of, or have a strong say in, a safety department. You should expect staff members to want to be educated in safety management, incident/accident investigation, report writing, presentation and so forth. Without this, you can quickly get surrounded by happy amateurs whose well-meaning efforts contribute little to the quality and credibility of the safety function.

And here is what you have to show in return:

- A sensitivity to legitimate organizational concerns about production pressures and economic constraints. Showing that you care only about safety can quickly make you irrelevant, as no organization exists just to be safe. They are there to attain other goals, and you are an inextricable part of that. Without achieving production goals for example, everybody may as well pack up and go home. Even the safety people, as soon there will be no more operation to monitor.
- Useable safety intelligence, as opposed to distant tabulations of statistical data. Show, for example, how safety margins may have been eroding over time, and where the gap lies between what causes problems in the operation and what management seems to believe. Go beyond statistics and come up with persuasive intelligence, for instance by using an "incident of the month" that you select for the point you want to get across.

A concerned outsider who knows the inside

Independent of how you get your safety intelligence, you should strive for a role as a concerned outsider who understands the inside. To see how safety margins may be eroding over time, you have to monitor the organization's understanding of itself. After all, without knowing it, the organization may have chosen to operate closer to the boundaries of safe performance than it would actually like to. The organization will not see this itself. It might interpret the inevitable signs of operating close to the boundaries in localized,

unthreatening ways (for example, inattention, complacency of operators, or any other 'human error'). You then have to show the organization how it may be misleading itself into believing that it has a narrow 'human error' problem. You have to show that this is actually the effect of problems deeper inside the operation, created by the organization itself. You may find fertile ground for your efforts: organizations can occasionally feel hopeless and helpless when 'human errors' just don't go away, however many reprimands, rules or regulations they issue.

Tame Your Safety Bureaucracy

Old View safety tends to locate a lot of the responsibility for safety in processes, protocols, procedures, databases and paperwork. And it typically makes a staff department responsible for administering all of it. This has grown dramatically over the past 30 years or so, though it still differs per industry.

> *One of the airlines I flew for, with a fleet of scores of passenger aircraft (both props and jets), had a safety department consisting of exactly one person. He was both safety manager and the department's entire personnel roll. Because of long experience in the industry, he was intimately familiar with many of the messy details of airline life, though he was not a pilot himself. His office was located right next to the crew room at the airline's main base, behind the little pantry. The door was open whenever he was there, which was almost always. Pilots would indeed walk in and out to discuss things with him, right there and then. He had the ear of the airline's senior managers and could easily raise issues directly with them when they'd come to his attention. He would also get back to the individual pilots who had brought the issues to him. In the years that I flew for this airline, the pilots might have received one or two explicit safety notices that flagged a particular issue or incident that had shown to be risky. More paperwork was not put out by this department, even though managers would have seen various statistics in regular updates.*
>
> *A few years later, I was working with a chemical plant that had roughly the same number of employees in total, though only a small core consisted of operators who directly interfaced with the safety-critical chemical processes at the plant. The safety department was up to 20 people strong, many of whom had not been in the core of the plant—recently, or ever. The safety manager was said to be in need of a map to find his way on the plant site, and particularly to find his way to the process operators, many of whom he had not spoken (or listened) to in years. The safety department was famous for*

cranking out reams of paperwork and bulleted lists. Most of it was directed at the plant's various managers, and consisted of data that was only a distant abstraction from what life looked like on the inside.

The growth of this machinery of safety—the measurement, the procedure-writing, the assessment and evaluation, the surveillance—is due in part to increasing regulation. In fact, some would say there is an over-regulation of safety. Many organizations today are accountable to a series of regulators, all of whom need their data sliced and parcelled up in particular ways. This bureaucratic accountability demand in turn demands more bureaucracy. Organizations are in need of additional safety people not so much to manage safety, but to feed the various bureaucracies that need their data. Such safety bureaucracies, however, tend to have some negative consequences:

Safety has increasingly morphed from operational value into bureaucratic accountability. Those concerned with safety are more and more removed—organizationally, culturally, psychologically—from those who do safety-critical work at the sharp end.

- Safety bureaucracies can sometimes institutionalize and confirm the Old View. They do this, for example, by counting and tabulation of negatives (incidents, non-compliance): the 'human errors' and violations of those at the sharp end.
- Incentive structures around the absence of negatives can get people to suppress bad news, or to fudge injury or incident numbers. This hampers learning and honesty.
- Safety bureaucracies are often organized around lagging indicators: measuring that which has already been. The predictive value of most lagging indicators is known to be poor.
- Safety bureaucracies tend to value technical expertise less than they value protocol and compliance. This can disempower not only the expert operators who do safety-critical work, but their supervisors and middle management as well. People may no longer feel as able or empowered to think for themselves, because first the safety people need to come in and have their say. This can stifle innovation. It can also erode problem ownership, with safety people in turn lamenting the "lack of engagement" with what they see as "safety." Of course, rather than inspiring people to take responsibility for safety, such arrangements seem to encourage people to shift that responsibility onto others.

- The gap between how an organization or operation works and how the bureaucracy believes it works can grow. If people at the sharp end come up with worker-led innovations, then these can quickly go underground if they are non-compliant with current procedure or protocol. This means that wider adoption or adaptation of such innovations is limited. Sources of future efficiency, safety and competitiveness are thus silenced. Instead of diversity and innovation, safety bureaucracies tend to prefer predictability and standardization.
- Paperwork gets in the way of talking and listening to people. Time spent filling in spreadsheets or making bullet lists is not time spent on the floor at the sharp end. Supervisors' or managers' time to communicate with their operational staff is often compromised by the bureaucratic demands of the safety organization. Meetings, email and other paperwork demands get in the way of being a supervisor or manager to those they are responsible for.
- There is a self-fulfilling nature about safety bureaucracy. Bureaucratic accountability is demanded because of bureaucratic accountability; paperwork begets paperwork; non-operational positions grow more non-operational positions.
- This is known as bureaucratic entrepreneurism. Fear of the consequences of curtailing a safety function is combined with a promise of future useful work and reminders of past successes. This is what most bureaucracies do (not just those associated with safety), as it helps continue their existence.

There are suggestions in research and from mishaps that growing your safety bureaucracy actually increases your risk of an accident. The more that safety processes and protocols are developed or enforced bureaucratically by those who are at a distance from the operation, the more they become "fantasy documents." Fantasy documents bear no relation to actual work or actual operational expertise.[2] An organization may have become so adept at generating such documents that it gets in the way of managing the risks that actually need managing.

> *In 2008, for example, two years before the Macondo well blowout, BP had identified what it called priority gaps in its Gulf of Mexico operations. The first of these was that there were "too many risk processes going on", which had collectively become "too complicated and cumbersome to effectively manage."*[3]

Recall from above that the bureaucratization of safety helps create what Vaughan called "structural secrecy." This is a by-product of the cultural,

organizational, physical and psychological separation between operations on the one hand, and safety regulators, departments and bureaucracies on the other. Under such conditions, critical information may not cross organizational boundaries. Once firmly committed to its existing processes and procedures, a safety bureaucracy may not know what it really needs to learn from the operation, and may not have ways of dealing with such knowledge if it did.

Having a large safety bureaucracy could actually increase the probability of an accident. A bureaucracy can produce "fantasy documents" and demand compliance with cumbersome processes while not dealing with real sources of risk.

Safety as Responsibility Down, Not Accountability Up

A focus on safety systems and procedural compliance, on surveillance and monitoring, and on achieving low numbers of negative events, can shift the very meaning of safety in an organization. Instead of seeing safety as an ethical responsibility for those who do the dirty and dangerous work, it can become a bureaucratic accountability that is managed upward in the organization. Those above need to be shown good numbers, numbers that do not get you in trouble, that do not make you look bad.

A few years ago, I heard of a woman who was slightly injured at work. She told her supervisor, showed the injury, and went to see the doctor that same afternoon. While in the waiting room, she got a call from school. Her son had fallen ill and been sent home. After her appointment and having her gash cleaned and glued by a nurse, she rushed home to take care of her boy. She later informed her supervisor.

News of the injury made its way to the company's safety manager. He was horrified. Not necessarily because of the injury or the employee's fate, but because he had been on a "winning streak." Next to the entrance of the plant, a sign announced that the company had been without injury for 297 days. 300 had been within reach! The number would have looked so good. It would have made him look so good.

The day after the incident, the safety manager went to see the supervisor before committing, with lead in his shoes, to change the sign to 0. Zero

days since the last injury. Then he learnt that the employee had gone home after the incident and doctor's visit. It was a gift from heaven. He called the HR manager and together they resolved to generously give the woman the previous afternoon off. It was no longer a Loss-Time Injury (LTI). The woman had simply gone home to take care of her child. He also called the clinic. Because no suturing was done and no doctor needed to touch the woman, this was not a medical treatment in the strict definition held by the company. So no Medical-Treatment Injury (MTI) either. The safety manager could breathe again.

A few days later, the sign next to the entrance proudly showed 300.

The "bureaucratic creep" that happens to safety can be explained in part by the systems of accountability and liability that surround safety nowadays. Many organizations, for example, have invested in safety management systems. These exist in principle to ensure that the structures and conditions are in place to enhance safety (more about this in the next chapter). Managers and regulators can take a look at what is in such a management system and (hopefully) get a good idea of how effective their management or regulation of safety is.

But safety management systems can sometimes become liability management systems if their chief role is to prove that management did something about a safety problem. That "something" can include telling everybody else to try harder, to watch out more carefully. It may be putting up posters exhorting people to do or not do certain things. It may be closing out on an investigation by recommending a new policy or stricter compliance with an existing procedure. Those are not really investments in safety, as much as they are investments in showing that you did something about the problem. That is when safety is once again an accountability up, rather than a responsibility down.

Table 6.1 shows a contrast in how you can organize safety—according to Old View and New View principles. In the Old View, power for what to say or do (or to decide what is safe or unsafe) is in the hands of a boss or safety manager, and the intelligence for those decisions is supplied by a staff organization (based on lots of quantifiable data). Safety is thought to be achieved through rules and compliance. These are rules which largely make it impossible (or sanctionable) for people to do the wrong thing. Safety is built on

Safety management systems can become liability management systems when their main role is to prove that management did something about a problem (even if that meant telling everybody else to try harder).

process and bureaucracy, with a strong emphasis on fixed rules, gathering and analyzing data, showing numbers and bullet points up the organizational hierarchy in order to meet bureaucratic accountability expectations and demands. Predictability and standardization—that is, one best method of accomplishing a task—is thought to equate with safe outcomes.

A New View of safety would replace or at least complement this, at times, with an alternative set of principles:

- power to decide in New View safety lies with experts, not bureaucrats;
- and is driven by insight and context-rich experience from the line, rather than by rule and regulation governed by a staff;
- creating safety is about giving people who do safety-critical work the room and possibility to do the right thing. This means giving them not only the discretionary space for decision making, but also providing them with error-tolerant and error-resistant designs, workable procedures and the possibility to focus on the job rather than on bureaucratic accountabilities;
- work, and its safety, are not just governed by process and rules, but adjusted by mutual coordination;
- and innovation and better solutions are understood to be viable only if there is a willingness to embrace diversity and occasional, safe-to-fail non-compliance.

Table 6.1 How to organize safety according to the Old View and New View

Old View Safety	New View Safety
Whoever is boss or safety manager, gets to say	Whoever is expert and knows the work, gets to say
Dominated by staff	Driven by line
Guided by rules and compliance	Guided by insight and context
Make it impossible for people to do the wrong thing	Give people space and possibility to do the right thing
Governed by process and bureaucracy	Adjusted by mutual coordination
Strives for predictability and standardization	Strives for diversity and innovation
Safety as accountability that is managed upward	Safety as a responsibility that is managed downward

This does not mean that a staff safety department has no role to play in New View safety. It does. In fact, research shows that operational experts are not always a trustworthy source of information about what is safe and not safe either.[4] Operational experts have a lot of experience in meeting their local conditions and expectations, in managing the goal conflicts and resource limitations imposed by their organization. They may be so successful with it, however, that they, too, can miss potentially dangerous side effects of the way they have adapted their work to meet those constraints. The complexity and large organizations, particularly if they involve webs of contractors and sub-contractors, can render risk invisible to experts too. Reflecting on the Challenger Launch decision, Jensen[5] describes it as such:

> *We should not expect the experts to intervene, nor should we believe that they always know what they are doing. Often they have no idea, having been blinded to the situation in which they are involved. These days, it is not unusual for engineers and scientists working within systems to be so specialized that they have long given up trying to understand the system as a whole, with all its technical, political, financial and social aspects (p. 368).*

So this is not an either–or question. Operational expertise is indispensible for getting a sense of the messy details of what it means to get the job done under pressure, goal conflicts and resource limitations. But an ability to step back from those messy details and view the complex organizational web, through which risk travels in perhaps unexpected ways, is necessary too. It is unlikely that the same group of people can do it all: a constructive interplay or conversation between line and staff, between experts and others is probably necessary.

Safety and the Line Organization

That said, let's look at the creation and assurance of safety in a line organization. Responsibility for safety in a local product, process or technology is foremost up to the line, as that is where people work with it every day. Operating safety-critical processes or technologies raises all kinds of questions, on a daily basis, that need to be tackled by the line organization—by managers and supervisors. You would think that this, indeed, is part of their jobs. Though it may be one of those parts that get lost in the machinery and bureaucracy of safety for which managers and supervisors get held accountable.

Recent research shows that active involvement in safety concerns by line management has various positive implications.[6] Positive perceptions of

leadership involvement in daily work operations not only increases worker willingness to comply, it also lays the basis for the sort of relationship that allows workers to share "bad news" with the boss, to tell them about safety concerns which may have implications for production or other organizational goals. Interestingly, and this is consistent with long-standing psychological findings about motivation, the topic of discussion or interest should not necessarily be safety. When the line gets involved, the topic should be people's *work*, not (just) their safety. Focusing on safety in line discussions can quickly become a paternalistic exercise. Other people, who just come to visit, supposedly know better what is safe and unsafe than the workers who do the job every day.

If all you do is look after people's safety, they may feel treated as children. If you look after people's work, they will likely feel treated as colleagues.

What do you do out in the field?

Suppose that you are in the line, and you want to see safety as an ethical responsibility. What do you do when you are out in the field? What do you look for and talk about? Looking at people's work in the way that a safety bureaucracy wants you to look at it, is unlikely to yield interesting insights. After all (recall this from the concluding sentences of the previous chapter), that sort of "job observation" is driven by pre-conceived notions of the "one right way" to work. It is motivated by making it impossible for people to do the wrong thing, rather than exploring the various ways to make it possible and legitimate for them to do the right thing. Rather than inviting bottom-up input about how work can be done, it imposes a top-down view of how work should be done. People might feel watched and judged rather than taken seriously and supported. Also, such job observations can become empty box-ticking exercises toward the end of the month, when a bureaucratic target of "x number of safe job observations" needs to be achieved and reported up the line.

This means that the line organization may need some help, some coaching, in how to approach such questions about people's work. And the safety bureaucracy needs to be kept in check while you do this. Even with support or coaching in place, there is no substitute for the authentic presence of line supervisors and managers. Authentic in this case means:

- that managers and supervisors have an inquisitive mindset;
- that they really want to learn about the work that is going on, and what makes that work frustrating or dodgy;

- that they can genuinely listen;
- that they have no ego in it: they accept being proven wrong about how they thought work should be done;
- that they do not stand at the ready with their judgments about how the work should be done instead, or how it does not meet various written requirements;
- that they don't worry, as their first priority, about their own liability or how to reduce it. Because, in any case: if supervisors or managers are held accountable for something going wrong in the workplace, their defense is not likely to hold merely on having told their people to follow applicable rules and procedures. Being able to prove a genuine interest in understanding and improving people's working conditions is likely a much better defense.

Engagement with the questions above does not necessarily need to happen during ongoing work. Debriefings afterward are possible in some settings too (they are routine in aviation, or the military). Debriefings can yield important insights into the workarounds, the improvisations, the innovations and frustrations associated with getting the job done that day. Make sure that relevant people are present who can do something with what is heard during the debriefings (and who can separate the wheat from the chaff). Such debriefings may require a neutral facilitator. This can help when people otherwise believe that their stories might get them in trouble. Of course, that they would feel that way is data in itself—data about the organization and its (lack of) a "just culture" (recall the end of Chapter 1).

Autonomy, mastery, purpose

"How do we get people engaged?" is a common question that gets asked by safety people. Perhaps it is the wrong question to ask. Perhaps we should turn it on its head:

- the issue is not that operational people are not engaged in safety work;
- it is that safety people are not engaged in operational work.

As was said early on in *The Field Guide*, people do not come to work to do a bad job. They typically come to work to do a good job, or at least a passable job, a normal job. And a safe job. Let's start with the assumption that such a motivation exists already: they, after all, want to go home safely too. Engaging people with an external drive "to be safe" through posters, reminders, stickers and campaigns denies or downplays the existence or relevance or power of

this internal motivation. Research shows clearly that people's motivation is supported by autonomy, mastery and purpose.[7] What does that mean?

- **Autonomy.** Most people have an inner drive to explore ways of doing things, and of doing them better or more efficiently. If, however, their job is put together by someone else who tells them exactly how to do something (and non-compliance is discouraged or sanctioned), such inner drive is cut off. Of course, such exploration may lead to less safe ways of doing things (though certainly not always). Where it does, this offers excellent opportunities for conversation and joint agreements about how best to work. Meaningful feedback, choice over how to do things, and real control over various aspects of their work, are all ways to affirm people's sense of autonomy.
- **Mastery.** Most people want to get better at doing things. A sense of progress in capabilities is a very strong inner motivator. Tasks given to people don't have to match their capabilities exactly (in fact, that probably renders people rather inert), but should be given with space and support that allows improvement, growth and an eventual sense of mastery of the task. This can make even routine work highly motivating.
- **Purpose.** Connecting the task to a cause larger than it, or the person doing that task, is a third inner motivator. Such purpose, with people jointly caring about and having a stake in the outcome, typically motivates them to overcome obstacles, and find new ways of tackling problems that may even lie outside their area of responsibility. A sense of purpose gets people to do things in the service of something larger than themselves.

High Performance Teams

When you look at this sort of research on motivation, it is remarkable how much of safety work is organized around, and based on, compliance. People think that the motivation to work safely needs to be imposed from the outside, and enforced with sticks and carrots. But is that how high-performance teams work too? There is, from sporting to executive coaching, a whole literature on high-performance teams. What makes them so good is not that those who run them have really good sticks and carrots. The attributes of high-performance teams are much more like this:[8]

- building trust, with a comfort about being vulnerable and honest with each other when it comes to weaknesses or mistakes;
- comfortable with what is known as constructive conflict, a willingness to engage in passionate dialogue about what matters to the team. There

is no hesitation to disagree, challenge and question—all in the spirit of finding the best answer or solution for that context;
- a decision process where people can participate and which they feel is something they have contributed to. Even if the outcome is not what they might have wanted, they still agreed to the process, and so will be more ready to offer the buy-in that the team needs;
- shared accountability after having committed to decisions and standards of performance. The team leader does not have to be the primary source of such accountability, peers do it instead. Such accountability is typically forward-looking, not backward-looking;
- a focus on results that allows individual agendas and needs to be set aside.

How many safety organizations do you recognize in this? And how many operational teams in your organization are engaged with safety along the lines of such high-performance commitments above? I guess very few. So there is room to develop and improve!

How do you motivate people to be safe when nothing has happened in a long time?

When your organization or industry is already quite safe, your people may have little familiarity with actual failures. They may go through operational life without anything "interesting" happening. As a result, as high-reliability theory might say, they forget to be afraid. They lose their chronic unease. Some might even say that your people become "complacent" (but don't say that: go back to Chapter 4 to see why not).

> *Such was the life of many of the Boeing captains I flew with. Plying pretty standard routes down to the Mediterranean and back up, flying an aircraft whose basic design went back decades, and flying in well-controlled radar airspace all the time, there was simply not much that was unpredictable any longer. Of course, even this aircraft recently suffered a big, surprising accident. It turned out that its automation was designed in ways that pilots were never told about or trained for. This particular aspect of its design was not in the handbooks or training materials for pilots. It led to an automation surprise that any crew could have been susceptible to: an unsuspected bleeding away of airspeed when low on final approach, leading to a stall and crash. The accident killed nine people and completely destroyed the aircraft. It revealed to the 737 community the kind of booby trap that might lurk in an otherwise well-known ultra-*

safe system. Telling people to be chronically uneasy, or exhorting them to not become complacent, does very little to forestall the risk of having such an event.

Invest time and effort in understanding the following sorts of things about their work:

- What is the *gap* between written guidance and actual work? Why is that gap there? The gap is not evidence of "violations" (that would simply blame your operators). Rather, it is evidence that your organization may be miscalibrated about how people have learned to cope with the messy details of work. These are the messy details that cannot be anticipated in standard rules or covered by pre-ordained solutions. By studying the gap, you learn a lot about your own (mis-) understanding of people's work.

To take responsibility for safety on the line, you should first and foremost look at people's work, more than (just) at people's safety.

- What does it take to get the job done on a daily basis? What are the "workarounds," innovations or improvisations that people have to engage in in order to meet the various demands imposed on them?
- What are the daily "frustrations" that people encounter in getting a piece of machinery, or technology, or even a team of people (for example, contractors), to work the way they expect?
- What do your people believe is "dodgy" about the operation? Ask them that question directly, and you may get some surprising results.
- What do your people have to do to "finish the design" of the tools and technologies that the organization has given them to work with? Finishing the design may be obvious from little post-it notes with reminders for particular switches or settings, or more "advanced" jury-rigged solutions (like an upside-down paper coffee cup on the flap handle of the 60-million dollar jet I flew, so as to not forget to set the flaps under certain circumstances). Such finishing the design can be a marker of resilience: people adapt their tools and technologies to forestall or contain the risks they know about. But it can also be a pointer to places where your system may be more brittle than you think.
- How often do your people have to say to each other: "here's how to make it work" when they discuss a particular technology or portion of your operation? What is the informal teaching and "coaching" that is going on in order to make that happen?

While the "physical cause" of the Columbia Space Shuttle accident can be traced by a sequence-of-events model, the organizational processes behind it can't. Even epidemiological models fall short. The challenge is to explain why people gradually come to see deviant behavior or system performance as normal or acceptable, not judge it deviant from the outside and hindsight.

The accident focused attention on the maintenance work that was done on the Shuttle's external fuel tank, revealing the differential pressures of having to be safe and getting the job done (better, but also faster and cheaper). A mechanic working for the contractor, whose task it was to apply the insulating foam to the external fuel tank, testified that it took just a couple of weeks to learn how to get the job done, thereby pleasing upper management and meeting production schedules. An older worker soon showed him how he could mix the base chemicals of the foam in a cup and brush it over scratches and gouges in the insulation, without reporting the repair.

The mechanic found himself doing this hundreds of times, each time without filling out the required paperwork. Scratches and gouges that were brushed over with the mixture from the cup basically did not exist as far as the organization was concerned. And those that did not exist could not hold up the production schedule for the external fuel tanks. Inspectors often did not check. A company program that once had paid workers hundreds of dollars for finding defects had been watered down, virtually inverted by incentives for getting the job done now.

These are not the sorts of things that show up in your incident-reporting system. In other words, they will likely not become part of your organization's system of bureaucratic accountability. They don't show up there because people see these things as part and parcel of their daily work. They are not judged to be incidents. And so they are not worth reporting in the formal sense. However, these are the sorts of things that, surprisingly, do show up as causal or contributory in the few big accidents that still happen, even to very safe industries and organizations.

Notes

1 Woods, D.D. How to design a safety organization: Test case for resilience engineering. In: Hollnagel, E., Woods, D.D., Leveson, N.G., editors. *Resilience Engineering: Concepts and precepts*. Aldershot, UK: Ashgate, 2006:296–306.

2 Clarke, L., Perrow, C. Prosaic organizational failure. *American Behavioral Scientist* 1996;39(8):1040–57.

3 Elkind, P., Whitford, D. BP: An accident waiting to happen. *Fortune*. New York: Fortune Features, 2011:1–14.

4 Amalberti, R. *Navigating safety: Necessary compromises and trade-offs—theory and practice*. Heidelberg: Springer, 2013.

5 Jensen, C. *No downlink: A dramatic narrative about the Challenger accident and our time, First Ed*. New York: Farrar, Straus, Giroux, 1996.

6 Dahl, O., Olsen, E. Safety compliance on offshore platforms: A multi-sample survey on the role of perceived leadership involvement and work climate. *Safety Science* 2013;54(1):17–26.

7 Pink, D.H. *Drive: The surprising truth about what motivates us*. New York: Riverhead Books, 2009.

8 Lencioni. *Overcoming the five dysfunctions of a team*. San Francisco, CA: Jossey-Bass, 2005.

7 Building a Safety Culture

Just as there are two ways to look at 'human error,' there are two ways to look at safety culture. The research on safety culture (and by now, there is a lot) has been saying this for a while.[1] Depending on what view you take, you will either believe or reject the idea that you can "build" a safety culture. Or that a "safety culture" is separable and recognizable as such at all. The two views are:

- the interpretivist view;
- the functionalist view.

Don't get put off by the big words, though. The two views are summarized in Table 7.1 overleaf. You will probably recognize what your organization believes about "safety culture."

Whichever you believe, or whether you sit somewhere in the middle, it is always wise to not overestimate your control over "culture." In fact, safety culture has become so popular over the past two decades, that it now covers pretty much anything. For example, "culture is the way we do things around here" excludes nothing. Everything is culture. But if everything is culture, then how can you control it, or even influence it? Where do you start? If it is "the way we do things around here," it covers your and your people's behaviors, attitudes, technologies and tools, operating assumptions, social relationships, communication patterns, professional and national backgrounds, roles, jobs and job tasks, protocols and much more. No wonder that it has become increasingly popular to blame a deficient safety culture for accidents. It covers everything. And it makes the work of an investigator really easy. Say "safety culture," and you've probably covered whatever it was that "caused" the accident.

No wonder we blame "safety culture" for more and more accidents. By blaming "safety culture," we pretty much cover everything. But it explains little.

Terms that try to explain so much, end up explaining little. That does not mean, however, that "culture" does not play a role in how an organization

Table 7.1 The interpretivist and functionalist views of safety culture

Interpretivist View	Functionalist View
At home in culture studies, like sociology, anthropology	At home in management studies, organizational psychology, engineering
Sees culture as something an organization does	Sees culture as something an organization has
Culture is complex and emerges from interactions between people	Culture can be reduced to the attitudes and behaviors of individual people
Culture can only be influenced, by what people anywhere in it do and how that interacts with others	Culture can be controlled. It can be imposed, changed, taken away, replaced, typically from the top down
Studies culture with qualitative methods such as observations, interviews, discussions, document study	Studies culture with quantitative methods such as surveys, measurements, questionnaires
Takes the "emic" or inside-out perspective	Takes the "etic" or outside-in perspective
Assumes a diversity of perspectives and ideas about safety	Assumes a homogeneity of views and attitudes ("vision zero," "safety first")
Leads to little other than more studying of culture	Leads to safety campaigns, behavior modification programs, posters
Typically accused of being non-pragmatic, no control over culture	Typically accused of being overly pragmatic, myth of control

relates to safety. Or that you can't do anything to influence the "culture" of your organization. In fact, you are probably influencing that culture very much today, every day. As are other people. Simply by what you, and they, are saying and doing right now.

Influencing Your Culture Toward the New View

If you want to influence your organization's culture toward a New View of safety, what are your choices? What languages should you use, what preferences and priorities should you express? Take a look at Table 7.2 opposite. It shows the basic contrast.

Table 7.2 Old View Safety and New View Safety

Old View Safety	New View Safety
People seen as a problem to control	People seen as a resource to harness
Focus on people's attitudes and behavior	Focus on people's working conditions
Safety defined as absence of negative events (incident/injury free)	Safety defined as presence of positive capacities to make things go right
Whoever is boss or safety manager, gets to say	Whoever is expert and knows the work, gets to say
Dominated by staff	Driven by line
Guided by rules and compliance	Guided by insight and context
Make it impossible for people to do the wrong thing	Give people space and possibility to do the right thing
Governed by process and bureaucracy	Adjusted by mutual coordination
Strives for predictability and standardization	Strives for diversity and innovation
Safety as accountability that is managed upward	Safety as a responsibility that is managed downward

People as a Problem to Control

The Old View, which sees 'human error' as the cause of trouble, tends to locate the safety problem in people. People are the problem that needs controlling. Recall the quote from the first chapter: "*people don't always do what they are supposed to do*. Some have negative attitudes to safety which adversely affect their behaviors…" This makes the target for safety intervention pretty obvious—people. Change people's attitudes, change their behaviors, get rid of those who don't want to play by your rules.

Resist behavioral safety interventions

Behavior-based safety programs typically claim that around 85 percent of worker injuries and incidents are caused by people's own "unsafe acts."

In other words, 'human error' is the main cause of trouble. People are the problem. That is Old View thinking. Recall the safety manager from the preface, whose company had concluded that the most common causes of injuries and incidents were worker carelessness, lack of awareness, disregard for safety procedures and distraction. Again, 'human error' becomes the target of any safety intervention.

Behavior-based safety programs focus on workers' "unsafe acts." They see 'human error' as the cause of trouble and place responsibility for safety on workers themselves.

Behavior-based safety programs focus attention on undesired behaviors by workers, and place the responsibility for safety on the workers themselves.[2,a] Remember that Heinrich, in the 1930s, concluded that 88 percent of workplace accidents were caused by workers' unsafe acts. He reached this conclusion on the basis of supervisors' accident reports. Of course—if you are a supervisor, you might want to say that you and the work environment you helped create are both blameless. Instead, the worker is the problem to control.

---

a A neo-liberal trend towards what is known as worker "responsibilization" in many Western and other countries seems to coincide with the restructuring and intensification of work under pressures of resource constraints and competition. This trend is aimed at helping workplaces become more competitive and productive, and the adverse health and safety impacts are increasingly attributed to workers' own behaviors rather than how work is organized or resourced. A recent study by David Gray in Canada shows how workers are assigned ever more responsibility for their own safety at work and are held accountable, judged and sanctioned as individual instigators of trouble—not collective recipients of it. In some places and industries, workers themselves are increasingly blamed (sanctioned, ticketed) for safety violations, with over two-thirds of all citations handed out by workplace safety inspectors directed at workers or immediate supervisors. This is a departure from the gradual moral and juridical enshrining of the notion of *employers* as safety offenders (that is, blaming systems, organizations and sometimes the people behind them) during the twentieth century. Now, individual responsibility once again falls heavier on workers who are "instructed to become prudent subjects who must 'practice legal responsibility'" Gray concludes. Workers, in other words, are enticed (by winning their "hearts and minds") to do the right thing: pay attention, wear protective equipment, ensure machine guarding, use a lifting device, ask questions, speak up. And if they don't, Gray says, "The failure to practise individual responsibility in the face of workplace dangers is often used to explain why workers who perform unsafe jobs become injured."

> *At one factory that had implemented a behavioral safety program, a union representative asked workers during shift meetings to raise their hands if they were afraid to report injuries. About half of 150 workers raised their hands. Worried that some workers feared even raising their hand in response to the question, the union representative asked a subsequent group to write "yes" on a piece of paper if they were afraid to report injuries. Seventy percent indicated they were afraid to report. Asked why, they said that "we know that we will face an inquisition," or "we would be humiliated," or "we might be blamed for the incident or injury."*[3]

Behavior-based safety interventions are based, like Heinrich, on a psychology that was dominant in the 1930s: behaviorism. They target behaviors by observations and by systems of rewards and sanctions. The theory is pretty simple (as it was in behaviorism): those who work safely and carefully do not cause incidents, and they should be rewarded. Those who work carelessly do cause incidents and they should be sanctioned. Behaviorism, as a school of psychology, was not interested in what happened inside the minds of people. It saw the mind as a black box. It was best left closed, because looking into it would only lead to unverifiable speculation. Similarly, behavior-based safety programs do not typically ask what goes on in people's minds. So here are a few good reasons to be weary of such interventions:

- Behavioral safety interventions are not always interested in finding out *why* people do things unsafely. They do not typically ask why it makes sense for people to do what they do. This leaves unexplored the organizational context of goal conflicts, adaptations and the various ways in which people have learned to cope with the contradictions and complexities of real work.
- Behavior-based safety programs do not ask *what* is responsible for the creation of risk. They ask *who* is responsible. And they already have the answer: the worker. Behavior-based safety programs ascribe responsibility for undesirable behavior, as well as responsibility for changing it, to the worker—not the conditions surrounding their work.
- A climate of fear and intimidation sometimes accompanies the implementation of behavior-based safety programs. Some companies have cash prizes for those who do not report incidents or injuries and offer bonus payments to managers whose departments report low numbers of negative events. Others disincentivize reporting of incidents and injuries by taking away perquisites (loss of overtime opportunities) or stigmatization like mandatory drug testing after an incident.

Consider the example of a food warehouse, where 150 workers load and unload trucks, lift boxes, drive fork trucks and move pallets. Each month that no one reports an injury, all workers receive prizes, such as $50 gift certificates. If someone reports an injury, no prizes are given that month. Management then added a new element to this safety incentive program: if a worker reported an injury, not only would co-workers forgo monthly prizes but the injured worker had to wear a fluorescent yellow vest for a week. This was easy to rationalize, of course. The company could say that it was concerned for the worker's safety, and that the vest was there to protect her or him better. But the vest identified the worker as the safety problem. Worse, it stigmatized the worker, alerting to colleagues that he/she lost everybody their prizes.[4]

If you are concerned about safety initiatives that suppress reporting, you are not alone. The US Government Accountability Office, or GAO, recently studied these issues in the US and asked whether some safety incentive programs and other workplace safety policies may actually discourage workers' reporting of injuries and illnesses.[5] It found the following:

Little research exists on the effect of workplace safety incentive programs and other workplace safety policies on workers' reporting of injuries and illnesses, but several experts identified a link between certain types of programs and policies and reporting. Experts and industry officials suggest that rate-based programs may discourage reporting of injuries and illnesses and reported that certain workplace polices, such as post-incident drug and alcohol testing, may discourage workers from reporting injuries and illnesses. Researchers and workplace safety experts also noted that how safety is managed in the workplace, including employer practices such as fostering open communication about safety issues, may encourage reporting of injuries and illnesses.

Behavioral safety interventions typically focus on *who* is responsible, not on *what* is responsible for creating risk.

Behavioral safety programs can harm social cohesion and workplace solidarity. When workers lose prizes if a co-worker reports an injury, peer pressure comes into play.

I learned of an oil installation where the "prize" for a particularly long injury-free period had gone from a six-pack of beers to a cooler to hold them, up to

a leisure boat to hold the cooler. A few days before the period was about to be expire and the boat awarded, a new worker suffered an injury that could not be hidden. She reportedly was almost lynched by her co-workers.

Safety Defined as the Absence of Negatives

Recall from Chapter 5 that we often think that safety means not having any incidents or accidents. That safety is the absence of negatives. Behavioral safety interventions are organized around that very idea. In fact, any initiative that tries to suppress the number of reported incidents is organized around that idea. The fewer incidents and injuries, the safer the organization looks. At least that is the way it looks to the insurer, the clients, the regulator, the managers, the board. And perhaps even to the workers. Of course we all want an absence of negative events. But is it smart to define safety as the absence of negatives? Let's look at that here.

Question your organization's "vision zero"

Defining safety as the absence of negatives is expressed most obviously by "zero visions." These are quite popular in a number of industries. Of course, a zero vision is a great commitment to make—to yourself, your organization, your employees. It is a commitment that may get you to do certain things, and *not* do certain other things, on a daily basis.

Initiatives that try to suppress the number of reported incidents see safety as an absence of negatives. These might actually harm safety in the long run.

But don't mistake a commitment for a statistical probability. That you are committed to having zero incidents or accidents, does not mean that you or your organization will. The world is too complex and dynamic and unpredictable for that. Here are some reasons to take another close look at the vision zero of your own organization—if indeed it has committed to something like that. These are not reasons to reject the vision, but to have more constructive conversations about it:

- Zero vision is defined by its dependent variable: the outcome. It is not defined by its control variables: the safety inputs that everyone in the organization makes. This means that a commitment to zero often leaves many people in the dark about what it is that they are supposed to do. It also encourages numbers games: a manipulation of the dependent variable.

- Zero vision implies that everything is preventable, otherwise 'zero' would be a nonsensical ambition. But if everything is preventable, then everything needs to be investigated and remedies found. This can be a waste of the limited resources your organization has available for investigations.
- Zero visions can stigmatize incidents and those involved in them. They suggest that 'human error' is the source of trouble and the target for intervention. "Errors" can become equated with moral lapses, failures of character.
- There is no support in the safety research literature that a zero vision is at all achievable. All accident theories from the past decades more or less acknowledge that a world of zero bad consequences is out of the question.
- A focus on zero can actually lead to a blindness to real risk. Most organizations which have suffered big calamities over the past decades had exemplary performance on incidents and injuries. Before they blew stuff up and killed scores of their employees, their numbers of minor negative events were really low (or even zero for some time). But while counting and tabulating largely irrelevant low-consequence events, the definition of acceptable engineering risk had been eroding under their noses. They were evidently counting what *could be* counted, not *what* counted.

Defining zero by its dependent variable

As explained in the first bullet above, zero vision has got things upside-down. It tells managers to manipulate a dependent variable. Let's look at that in a bit more detail. If we want to make investments in safety, we normally specify the kinds of things that we want engineers, experts, managers, directors, supervisors and workers to do to organize work, communicate about it, write standards for it. We focus on what they need to manipulate, or on the manipulated or independent variable, in other words. Outcomes, measured in terms of incidents or accidents, or in terms of indicators of resilience, then are what they are. In retrospect—and the study of past accidents is often what drives theorizing on safety—outcomes can be traced back to manipulated variables, whether validly or not. Zero vision turns all of this on its head. Managers are expected to manipulate a dependent variable—this is an oxymoron. Manipulating a dependent variable is something that science considers to be either impossible or unethical. And that is what zero vision can inflict as well. With a focus on the dependent variable—in terms of how bonuses are paid, contracts are awarded, promotions are earned—manipulation of the dependent variable

(which is, after all, a variable that literally depends on a lot of things *not* under one's control) becomes a logical response. Honesty can suffer, as can learning. And indeed, in the longer run, safety itself can be the victim.

Also, defining a goal by its dependent variable tends to leave people in an organization in the dark about what to do (which variables to manipulate) to get to that goal. Workers can become skeptical about zero sloganeering without evidence of tangible change in local resources or practices. It is easily seen as leadership double-speak. A recent survey of 16,000 workers revealed wide-spread cynicism in the face of zero vision.[6] Not only is it in itself unable to practically engage workers, there is nothing actionable (no manipulated variables) in a mere call to zero that they can identify and work with.

Suggesting that everything is preventable

Investigative resources can be misspent as a result of a zero vision. If zero is assumed to be achievable, then everything is preventable. And if everything is preventable, everything needs to be investigated. This includes minor sprains, rolled ankles and paper cuts. If your organization doesn't investigate, it can even have direct legal implications. A documented organizational commitment to zero harm can lead the regulator or judiciary to claim that if the organization and its managers and directors really believed that all harm was preventable, then such prevention was reasonable practicable. They are liable if harm occurs after all, since they or their workers must have failed to take all reasonably practicable steps to prevent it.

A commitment to zero suggests that we must manipulate a dependent variable. This has got things upside-down.

Stigmatizing incidents and those involved in them

A zero vision can also encourage the stigmatization of workers involved in an incident.

> *One of the more enduring instances of this can be found in medicine. The 2000 Institute of Medicine report[7] was accompanied by a political call to action to obtain a 50 percent reduction in medical mistakes over five years. This was not quite a zero vision, but halfway there. Many in this world are still battling the very idea that errors don't occur. When in medicine, you practice perfection, otherwise you don't practice. Or shouldn't. Many in that*

world are faced daily with situations where errors are not the consequence of systemic complexities, but considered to be the cause of harm. They are often seen as shameful lapses, as moral failures or failures of character in a practice that should aim to be perfect. 'Human error,' in such situations, is not seen as the systematic by-product of the complexity and organization of care, but as the result of "human ineptitude;"[8] as a result of some people lacking the "strength of character to be virtuous."[9] It is not surprising that the medical world is foremost in producing what are known as second victims—practitioners involved in an incident for which they are left to feel personally responsible and guilty, without much meaningful support.[10] The more "errors" are morally outlawed by an organization's commitment to zero, the more likely it is that practitioners involved in incidents will be turned into second victims.

Many organizations don't realize (or do not sufficiently acknowledge) that operators involved in an incident are victims too. In healthcare, but in other worlds too, they may be *second* victims: involved in an event that (potentially) harms or kills someone else, and for which they feel guilty and personally responsible. If operators are confronted with having violated the organization's commitment to zero as well, they can feel even worse. And they might be less inclined to talk about the event or seek support. This in turn can harm their chances of recovery and re-integration.[11] Stigmatizing an incident as something for which people should feel guilty—because the organization has publicly committed to not having any incidents—is not a good way to keep your people. It is not a good way to keep them healthy. And it is not a good way for your organization to keep itself informed about what is going on.

No support for zero vision in the safety literature

Safety theorizing of the past few decades is generally pessimistic about achieving an incident- and accident-free organization. Man-made disaster theory, for example, has concluded that "despite the best intentions of all involved, the objective of safely operating technological systems could be subverted by some very familiar and 'normal' processes of organizational life."[12] According to the theory, such subversion of the best intentions (the zero vision) occurs through usual organizational "errors" such as:

- information not being fully appreciated;
- information not correctly assembled;
- information conflicting with prior understandings of risk.

Barry Turner, founder of man-made disaster theory (you might recall this from earlier chapters), noted that people were prone to discount, neglect or not take into discussion relevant information. These are of course hindsight-driven judgments on his part. The point he was trying to make, though, was that no matter what vision managers, directors, workers or other organization members commit to, there will always be:

- erroneous assumptions and misunderstandings;
- rigidities of human belief and perception;
- disregard of complaints or warning signals from outsiders;
- decoy phenomena that suggest that the biggest risk is in a place where it actually isn't at all;
- a reluctance to imagine worst outcomes.

These, Turner suggested, may seem like 'human errors' at first sight. The target for intervention is the behavior and attitudes of managers in the organization. They need to be told to try harder, to not make such errors. They need to be reminded to pay more attention, to not get distracted, to not lose awareness of what really matters. But on closer inspection, these things are the normal by-product of humans bureaucratically organizing their work.[13]

More recent research has confirmed this. Vaughan's analysis of the 1986 Space Shuttle Challenger launch decision confirmed that the potential for having an accident grows as a normal by-product of doing business under normal pressures of resource scarcity and competition. Telling people not to have accidents, or getting them to commit to not having an accident, is not very promising. The potential for mistake and disaster, said Vaughan, is socially organized, embedded deeply in the way your organization does its work. This has to do with:

- *cultures of production* where problem-solving under pressure and constraints is highly valued;
- *structural secrecy* associated with bureaucratic organization, where information does not cross the boundaries of the various silos in which work is done and administered;
- *gradual acceptance* of more risk as bad consequences are kept at bay. The potential for an accident can actually grow underneath the very activities that your organization undertakes in order to tell itself and others that risk is under control (for example, measuring and tabulating injury numbers).

High-reliability organization (HRO) theory is generally known as a more optimistic way of looking at your organization's capacity to prevent accidents.

But it is very ambitious in its requirements for leadership and organizational design. In fact, so ambitious that a total elimination of accidents (that vision zero) is out of reach. This includes:

- leadership safety objectives;
- maintenance of relatively closed operational systems;
- functional decentralization;
- the creation of a safety culture;
- redundancy of equipment and personnel;
- systematic learning.

All these things are required for your organization to achieve HRO status.[14] While some organizations may come closer to some of these ideals than others, there is none that has closed the gap perfectly. What is more, there are no guarantees that manipulating these attributes will keep an organization at zero.[15]

Decades before, Charles Perrow came out with his book *Normal Accidents*.[16] He proposed that the risk of having an accident is not dependent on anybody's commitment to zero. The risk of having an accident is a fixed, structural property of the complexity of the systems we choose to build and operate. If we build really complex things, like nuclear power plants, then we should expect those to produce interactions and couplings that are really hard for us to understand. (This has been true more recently of surprisingly complex things like collaterized debt obligations in financial markets as well.[17] These are things that we can build and understand in isolation. But when we release them into a complex, dynamic world, their interconnections and interactions and couplings multiply beyond our ability to predict or understand.) Perrow called the features of these systems that make them more susceptible to an accident:

- *Interactive complexity.* Interactive complexity makes it difficult for humans to trace and understand how failures propagate, proliferate and interact.
- *Tight coupling.* Tight coupling means that the effects of single failures reverberate through a system—sometimes so rapidly or on such a massive scale that intervention is impossible, too late, or futile.

The only way to achieve a zero vision in such a system is to dismantle it, and not use it altogether. Which is what Perrow essentially recommended us to do with nuclear power generation.[b]

b Some would argue that Perrow's prediction has not been borne out quantitatively since the theory was first publicized in 1984. Perrow's epitome of extremely complex and highly

The most important realization from this is that a vision zero does not come out of safety research. That vision is exactly that: a vision. It is an ethical commitment to do no harm. As such, it is laudable. But it is not a research result, it is not a scientific prediction. There is no accident theory underpinning it or supporting it. And in fact, making your commitment to zero a firm one may actually harm your organization's safety.

A blindness to real risk

In 2005 BP was touting an injury rate many times below the national average at its Texas City facility, but then an explosion there took the lives of 15 workers and injured 180. More recently, BP (whose oil rig Deepwater Horizon exploded in the Gulf of Mexico on April 20, killing 11 and causing untold damage in the Gulf) had been a recipient of safety awards for having low recorded injury rates in their facilities.

Similarly, the resources industry in Australia operated an industry award system (it was called the Minex awards). Awards were given each year to the best (that is, safest) mine in Australia. It was given only after a stringent and rigorous auditing of the mine's safety systems, processes, culture and performance. This was done iteratively to ensure that the award would go to only the truly deserving top mine. Northparkes Mine in central New South Wales was given a high commendation by the Minex panel. Safety culture surveys put it in the top third of "positive responses" among its peers. Four months later, in 1999, four men were killed at one of its underground lifts as a result of a massive rock collapse and subsequent air concussion.

A zero vision does not come out of safety research. It is neither a research result, nor a scientific prediction. There is no accident theory underpinning it.

coupled systems—nuclear power generation—has produced only a few accidents, for example. Yet the 2011 earthquake-related disaster at Fukushima closely followed a Perrowian script. The resulting tsunami flooded low-lying rooms at the Japanese nuclear plant, which contained its emergency generators. This cut power to the coolant water pumps, resulting in reactor overheating and hydrogen-air chemical explosions and the spread of radiation. Also, increasingly coupled and complex systems like military operations, spaceflight and air traffic control have all produced Perrowian accidents since 1984.

> *As a third example, back in 1988, the Piper Alpha Oil Rig in the North Sea had been considered one of the company's safest and most productive platforms. Its permit-to-work system had been named as its most outstanding successful system. On this rig, an explosion and subsequent fire occurred when a pipe starting leaking gas, which ignited. A temporary flange, with no safety valve, had been used to block off the pipe during a maintenance operation the previous shift. The permit to advise operators not to start the pumps on this line (part of that celebrated permit-to-work system) was misplaced or lost. The water deluge system was inoperarable at the time, and 167 men, most of them huddled and waiting for further instructions in the accommodation block, perished in the fire. During the inquiry after the accident, the rig manager said, "I knew everything was all right, because I never got a report that anything was wrong."*

Why do these catastrophes happen to seemingly excellent organizations? Low numbers of negatives might feed the illusion that risk is under control. And in a narrow sense, some risk may indeed be under control. Or at least it may look like it is. But that risk may ultimately be irrelevant to the survival of the organization or the safety of its key process. Diverting so much of your organization's safety resources to managing numbers of injuries and incidents may mean that you are losing sight of the process risks that matter. Remember the greatest residual risk of already safe organizations: the gradual drifting into failure of your entire operation. This typically happens under the assumption that no negative events represent a control of all risk. That no bad news is only good news. Which, of course, is an illusion. If you are measuring risk only very narrowly and suppressing the incidents associated with that risk, it means nothing about the risk that you are not measuring, that you have lost sight of. As, indeed, some of the organizations in the examples above might have.

The safety risk of low incident numbers

Lower reporting rates can have benefits for the organization, of course. This may include the following:

- you may save money on healthcare, insurance and other compensation costs, if you can show that (or the worker can't disprove that) the injury or incident didn't happen at work;
- the chance to get contracts renewed or get additional work goes up if you can show low incident numbers;
- you may be less likely to get regulatory inspections if you have low incident numbers;

- If your organization or industry is self-insured for the costs of incidents and injuries, insurance auditors probably want to see low numbers of negative events to determine your premium rates.

To ensure low incident numbers, or to be able to show clearly where incidents are still happening—organizations are increasingly able to monitor and track compliance to their targets. They can do that through a variety of "panoptisms" (or "see-all" technologies). These range from cockpit voice recorders, to intelligent vehicle monitoring systems, to video recorders in operating theaters in some hospitals. Safety audits, safe work observations, local supervision, and safety culture measurements and safety climate assessments add to the mix. Together, these things create a form of social control. Such social control connects individual behavior to organizational norms and expectations. It also exerts a kind of disciplinary effect. Behavior and non-compliance get recorded in files, reports, case notes and more. There are additional ways by which these lower numbers are achieved. These can nowadays include:[18]

- safety incentive programs, where workers receive prizes or rewards when they don't report work-related injuries;
- injury discipline policies, where workers are threatened with or receive discipline (such as fines, warnings or even termination) when they do report injuries;
- post-injury drug or alcohol testing, where workers are automatically drug- or alcohol tested when they report an injury;
- workplace signs that track the number of hours or days without a lost-time or recordable injury. This encourages numbers games;
- other posters or signs, such as those stuck to washroom mirrors stating, "You are looking at the person most responsible for your safety;"
- targets for numbers of incidents of a particular category that an organization or unit is allowed to have each year. One air traffic control organization, for example, allows only six incidents per unit per year. If they go over that, the unit is stigmatized and penalized. No unit has recently had more than six incidents, so the managers can claim success (right?);
- programs where workers observe co-workers and record their "safe behaviors" or "unsafe acts." This focuses attention away from workplace hazards and reinforces the Old View idea that 'human error' is the cause; that incidents result from workers' bad behavior rather than hazardous conditions.

A Louisiana man is spending time in prison for lying about worker injuries at a local power utility, which allowed his company to collect $2.5 million

in safety bonuses. A federal court news release says that the 55-year-old was sentenced to serve six and a half years in prison followed by two years of supervised release.

He was the safety manager for a construction contractor. He was convicted in November of not reporting injuries at two different plants in Tennessee and Alabama between 2004 and 2006. At his federal trial, jurors heard evidence of more than 80 injuries that were not promptly recorded, including broken bones, torn ligaments, hernias, lacerations and injuries to shoulders, backs and knees. The construction contractor paid back double the bonuses.[19]

The underreporting that results from the implementation of such safety programs really means that you are shooting yourself in the foot. Your organization does not have a great safety culture because it has low numbers of incidents. In fact, the opposite is true. This has been shown in various industries already. A study of Finnish construction and manufacturing from 1977 to 1991, for example, showed a strong correlation between incident rate and fatalities, but reversed ($r = -.82$, $p<0.001$).[20] In other words, the fewer incidents a construction site reported, the higher its fatality rate was (see Figure 7.1).

Figure 7.1 shows this nicely. The horizontal (x) axis shows the fatality rate for the hours worked on construction sites in a given year, the vertical (y) axis shows incident frequency for that same year. As the incident rate increases, the fatality rate declines. Efforts to reduce incident rates (which may involve discouraging or suppressing their reporting, see above) are strongly correlated with a higher fatality rate.

This negative correlation between number of incidents and fatalities has been shown in aviation as well. Table 7.3 (on page 178) is based on an analysis by Barnett and Wang,[21] and shows correlations between nonfatal accidents/incidents per 100,000 departures for individual major carriers with their passenger-mortality risks. As in the construction site data above, all the correlations are negative. This means that carriers with higher rates of nonfatal accidents/incidents had lower mortality risks.

Interestingly, the correlations become increasingly negative as the events become more severe: from –0.10 for incidents only to –0.34 for serious nonfatal accidents only. In other words, the closer to death an airline has brought its passengers, the less likely it is to actually get them there from then on. This supports the importance (as emphasized by HRO theory, and many others) of learning from near misses. Suppressing such learning opportunities, at whatever level, and by whatever means, is not just a bad idea. It is dangerous.

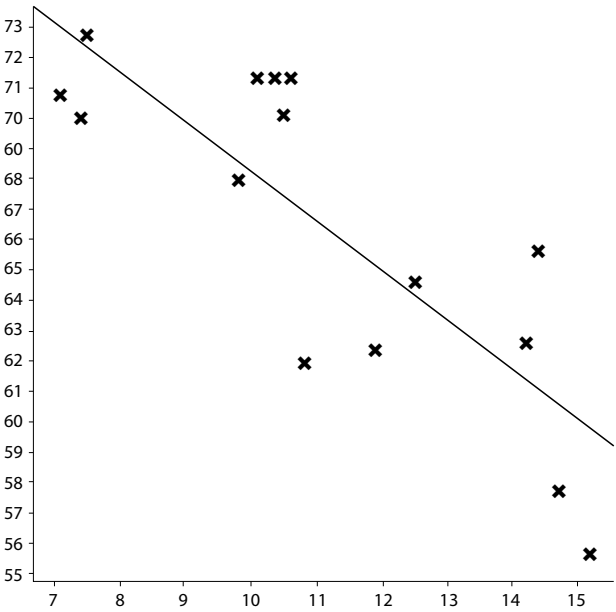


Figure 7.1 As the incident rate declines, the fatality rate increases. These data come from a construction industry study on sites between 1977 and 1991.* The horizontal (x) axis shows the fatality rate for all the manhours worked in a given year (represented by the data points in the plot). The vertical (y) axis shows incident frequency for that same year. Efforts to reduce incident rates (which may involve suppressing their reporting) are strongly correlated with a higher fatality rate

Source: Saloniemi A, Oksanen H. Accidents and fatal accidents: Some paradoxes. *Safety Science* 1998;29:59–66.

How Safe is the Activity Already?

The actions you can successfully implement toward a New View safety culture depend on how safe that part of your system already is. Unsafe systems typically kill or gravely injure one in ten to 1,000 people per activity. Safer systems one in 10,000, safe systems one in 100,000 and ultra-safe systems one in 1,000,000 or better.[22] The way safety is made and broken inside of these systems is quite different. And the sorts of things you might want to recommend these systems to do to become safer should depend on where they are at that moment.

Table 7.3 Correlation of major US jet air carrier nonfatal accident/incident rates and passenger-mortality risk from 1990 to 1996 (the serious nonfatal accident category excludes turbulence-related accidents and serious accidents in flight, at the gate or on the ramp)

Type of Nonfatal event	Correlation with passenger mortality
Incidents only	–0.10
Incidents and nonfatal accidents	–0.21
Nonfatal accidents only	–0.29
Serious nonfatal accidents only	–0.34

- **Unsafe systems,** such as certain types of mountain climbing or surgery (for example, transplant). The risk of failure (including fatalities) is inherent in the activity and is accepted, as it is the other side of trying to extract maximum performance. Or the activity may simply be worth the gamble because the alternative is even worse. Outstanding performance and a constant search for maximum performance make unsafe systems work. Greater safety can sometimes be expected through increasing practitioner competence. That often takes care of itself, though, as participants in such activities tend to be competitive and constantly out to improve their own performance.
- **Safer systems**, such as road traffic or certain types of healthcare. Safety improvements here can come from standardization—of participants (through training), of the work (through rules and procedures) and of the technology used (through ergonomics). People in safer systems make their own choices, but such choices can be led in better ways through standardization and also by enforcing compliance (for example, making sure people wear seatbelts or bike helmets, even if the statistics on these is sometimes doubtful). Quality control can help monitor whether components and processes meet pre-specified standard criteria. (These hard fixes may be tried in unsafe systems as well, but they could well be resisted because of the importance of individual competence there.)
- **Safe systems,** such as our food supply or charter airline flying. Safety in these systems comes in part from top management commitment to safety. This commitment is made obvious by a written safety policy and an explicit advertisement of how much resources management spends on safety. The organization does safety monitoring beyond quality control (for example, through event reporting and deeper analysis), safety management (for example, by safety departments)

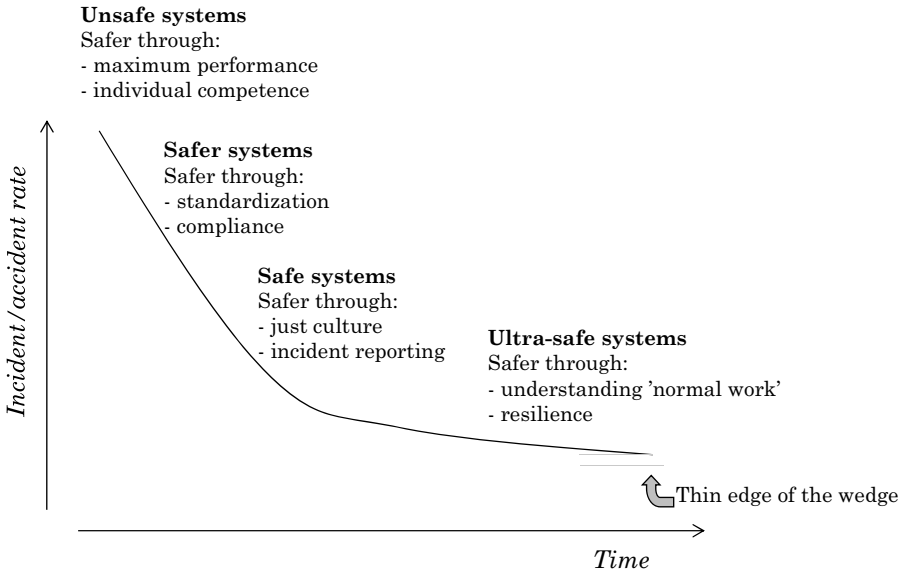


Figure 7.2 Which safety interventions make sense depends on how safe the system already is (based on René Amalberti's research).* Over time, our industries have become safer, with many now living at the "thin edge of the wedge." The safer you become, the more difficult it is to achieve safety improvements

Source: Amalberti R. *Navigating safety: Necessary compromises and trade-offs—theory and practice*. Heidelberg: Springer, 2013.

and emphasizes skills that go beyond competent individual practice (for example, teamwork or other "soft" skills). In these safe systems, incident reporting is worthwhile, as the ingredients of incidents (dress rehearsals) tend to show up in accidents. Recommendations for making safe systems even safer include:

- **Ultra-safe systems**, such as scheduled international airline flying or nuclear power generation. The route to risk in ultra-safe systems is one where the entire system gradually changes its definition of what is safe or acceptable practice. A slow drift into failure is the biggest source of residual risk here. This means that incident reporting (still effective at the previous level) may no longer help predict or prevent accidents, as the definition of an "incident" has shifted for those who do the work. The sorts of things that show up in accidents will likely not have been reported as incidents, because they weren't seen as such at the time. This means that recommendations have to focus on understanding

what people see as "normal work" including the daily frustrations, workarounds and normalizations of practice that they have to engage in to get the job done. Ultra-safe systems require an understanding of what makes them resilient: how are they able to continue their functioning despite production pressures, design issues, organizational conditions and usual disruptions?

You want to peg your investments in safety to the level of safety the particular activity has already attained. That way, it will not be either unattainable or irrelevant. Telling people in an ultra-safe activity that they need more rules and better compliance, for example, is not going to make things any safer. That would fit better in a system that is merely safer than the unsafe.

Life at the thin edge of the wedge

Many industries have become considerably safer over the past century. For sure, there are activities that lie at or near the unsafe system definition (certain forms of surgery, for example, or recreational activities such as base jumping). But many organizations are now ultra-safe or near-zero. They live at the thin edge of the wedge. Creating even more progress on safety there is quite difficult. It is easy to see, however, that doing more of the same does not lead to something different: it will maintain the status quo. In other words, emphasizing standardization and compliance, and investing in incident reporting, will help the organization maintain its safety level. It will do little, however, to improve safety even further. Let's focus on ultra-safe or near-zero organizations for the moment. This is important because:

- Safety improvements over the last decades, including technologies, regulation, training, reporting systems, social awareness and more, have brought many industries into or toward the ultra-safe band. Parts of your industry may be close to, or in it, as well.
- Making an ultra-safe or near-zero organization even safer is one of the most difficult challenges. There is little theory to guide you. Safety progress in such organizations has gone asymptotic: that is, the curve of progress is plateauing.
- At the same time, ultra-safe or near-zero organizations still suffer surprisingly big accidents from time to time. These accidents seem to come from nowhere: the safety statistics, incident reports, or intelligence about holes or weaknesses picked up by the safety management system did not predict what was going to happen.

In many cases, actions that get taken to improve the safety of ultra-safe organizations are those that would fit less safe organizations. For example, these may be recommendations to create a more detailed policy on some aspect of the operation or design of the system. Or a recommendation to intervene in the behavior or attitude of people through an awareness campaign. They may be new calls for standardization, the broadcasting of a particular procedure, or stringent reminders to follow an existing rule or procedure. Such recommendations are fit for less safe systems, or even unsafe ones. That is where they have had demonstrated effect. But in ultra-safe organizations, such recommendations have no longer much effect in increasing safety.

> *For example, as Rene Amalberti observes, the rate of production of new guidance materials and rules in the European Joint Aviation Regulations is significantly increasing. But for years, aviation safety has remained on a plateau at 10^{-6} (even though over 200 new policies/guidance/rules are issued each year). Since nobody knows really what rules or guidance materials are really linked to the final safety level, the system is purely additive. New rules and guidance are added, and old rules and guidance material are never cleaned up. It is no surprise that regulations become inapplicable sometimes, or start conflicting with other regulations. Those who work in the field, as a result, exhibit more and more violations in reaction to this increasing legal pressure.*[23]

Safety actions taken in ultra-safe (near-zero) organizations are often repetitions or retreads of those taken in a less safe organization. They miss the point and do not help in creating additional safety.

If you look at Figure 7.2, you will see that life in an ultra-safe or near-zero organization is like living on the thin edge of the wedge. Your organization is ultra-safe. All the things that it has done to improve its safety have had great success so far. But there is no guarantee that you have all the risk under control. At the thin edge of the wedge:

- That which is going to bite you—the unlikely but still possible accident—will likely be something you were not measuring;
- It will likely be something that has not been reported as an incident by one of your workers.

Decoy phenomena

What can offer some insight into managing risk at the thin edge of the wedge? One is "decoy phenomena." Barry Turner identified these in the 1970s in his studies of industrial disasters and accidents (he was, as you might recall, the founder of man-made disaster theory). Decoy phenomena are apparent safety issues that take up a lot of the organization's attention. They may have been singled out as a safety priority. Your organization may have labeled them as one of the "top five" or "top three" risk priorities that needs to be controlled in order to prevent an accident. Such decoy phenomena are strong in pulling both attention and organizational resources away from other possible sources of risk. These may as yet be unknown. Or they may be known but not believed to be very dangerous. Decoy phenomena are typical for organizations at the thin edge of the wedge. Pursuing what they see as the last frontiers in improving their safety performance, they may devote so much time to one or a few issues that they lose sight of the rest.

A student of mine was deeply involved in the investigation of an aviation accident (also alluded to in the beginning of The Field Guide*). At the time, his airline (and the regional airline industry around it) was focused on a number of high-altitude incidents that had occurred to the sorts of small jets that were popular in the industry at the time. These incidents showed that pilots were not always familiar with the high-altitude (thin air) aerodynamics of these jets. One incident, for example, led to the flame-out of the jet's engines. The jet, on a ferry flight (with no passengers onboard) crashed, killing the two pilots. The airline, in response to the accident, restricted the altitude to which pilots could take its aircraft. And the industry focused on other high-altitude incidents, trying to predict the next accident.*

One morning, two years later, my student got a call from a colleague, who said "We've lost one, we've lost one." Immediately, my student thought of a high-altitude accident. It wasn't. The jet his airline had lost, with 49 fatalities, crashed on takeoff from a badly lit regional airport that was under construction at the time. The airport charts available in the cockpit were confusing and inaccurate enough to direct the pilots to the wrong runway. That runway proved too short for the pilots to get their jet safely airborne.

High-altitude incidents had become the industry's decoy phenomenon. Sure, there may have been some allusions in the industry's reporting database about airport confusions and construction, but it was never elevated to the level of accident risk. Attention and resources were focused elsewhere. This

is how an industry at the thin edge of the wedge, an ultra-safe industry, suffered that big, surprising accident.

What is a good way to deal with decoy phenomena? One thing to think about is that the incidents you are still having at the thin edge of the wedge may actually be markers of resilience you should celebrate; more than markers of risk that you should try to further control. Incidents that happen at the thin edge of the wedge, but which do not have far-reaching or fatal consequences, could constitute good stories of how your operators are successfully dealing with the pressures, constraints and surprises of their daily work. If you, as in the example above, try to pursue these incidents with all your energy and focus, you may actually lose sight of the places where real risk is brewing.

Don't be overconfident in your safety management system

Safety management systems have a very important role to play in the transitions from safer to safe to ultra-safe organizations. Among many other things, they:

- support the standardization of important things (procedures, designs, training, as well safety-related positions and responsibilities throughout the organization);
- help draw up and track the rules and regulations that are applicable;
- assist in monitoring compliance and tracking deviations;
- drive and guide the establishment incident-reporting systems and just cultures that make sharing and analysis of safety-related information possible.

Safety management systems were born out of the defense-in-depth thinking that was dominant in the late 1980s and the 1990s. The idea at the time, consistent with New View thinking and driven by earlier ideas about man-made disasters, was that we shouldn't just look for sharp-end interventions to improve safety. Instead, we should shift our gaze to the administrative, or blunt, end of the organization. That, after all, is where many layers of defense against failure are introduced, constructed, resourced, monitored and assured (for example, maintenance, procedure design, line supervision, training). The problem, identified at the time, was the imperfection of those layers of defense—their porousness. One way to think about it is that safety management systems were introduced in part to help find the holes in those layers of defense and fix them before they could contribute to an accident.

Such thinking has been responsible for continuous safety improvement. It has helped popularize the New View idea that 'human error' at the sharp end

is not the cause of trouble, but a consequence of trouble that can be found by digging more deeply into the blunt end of the organization. At the thin edge of the wedge, however, this logic cannot improve safety even further.[24] For sure—the efforts that have gone into it so far have to be sustained. That way, you can maintain the level of safety that has been achieved. But at the thin edge of the wedge, doing more of the same will not lead to additional safety improvement. Doing more of the same will not lead to something different. You will stay where you are.

Don't just look for the holes to plug

The reason for that seems to be this. At the thin edge of the wedge, an organization has managed to fine-tune many of its processes, and it has got its known sources of risk under acceptable control. The surprising accidents and fatalities that still occur at the thin edge of the wedge, however, seem to be preceded not by "holes" in the layers of defense nor by incidents that have been formally reported as risky events that foreshadow an accident. Instead, they seem preceded by normal work, done by normal people in what looks to everybody to be a normal organization. This will likely include the workarounds and daily frustrations, the improvisations and adaptations to get designs and procedures to work in practice, the slightly unreliable results or readings from various tests (for example, end-play measurements, negative pressure tests), the shortcuts that get taken to accommodate production pressures. Those things are no longer regarded as worthy of reporting. They are not seen as "holes." They are seen as normal, daily work, as all-in-the-game. And mostly they do not lead to trouble. Mostly, people in the organization have learned how to smoothly adapt around such difficulties, hiccups and pressures. They have learned how to get the job done. They are, however, often precisely the things that show up as crucial in the fatalities and accidents that still do happen. The way workers "finished the design" of the external fuel tank of the Space Shuttle by covering up scratches in the foam insulation, for example;[25] or the problem with unreliable end-play measurements and underspecified lubrication procedures in the tailscrew of Alaska Airlines 261.[26] Neither problem can be found as formally reported incidents. Nobody saw these problems as holes in defensive layers at the time. At the thin edge of the wedge, holes and incidents are not the herald of accidents. Normal work is.

> *The story of Abraham Wald demonstrates this. Wald was born in the Austrian–Hungarian empire around the turn of the twentieth century. After studying mathematics in Vienna, he could not get a university position because of his Jewish heritage, and in 1938 he was able to emigrate to the*

United States. There, he applied his statistical skills to the problem of Allied bomber losses due to enemy fire (ground-based anti-aircraft flak as well as bullets from attacking fighter aircraft). A study had been made of the patterns of damage that aircraft returned with, and it was proposed that armor should be added to those places that showed the most bullet damage. Armor, of course, increases the weight of an airplane dramatically and cuts into its payload or range. So you have to be really picky about where you put it, and how much you put on.

Wald, after doing his own extensive statistical analyses of returning bomber aircraft, came to quite a different conclusion. The airplanes that made it back with holes in them, he concluded, were the ones who had taken hits in areas where they could survive and return. Adding armor to those places would not do anything to help them. Instead, he said, we should add armor to those places that did not show holes. Because those were the airplanes that didn't come back. His statistical analysis identified the weak spots in non-returning airplanes. These were the weak spots that led to the loss of the bomber when hit. Those areas had to be reinforced, he argued, not the areas with holes in them. His insight became seminal for what is today known as operations research. In a sense, the areas with holes in them were the decoy phenomenon. They were evidence of the survivable incident; a marker of resilience. These were not markers of fatal risk that needed to be further controlled. That risk, instead, was in the areas on the returning bombers that did not show bullet holes.

At the thin edge of the wedge, holes in layers of defense and formally reported incidents are no longer the herald of accidents or fatalities. Normal work is.

Ironically, Wald himself died in a plane crash (of quite a different kind) in southern India while on a lecture tour in 1950.

Wald's advice may well be called for in organizations at the thin edge of the wedge. If you want to seek out and prevent what is going to lead to a fatality or accident, then don't (just) look for the holes you know about, or the problems that show up in your incident reports or safety management system. Look in the places where there are *no* holes; where your people do not see holes, where they do not see things that are worthy of reporting. In other words, look at normal work. Get a sense of their daily experiences and frustrations, their workarounds, adaptations, of the places where they finish imperfect designs

and procedures in order to get the job done. It is those places, where there are no holes, that may one day play a role in your organization's fatality or accident.

What to put in those places where you see no "holes"

Go back to the analogy with Wald's holes in the fuselages of Allied bombers. What is the sort of "armor" that *you* should put in those places where there are no holes? Here are some key questions you should ask of yourself, your teams and your organization. The answers probably hold some of the armor that you and your organization can put in the places where it doesn't see holes:

- **Monitoring of safety monitoring** (or meta-monitoring). Does your organization or team invest in an awareness of the models of risk it embodies in its safety strategies and risk countermeasures? Is it interested to find out how it may have been ill-calibrated all along, and does it acknowledge that it needs to monitor how it actually monitors safety? This is important if your organization or team wants to avoid stale coping mechanisms, misplaced confidence in how it regulates or checks safety, and does not want to miss new possible pathways to failure.
- **Do not take past success as guarantee of future safety**. Does your organization or team see continued operational success as a guarantee of future safety, as an indication that hazards are not present or that countermeasures in place suffice? In this case, its ability to deal with unexpected events may be hampered. In complex, dynamic systems, past success is no guarantee of continued safety.
- **Resist distancing through differencing**. In this process, organizational or team members look at other failures and other organizations or teams as not relevant to them and their situation. They discard other events because they appear to be dissimilar or distant. But just because the organization or section has different technical problems, different managers, different histories, or can claim to already have addressed a particular safety concern revealed by the event, does not mean that they are immune to the problem. Seemingly divergent events can represent similar underlying patterns in the drift toward hazard.
- **Resist fragmented problem solving**. It could be interesting to probe to what extent problem-solving activities are disjointed across organizational departments, sections or subcontractors, as discontinuities and internal handovers of tasks increase risk. With information incomplete, disjointed and patchy, nobody has the big picture, and nobody may be

able to recognize the gradual erosion of safety constraints on the design and operation of the original system that move an organization closer to the edge of failure.

- Knowing the **gap between work-as-imagined and work-as-done**. One marker of resilience is the distance between operations as management (including you?) imagines they go on and how they actually go on. A large distance indicates that your organizational leadership may be ill-calibrated to the challenges and risks encountered in real operations. Also, they or you may miss how safety is actually created as people conduct work and gather experiences from it.
- Keeping the **discussion about risk alive** even (or especially) when everything looks safe. One way is to see whether activities associated with recalibrating models of safety and risk are going on at all. This typically involves your people discussing risk even when everything looks safe (as gets done, for example, in a preflight briefing that goes through the threats and risks to safety right there and then, even if all looks fine and it has always gone well up to then). Indeed, if discussions about risk are going on even in the absence of obvious threats to safety, you could get some confidence that your organization or team is investing in an analysis, and possibly in a critique and subsequent update, of its models of risk.
- Having a person or function within the system with the **authority, credibility and resources** to go against common interpretations and decisions about safety and risk. Historically, "whistleblowers" may come from lower ranks where the amount of knowledge about the extent of the problem is not matched by the authority or resources to do something about it, or where there is no power to have the system change course. Your organization shows a level of maturity if it succeeds in building in this function or possibility at meaningful organizational levels. This also relates to the next point.
- The ability and extent of bringing in **fresh perspectives**. Organizations or teams that are able to apply fresh perspectives (for example, people from another backgrounds, with diverse viewpoints) on problem-solving activities seem to be more effective in managing risk. They tend to generate more hypotheses, cover more contingencies, openly debate rationales for decision making and help reveal hidden assumptions. With a neutral observer or commentator thus "institutionalized," you could be slightly more confident that your organization or team can "self-regulate" its safety, even or particularly in those places where no known risk exists or where no "holes" are seen.

Notes

1 Guldenmund, F.W. The nature of safety culture: A review of theory and research. *Safety Science* 2000;34(1):215–257.

2 Gray, G.C. The responsibilization strategy of health and safety. *British Journal of Criminology* 2009;49(3):326–42.

3 Frederick, J., Lessin, N. The rise of behavioural-based safety programmes. *Multinational Monitor*, 2000;41(1):1–7.

4 Ibid.

5 GAO. Workplace safety and health: Better OSHA guidance needed on safety incentive programs (Report to Congressional Requesters). Washington, DC: Government Accountability Office, 2012.

6 Donaldson, C. Zero harm: Infallible or ineffectual. *OHS Professional*. Melbourne: Safety Institute of Australia, 2013:22–7.

7 IOM. Patient safety: Achieving a new standard for care. Washington, DC: National Academy of Sciences, Institute of Medicine, 2003.

8 Gawande, A. *The checklist manifesto: How to get things right*, First Ed. New York: Metropolitan Books, 2010.

9 Pellegrino, E.D. Prevention of medical error: Where professional and organizational ethics meet. In: Sharpe VA, editor. *Accountability: patient safety and policy reform*. Washington: Georgetown University Press, 2004:83–98.

10 Dekker, S.W.A. *Second victim: Error, guilt, trauma and resilience*. Boca Raton, FL: CRC Press/Taylor & Francis, 2013.

11 Ibid.

12 Pidgeon, N.F., O'Leary, M. Man-made disasters: Why technology and organizations (sometimes) fail. *Safety Science* 2000;34(1–3):15–30.

13 Turner, B.A. *Man-made disasters*. London: Wykeham Publications, 1978.

14 Rochlin, G.I., LaPorte, T.R., Roberts, K.H. The self-designing high reliability organization: Aircraft carrier flight operations at sea. *Naval War College Review* 1998;51(3):97–113.

15 Sagan, S.D. *The limits of safety: Organizations, accidents, and nuclear weapons*. Princeton, NJ: Princeton University Press, 1993.

16 Perrow, C. *Normal accidents: Living with high-risk technologies*. New York: Basic Books, 1984.

17 Dekker, S.W.A. *Drift into failure: From hunting broken components to understanding complex systems*. Farnham, UK: Ashgate, 2011.

18 Frederick, *op. cit.*

19 Anon. Jail for safety manager for lying about injuries. *Associated Press* 12 April 2013.

20 Saloniemi, A., Oksanen, H. Accidents and fatal accidents: Some paradoxes. *Safety Science* 1998;29:59–66.

21 Barnett, A., Wang, A. Passenger mortality risk estimates provide perspectives about flight safety. *Flight Safety Digest* 2000;19(4):1–12.

22 Amalberti, R. The paradoxes of almost totally safe transportation systems. *Safety Science* 2001;37(2–3):109–26.

23 Ibid.

24 Ibid.

25 CAIB. Report Volume 1, August 2003. Washington, DC: Columbia Accident Investigation Board, 2003.

26 NTSB. Loss of control and impact with Pacific Ocean, Alaska Airlines Flight 261 McDonnell Douglas MD-83, N963AS, about 2.7 miles north of Anacapa Island, California, January 31, 2000 Washington, DC: National Transportation Safety Board, 2002.

This page has been left blank intentionally

8 Abandoning the Fallacy of a Quick Fix

So how do you make a 'human error' problem go away? The answer isn't as simple as the question. 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 or the accident. 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 Old View countermeasures. Here are some really obvious, but false, quick fixes:

- reprimanding the people involved;
- retraining the people involved in a mishap;
- writing a procedure;
- adding just a little bit more technology.

These quick fixes are false because:

- Reprimanding people involved in the incident typically stops learning from that incident. Your organization cannot learn and punish at the same time.
- Who says it's just them? Retraining the people involved will have limited reach into your organization. Their performance could be symptomatic of deeper problems and issues that all your people or similar practitioners are exposed to.
- Writing an extra procedure often only deals with the latest "hole" uncovered by the mishap. It simplistically assumes a linear trajectory toward the failure, for which one best method (as specified in the procedure) is the remedy. And when you don't take anything away, additional procedures also tend to create less transparency and more non-compliance.
- Adding just a little bit more technology will likely create new work for people, new error opportunities and new pathways to breakdown.

A 'human error' problem is an organizational problem. A 'human error' problem is at least as complex as the organization that helped create it. If you want to do something about your 'human error' problem, you will have to start seeing it as the effect of problems deeper inside your organization. Not as the simple cause of all your trouble.

So what should you expect your organization to do instead? Quick fixes may be no good. They may in fact be a sign of your organization trying not to learn. But what, then, are the hard fixes? Hard fixes are evidence that people inside the organization are taking the surprise of a problem seriously. Rather than trying to reduce that surprise by pinning the responsibility for a problem on a few Bad Apples, they see that the problem tells them something interesting about their organization; about how it has managed risk and produced safety so far. The failure may show, for example:

How do you make 'human error' go away? The answer isn't as simple as the question. A 'human error' problem, after all, is an organizational problem. It is at least as complex as the organization that has helped create it.

- how otherwise legitimate trade-offs between safety and goals such as production or efficiency may have made it increasingly difficult for people to create safety through practice;
- how the entire system may have drifted toward the boundaries of safe performance, taking the definition of "acceptable risk" or acceptable practice along with it;
- how incentive structures governing people's performance may actually have encouraged a focus on economic or efficiency goals, with safety being taken for granted;
- how the priorities and preferences that employees express through their own practice may be a logical reproduction of that which the entire organization finds important;
- how managerial attention to safety has eroded, or never really developed, leaving the organization without explicit safety policy and no clarity about the resources committed to safety;
- how organizational models about the sources of risk, and organizational plans for how to deal with risk, may have been wrong or incomplete.

Hard fixes change something fundamental about the organization. This is what makes them hard. But it is also what makes them real fixes. Once

organizations start reflecting on the question of how to make a 'human error' problem go away, there could be a window in which there is more openness to learning. The shock of a failure may indeed help open such a window, when:

- 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.

Of course, this atmosphere of openness, of willingness and commitment to learn and improve, can quickly become compromised by calls for accountability, by primitive knee-jerk reactions toward Bad Apples. How do you help avoid those? To the extent possible, walk your colleagues, your managers or your stakeholders through some of the material in chapters 1 and 2 of The Field Guide. Help them understand the different views of human error, help them recognize their own reactions to failure and how those stand in the way of making real progress on safety. The steps in the remainder of this chapter can assist you.

Hard fixes change something fundamental about, or in, the organization. This makes them hard. But it also makes them real fixes.

Reminders for in the Rubble

The last part of this chapter wraps together some the most important lessons from the Field Guide. It gives you a summary of the New View and its implications. It presents you with five sets of reminder. They are about:

- your own organization and the nature of safety and risk in it;
- what to think about when investigating 'human error;'
- doing something about your 'human error' problem;
- how to recognize Old View thinking;
- how to create progress on safety with the New View.

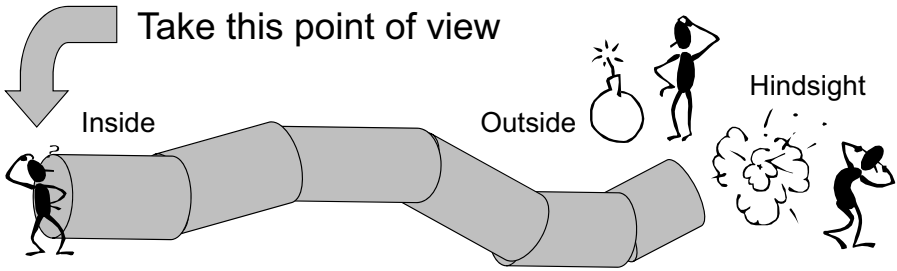


Figure 8.1 If you want to understand 'human error,' see the unfolding world from the point of view of people inside the situation—not from the outside or from hindsight

Whatever you try to understand 'human error,' do not forget to take the point of view of the person inside the situation.

Your organization and 'human error'

- Your organization is not basically or inherently safe. People have to create safety by putting tools and technologies to use while negotiating multiple system goals at all levels of your organization.
- The priorities and preferences that people express through their practice may be a logical reproduction of what the entire organization finds important.
- Human error is the inevitable by-product of the pursuit of success in an imperfect, unstable, resource-constrained world. The occasional human contribution to failure occurs because complex systems need an overwhelming human contribution for their safety.
- So 'human error' is never at the root of your safety problems. 'Human error' is the effect of trouble deeper inside your system.
- It also means that 'human error' is not random. 'Human error' is systematically connected to features of people's tools, tasks and operating environment.

What to think of when investigating 'human error'

- As far as the people involved were concerned, the outcome was not going to happen. If they knew it was going to, they would have done something else.
- Nobody comes to work to do a bad job. This is the local rationality principle. People do what makes sense to them at the time given their

focus of attention, their knowledge and their goals (which may well be the organization's goals, stated or unstated).

- Human error is not the cause of failure, but the effect. So 'human error,' under whatever label ("loss of situation awareness," complacency, inadequate CRM) can never be the conclusion of your investigation. It is the starting point.
- Explaining one error (for example, operator error) by pointing to another (inadequate supervision, deficient management, bad design) does not explain anything. It only judges other people for not doing what you, in hindsight, think they should have done.
- To understand what went on in somebody's mind, you have to reconstruct the situation in which that mind found itself.
- There is no such thing as the cause of a mishap. This is like looking for the cause of not having a mishap. What you deem causal depends on your accident model.

Reprimanding "Bad Apples" is like peeing in your pants. You did something about the problem and feel relieved. But then it gets cold and uncomfortable. And you look like a fool.

Doing something about your 'human error' problem

- Do not get deluded by the fallacy of a quick fix. 'Human error' problems are organizational problems, and so at least as complex as your organization which helped create them.
- Reprimanding supposed Bad Apples (errant operators) may seem like a quick, rewarding fix. But it is like peeing in your pants. You feel relieved and perhaps even nice and warm for a little while. But then it gets cold and uncomfortable. And you look like a fool.
- If you are a manager or supervisor, you cannot expect your employees to be more committed to safety than you yourself are, or appear to be.
- Problems result from your organization's complexity—not from apparent simplicity (for example, somebody's inattention).
- Do not expect that you can hold people accountable for their errors if you did not give them enough authority to live up to the responsibility you expect of them.

Recognizing old view thinking

- Old View thinking sees 'human error' as the major threat to basically safe systems. Unreliable, erratic people undermine systems of multiple defenses, rules, procedures and other safeguards.
- Old View thinking will try to count and categorize errors, and endeavor to get the number of 'human error' incidents down. It assumes that safety, once established, can be maintained by monitoring and keeping people's performance within pre-specified boundaries.
- Old View thinking will (unsuccessfully) try to revert to more automation, tighter procedures, closer supervision and reprimands to control erratic human performance.
- During downsizing, budget trimming and increased production pressures, Old View thinking will misinterpret 'human errors' as a source of trouble, when they are likely the inevitable downstream consequences of trying to do more with less.
- Old View thinking judges rather than explains human performance. It uses language such as "they should have…" "if only they had…" and "they failed to…" But by saying what people should have done, you don't explain at all why they did what they did.

Creating progress on safety with the new view

- To create safety, you don't need to rid your system of 'human errors'. Instead, you need to realize how people at all levels in the organization contribute to the creation of safety and risk through goal trade-offs that are legitimate and desirable in their setting.
- Rather than trying to reduce "violations," New View strategies will find out more about the gap between work-as-imagined and work-as-done—why it exists, what keeps it in place and how it relates to priorities among organizational goals (both stated and unstated).
- New View thinking wants to learn about authority–responsibility mismatches—places where you expect responsibility of your people, but where their situation is not giving them requisite authority to live up to that responsibility.
- You know your organization is maturing toward the New View once it actively tries to learn how it is learning about safety. This means your organization is calibrating whether its strategies for managing safety and risk are up-to-date.
- Every organization has room to improve its safety. What separates a strong safety culture from a weak one is not how large this room is.

What matters is the organization's willingness to explore this space, to find leverage points to learn and improve.

Where to Go from Here

One of the questions that people often raise after reading *The Field Guide* is: where do I go now? What should I read in order to continue my learning about a different way of thinking about safety? What if I am concerned about issues of justice and learning in my workplace? What is a good example of a New View investigation? The final section of this chapter is intended to help you on your way. Don't see what is proposed here as comprehensive or exclusive—there surely is a lot more out there. But here are some pointers.

Examples of new view investigations

A good example of an official New View investigation is available from the Transportation Safety Board of Canada. It is the investigation into the 1998 Swissair 111 accident that happened off the coast of Nova Scotia. It qualifies as a New View investigation for a number of reasons, but most prominent among them are:

- the efforts the investigators have gone through to understand the unfolding situation from the point of view of the crewmembers (who died in the accident);
- the extent to which investigators have tried to avoid hindsight and outcome biases by discounting knowledge of outcome in their assessment of crew decision making;
- the decision to not have "probable cause" statements (which always end up being narrow choices and constructions, as you have learned in *The Field Guide*), but rather go through their "findings as to causes and contributions," "findings as to risk" and other findings. The report subsequently spends a good number of pages on "safety action," which is the whole point of having done the investigation: prevention.

TSB (2003). Aviation investigation report: In-flight fire leading to collision with water, Swissair Transport Limited, McDonnell Douglas MD-11 HB-IWF, Peggy's Cove, Nova Scotia 5 nm SW, 2 September 1998. Gatineau, QC, Transportation Safety Board of Canada.

Another great example of a New View investigation is the revisionist account of the shootdown of two Black Hawk helicopters over northern Iraq in the 1990s. The Black Hawks were ferrying UN peacekeepers and were mistakenly downed by two US fighter jets. The account is written by Scott Snook, himself an ex-military man with plenty of insight into the messy details of the operation (Operation Provide Comfort). He went back to the official investigation and started to question not only their conclusions, but where they looked for evidence. It is a book in the best tradition of revisionist accounts. This tradition includes an early one into the Mount Erebus DC-10 disaster in Antarctica (Vette, G. (1984). *Impact Erebus*. New York: Sheridan House Inc.). It is remarkable how official Old View investigations inspire researchers, journalists and would-be sleuths around the world to try to do a better job—pinning blame on the front-line operators just does not cut it for them, as it doesn't for many others. Snook's book is not too long, it is well-organized and eminently readable. It should give you plenty of inspiration for how to set up your own New View investigation, where to look, and what to ask.

Snook, S.A. (2000). *Friendly fire: The accidental shootdown of US Black Hawks over Northern Iraq*. Princeton, NJ: Princeton University Press.

If you want to become even more versed in New View 'human error' investigations, there is a book that takes you much further behind that label than *The Field Guide*. A deeper treatment of 'human error,' and the cognitive and organizational factors behind human performance can be found in:

Woods, D.D., Dekker, S.W.A. et al. (2010). *Behind 'human error'*. Aldershot, UK: Ashgate.

This book explores how systems thinking has radically changed our understanding of how accidents occur, and explains the role of cognitive system factors—bringing knowledge to bear, changing mindset as situations and priorities change, and managing goal conflicts—in operating safely at the sharp end of systems. It also examines how the clumsy use of computer technology can increase the potential for erroneous actions and assessments in many different fields of practice, and examines in detail how the hindsight bias enters into attributions of error. As with Paul Fitts in 1947, the label 'human error' is still the result of a social and psychological judgment process. It limits our focus on only a subset of a complex of interacting contributors.

Instead of taking a deeper dive, there is an opportunity for a quick skim across the surface of the New View to 'human error' as well. This is offered

by Todd Conklin, whose *Pre-Accident Investigation Guide* gives you (or your manager) just that. It suggests that managers can get in on the safety action before anything ever happens by, for example, asking workers where the next incident will likely happen, and by using learning teams to investigate the accident that has not happened yet.

Conklin, T. (2012). *Pre-accident investigations: An introduction to organizational safety*. Farnham, UK: Ashgate.

Understanding Complexity and Drift into Failure

The New View (as did man-made disaster theory in the 1970s) helps you understand how failure is not the short-term outcome of people who did things wrong at the sharp end, but as something that gets incubated over a much longer period. During this incubation period, people do what makes sense to them at the time. They do what they believe is normal work in what looks like a normal organization. Drift into failure lies in the small, incremental shifts in how safety is traded off against other goals that are important to the organization, or, to speak with Diane Vaughan, in the gradual normalization of what was previously seen as deviant or unacceptable. To be sure, such gradual shifts are necessary for an organization to adapt and survive in a competitive environment, to explore and experiment with better, more efficient ways of doing things. Here are some pointers to books on drift into failure and the complexity of safety. Barry Turner laid an important piece of conceptual groundwork for these ideas. Charles Perrow explained how the potential for failure is a structural property of complex, interactive and tightly coupled systems. And Diane Vaughan has done more than any other social scientist to understand the phenomenon of drift in a complex system from the point of view of people inside the organization at the time.

Turner, B.A. (1978). *Man-made disasters*. London: Wykeham Publications.
Perrow, C. (1984). *Normal accidents: Living with high-risk technologies*. New York: Basic Books.
Vaughan, D. (1996). *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. Chicago, IL: University of Chicago Press.
Dekker, S.W.A. (2011). *Drift into failure: From hunting broken components to understanding complex systems*. Farnham, UK: Ashgate.

Steven Mandis' recent book *What Happened to Goldman Sachs* is an insider account of organizational drift and its unintended consequences. It makes rich use of the ideas in the *Challenger Launch Decision* and *Drift into Failure*.

Again written by someone with the privileged insight into the messy details of, in this case, investment banking, Steve traces how his firm morphed from adherence from ethical standards to merely legal ones as it became a global investment corporation. His analysis tries to show how it is possible for an organization's success to become its failure. He traces the pressures that forced Goldman to slowly drift away from the very principles, values and partnership culture on which its reputation and success was built, and how it is so very difficult for insiders to see this drift happen.

Mandis, S.G. (2013). *What* happened *to Goldman Sachs: An insider's story of organizational drift and its unintended consequences*. Boston, MA: Harvard Business Review Press.

Safety Differently and Resilience

The idea that we should be looking at safety as the presence of positive capacities rather than the absence of negative events, has recently been taken up by a number of authors and groups of thinkers. Resilience—as the ability of a system or team or individual to recognize, adapt to, and absorb disruptions that fall outside the design or preparation base, and to sustain or even improve its functioning—is one example of this. Not only were these ideas launched and collated in the mid-2000s, it is also an outgrowth of HRO thinking:

Hollnagel, E., Woods, D.D. et al. (2006). *Resilience engineering: Concepts and precepts*. Aldershot, UK: Ashgate.

Weick, K.E. and Sutcliffe, K.M. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty*. San Francisco, CA: Jossey-Bass.

Punchy as well as groundbreaking ways of dealing with these topics come of course from Erik Hollnagel, whose two recent books are short and very readable. The latest one takes to task the idea that safety is a condition where the number of adverse outcomes was as low as possible. In that case, the purpose of safety management is to make sure that the number of accidents and incidents is kept as low as possible. Safety is measured by counting the number of cases where it fails rather than by the number of cases where it succeeds. This unavoidably leads to a reactive approach based on responding to what goes wrong or what could go wrong. Instead, focusing on what goes right allows organizations to understand their ability to succeed under varying conditions, so that the number of intended and acceptable outcomes is as high as possible. Safety management, in this case, needs to ensure that as much as

possible goes right, and cannot only be reactive. It must also be proactive with regard to how actions succeed rather than with regard to how they can fail, as traditional risk analysis does.

Hollnagel, E. (2009). *The ETTO Principle: Efficiency-Thoroughness Trade-Off. Why things that go right sometimes go wrong*. Aldershot, UK: Ashgate.
Hollnagel, E. (2014). *Safety I and Safety II: The past and future of safety management*. Farnham, UK: Ashgate.

Finally, concern with safety is often inextricably connected to questions of accountability, learning and justice. A selection of books deals with this. One of the difficult questions is how to make accountability and learning work in each other's favor. It is, after all, really easy to hold people accountable while putting a damper on people's willingness to honestly share their safety concerns. Creating a "just culture" is incredibly difficult, but not impossible. Justice, however, is one of those essentially contested categories: reasonable people can always end up disagreeing about what it means. In any case, organizational justice does involve paying attention to second victims—practitioners involved in an incident that (potentially) harms or kills other people, and for which they feel personally responsible. There is a relationship between resilient individuals (who are supported in recovering from or even growing in the face of such incidents) and resilient organizations (which are able to face up to their vulnerabilities and learn from them).

Sharpe, V.A. (2004). *Accountability: Patient safety and policy reform*. Washington, DC: Georgetown University Press.
Dekker, S.W.A. (2012). *Just culture: Balancing safety and accountability* (Second Ed.). Farnham, UK: Ashgate.
Dekker, S.W.A. (2013). *Second* victim*: Error, guilt, trauma and resilience*. Boca Raton, FL: CRC Press/Taylor & Francis.

Further Reading

There are evidently many more books and directions for you to take in this than I can meaningfully discuss here. So let's finish this chapter with a bibliography of works that you might want to read at some point. These are among the books that have inspired me in articulating the New View. See if you have an opportunity to enjoy some of these, and be inspired too:

Amalberti, R. (2013). *Navigating safety: Necessary compromises and trade-offs: Theory and practice*. Heidelberg: Springer.

Berlinger, N. (2005). *After harm: Medical error and the ethics of forgiveness*. Baltimore, MD: Johns Hopkins University Press.

Billings, C.E. (1997). *Aviation automation: The search for a human-centered approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Burnham, J.C. (2009). *Accident prone: A history of technology, psychology and misfits of the machine age*. Chicago, IL: The University of Chicago Press.

Capra, F. (1982). *The turning point*. New York: Simon & Schuster.

Casey, S.M. (1993). *Set phasers on stun and other true tales of design, technology, and 'human error'*. Santa Barbara, CA: Aegean.

Cilliers, P. (1998). *Complexity and postmodernism: Understanding complex systems*. London: Routledge.

Cook, R.I., Woods, D.D. and Miller, C. (1997). *A tale of two stories: Contrasting views of patient safety*. Chicago, IL: National Patient Safety Foundation.

Degani, A. and Wiener, E.L. (1990). *Human factors of flight-deck checklists: The normal checklist*. Moffett Field, CA: NASA Ames Research Center.

Dörner, D. (1989). *The logic of failure: Recognizing and avoiding error in complex situations*. Cambridge, MA: Perseus Books.

Douglas, M. (1992). *Risk and blame: Essays in cultural theory*. London: Routledge.

Evans, C. and Holmes, L. (Eds). (2013*). Re-Tayloring management: Scientific management a century on*. Farnham, UK: Gower.

Feltovich, P.J., Ford, K.M. et al. (1997). *Expertise in context: Human and machine*. Menlo Park, CA: AAAI Press.

Feyerabend, P. (1993). *Against method*. London: Verso.

Feynman, R.P. (1988). *"What do you care what other people think?": Further adventures of a curious character*. New York, Norton.

Flach, J.M., Hancock, P.A. et al. (1996). *An ecological approach to human-machine systems I: A global perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Flin, R.H. and Mitchell, L. (2009). *Safer surgery: Analysing behaviour in the operating theatre*. Farnham, UK: Ashgate.

Gawande, A. (2002). *Complications: A surgeon's notes on an imperfect science*. New York: Picado.

Gawande, A. (2008). *Better: A surgeon's notes on performance*. New York: Picador.

Gawande, A. (2010). *The checklist manifesto: How to get things right*. New York: Metropolitan Books.

Geertz, C. (1973). *The interpretation of cultures*. New York: Basic Books.

Gergen, K.J. (1999). *An invitation to social construction*. Thousand Oaks, CA: Sage.

Gibson, J.J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Goode, E. and Ben-Yehuda, N. (1994). *Moral panics: The social construction of deviance*. Oxford, UK: Blackwell.

Gras, A., Moricot, C. et al. (1994). *Faced with automation: The pilot, the controller, and the engineer* (translated by J. Lundsten). Paris: Publications de la Sorbonne.

Hancock, P.A. (2009). *Mind, machine and morality: Toward a philosophy of human-technology symbiosis*. Aldershot, UK, Ashgate.

Heft, H. (2001). *Ecological psychology in context: James Gibson, Roger Barker and the legacy of William James's radical empiricism*. Mahwah, NJ: Lawrence Erlbaum Associates.

Heinrich, H.W., Petersen, D. et al. (1980). *Industrial accident prevention* (Fifth Ed.). New York: McGraw-Hill Book Company.

Helander, M.G. (2006). *A guide to human factors and ergonomics*. London: Taylor & Francis.

Hidden, A. (1989). Clapham Junction accident investigation report. London: HMSO.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method: CREAM*. Oxford; New York: Elsevier.

Hollnagel, E. (2004). *Barriers and accident prevention*. Aldershot, UK: Ashgate.

Hollnagel, E. (2009). *The ETTO Principle: Efficiency-Thoroughness Trade-Off. Why things that go right sometimes go wrong*. Aldershot, UK: Ashgate.

Hollnagel, E. (2014). *Safety I and Safety II: The past and future of safety management*. Farnham, UK: Ashgate.

Hollnagel, E., Nemeth, C.P. et al. (2009). *Resilience engineering: Preparation and restoration*. Aldershot, UK: Ashgate.

Hollnagel, E., and Woods, D.D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: Taylor & Francis.

Hollnagel, E., Woods, D.D. et al. (2006). *Resilience engineering: Concepts and precepts*. Aldershot, UK: Ashgate.

Hopkin, V.D. (1995). *Human factors in air traffic control*. London; Bristol, PA: Taylor & Francis.

Huey, B.M., Wickens, C.D. et al. (1993). *Workload transition: Implications for individual and team performance*. Washington, DC: National Academy Press.

Hutchins, E.L. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.

Jagacinski, R.J. and Flach, J.M. (2003). *Control theory for humans: Quantitative approaches to modeling performance*. Mahwah, NJ: L. Erlbaum Associates.

Janis, I.L. (1982). *Groupthink*, Second Ed. Chicago, IL: Houghton Mifflin.

Jensen, C. (1996). *No downlink: A dramatic narrative about the Challenger accident and our time*. New York: Farrar, Straus, Giroux.

Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities, and software*. New York: Scribner.

Kennedy, V. and Walker, D. (2007). *Dancing with Dr Death*. Sydney: New Holland Publishers.

Kern, T. (1998). *Flight discipline*. New York: McGraw-Hill.

Klein, G.A. (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex Pub.

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

Kohn, L.T, Corrigan, J. et al. (2000). *To err is human: Building a safer health system*. Washington, DC: National Academy Press.

Kuhn, T.S. (1970). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Langewiesche, W. (1998). *Inside the sky: A meditation on flight*. New York: Pantheon Books.

Lanir, Z. (1986). *Fundamental surprise*. Eugene, OR: Decision Research.

Leonhardt, J. and Vogt, J. (2006). *Critical incident stress management in aviation*. Aldershot, UK: Ashgate.

Leveson, N.G. (2012). *Engineering a safer world: Systems thinking applied to safety*. Cambridge, MA: MIT Press.

Lovell, J. and Kluger, J. (1994). *Lost moon: The perilous voyage of Apollo 13*. San Francisco, CA: Houghton Mifflin.

Mahler, J.G. (2009). *Organizational learning at NASA: The Challenger and Columbia accidents*. Washington, DC: Georgetown University Press.

Maurino, D.E., Reason, J.T. et al. (1995). *Beyond aviation human factors*. Aldershot, UK: Avebury Aviation.

McLean, B. and Elkind, P. (2004). *The smartest guys in the room: The amazing rise and scandalous fall of Enron*. New York: Portfolio.

Merry, A.F. and McCall Smith, A. (2001). *Errors, medicine and the law.* Cambridge, UK: Cambridge University Press.

Michaelides-Mateou, S. and Mateou, A. (2010). *Flying in the face of criminalization*. Farnham, UK: Ashgate.

Montgomery, K. (2006). *How doctors think: Clinical judgment and the practice of medicine*. Oxford, UK, Oxford University Press.

Morath, J.M. and Turnbull, J.E. (2005). *To do no harm: Ensuring patient safety in health care organizations*. San Francisco, CA: Jossey-Bass.

Moray, N. and Huey, B. (1988). *Human factors research and nuclear safety*. Washington, DC: National Academy Press.

Moshansky, V.P. (1992). Final report of the commission of inquiry into the Air Ontario crash at Dryden, Ontario. Ottawa, Canada: Ministry of Supply and Services.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco, CA: W.H. Freeman.

Nevile, M. (2004). *Beyond the black box: Talk-in-interaction in the airline cockpit*. Aldershot, UK: Ashgate.

Norman, D.A. (1988). *The psychology of everyday things*. New York: Basic Books.

Norman, D.A. (1993). *Things that make us smart: Defending human attributes in the age of the machine*. Reading, MA: Addison-Wesley Pub. Co.

Norman, D.A. (1998). *The invisible computer: Why good products can fail, the personal computer is so complex, and information appliances are the solution*. Cambridge, MA: MIT Press.

Norros, L. (2004). *Acting under uncertainty: The core-task analysis in ecological study of work*. Finland: VTT.

Noy, Y.I. and Karwowski, W. (2005). *Handbook of human factors in litigation*. Boca Raton, FL: CRC Press.

O'Hare, D. and Roscoe, S.N. (1990). *Flightdeck performance: The human factor*. Ames, IA: Iowa State University Press.

Owen, C., Béguin, P. et al. (2009). *Risky work environments: Reappraising human work within fallible systems*. Farnham, UK: Ashgate.

Perrow, C. (1984). *Normal accidents: Living with high-risk technologies*. New York: Basic Books.

Pink, D.H. (2009). *Drive: The surprising truth about what motivates us*. New York: Riverhead Books.

Pronovost, P.J and Vohr, E. (2010). *Safe patients, smart hospitals*. New York: Hudson Street Press.

Rasmussen, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. New York: North-Holland.

Reason, J.T. (1990). *'Human error'*. New York: Cambridge University Press.

Reason, J.T. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate.

Reason, J.T. and Hobbs, A. (2003). *Managing maintenance error: A practical guide*. Aldershot, UK: Ashgate.

Roed-Larsen, S., Stoop, J. et al. (2005). *ESReDA shaping public safety investigations of accidents in Europe*. Hovik, Norway: Det Norske Veritas.

Rogers, W.P. et al. (1986). Report of the Presidential Commission on the Space Shuttle Challenger Accident. Washington, DC.

Rosness, R., Guttormsen, G. et al. (2004). *Organisational accidents and resilient organizations: Five perspectives (Revision 1)*. Trondheim, Norway: SINTEF Industrial Management.

Salas, E. and Maurino, D.E. (Eds). (2010). *Human factors in aviation*. San Diego, CA: Academic Press.

Scott, J.C. (1998). *Seeing like a State: How certain schemes to improve the human condition have failed*. New Haven, CT: Yale University Press.

Sharpe, V.A. (Ed.). (2004). *Accountability: Patient safety and policy reform*. Washington, DC: Georgetown University Press.

Simon, H.A. (1957). *Models of man: Social and rational; mathematical essays on rational human behavior in a social setting*. New York: Wiley.

Simon, H.A. (1981). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Snook, S.A. (2000). *Friendly fire: The accidental shootdown of US Black Hawks over Northern Iraq*. Princeton, NJ: Princeton University Press.

Stanton, N. and Edworthy, J. (1999). *Human factors in auditory warnings*. Aldershot, UK; Brookfield, VT: Ashgate.

Starbuck, W.H. and Farjoun, M. (2005). *Organization at the limit: Lessons from the Columbia disaster*. Malden, MA: Blackwell Pub.

Stokes, A. and Kite, K. (1994). *Flight stress: Stress, fatigue, and performance in aviation*. Aldershot, UK: Avebury Aviation.

Suchman, L.A. (1987). *Plans and situated actions: The problem of human-machine communication*. New York: Cambridge University Press.

Townsend, A.S. (2013). *Safety can't be measured*. Farnham, UK: Gower.

Tuchman, B.W. (1981). *Practicing history: Selected essays*. New York: Knopf.

Tufte, E.R. (1990). *Envisioning information*. Cheshire, CT: Graphics Press.

Turner, B.A. (1978). *Man-made disasters*. London: Wykeham Publications.

Varela, F.J, Thompson, E. et al. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.

Vaughan, D. (1996). *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. Chicago, IL: University of Chicago Press.

Vette, G. (1984). *Impact erebus*. New York: Sheridan House Inc.

Vicente, K.J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Mahwah, NJ: Lawrence Erlbaum Associates.

Vincent, C. (2006). *Patient safety*. London: Churchill Livingstone.

Wachter, R.M. (2008). *Understanding patient safety*. New York: McGraw-Hill.

Wagenaar, W.A. (2006). *Vincent plast op de grond: Nachtmerries in het Nederlands recht (Vincent urinates on the ground: Nightmares in Dutch law)*. Amsterdam: Uitgeverij Bert Bakker.

Weick, K.E. and Sutcliffe, K.M. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty*. San Francisco, CA: Jossey-Bass.
Wickens, C.D. (1992). *Engineering psychology and human performance*. New York: HarperCollins Publishers.
Wiener, E.L. and Nagel, D.C. (Eds). (1988). *Human factors in aviation*. Academic Press Series in Cognition and Perception. San Diego, CA: Academic Press.
Wildavsky, A.B. (1988). *Searching for safety*. New Brunswick: Transaction Books.
Wilkinson, I. (2001). *Anxiety in a risk society*. New York: Routledge.
Woods, D.D, Dekker, S.W.A. et al. (2010). *Behind 'human error'*. Aldershot, UK: Ashgate.
Woods, D.D. and Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. Boca Raton, FL: CRC/Taylor & Francis.
Zimbardo, P. (2008). *The Lucifer effect: Understanding how good people turn evil*. New York: Random House.

This page has been left blank intentionally

Epilogue: Speaking for the Dead

"What you're doing, Sidney," Jim Reason said, looking at me intently, "is trying to crawl into the skull of a dead man."

He turned to gaze at the teacup in his hand, shook his head, and grumbled. "How is that possible?"

The way he looked, it was not a question.

We were in a side-discussion during a conference more than a decade ago. The topic was (what is formally known as) process-tracing methods. I had attempted to popularize these methods in the first version of the *Field Guide* in 2002, and have done so in subsequent versions of it. How far can we take these methods to understand why it made sense for practitioners to do what they did—if those practitioners are no longer around? Is it possible to "reconstruct the mindset" as I had called it at the time?

A broad issue in the New View is at stake in this question. If we want to learn from and about the human role in system success and failures, we need to study what they do, or did. Process-tracing methods are part of a larger family of cognitive task analysis, but aim specifically to analyze how people's understanding evolved in parallel with the situation unfolding around them during a particular problem-solving episode.[1] Process-tracing methods are extremely useful, if not indispensable, in the investigation of incidents and accidents. A scenario that leads to these develops both autonomously and as a result of practitioners' management of the situation, which gives it a particular direction toward an outcome. Typically, process-tracing builds two parallel accounts of the problem-solving episode: one in the context-specific language of the domain, and one in concept-dependent terms. These two interlink: It is in and through the latter that we can discover or recognize (even from other domains) regularities in the particular performance of practitioners during the episode. What we look at is not a meaningless mass of raw data, but a set of patterns of human performance. The context-specific or domain-dependent particulars make sense through the concepts applied to them. What looks like a flurry of changing display indications and confused questions about what the technology is doing, for example, can be made sense of by reading into them the conceptual regularities of automation surprise.[2] At the same time, the

conceptual can only be seen in the particulars of practitioners' performance; its regularities begin to stand out across different instances of it.

The question is: does an authentic representation of the process require us to engage directly with the practitioners who were there? Can process-tracing speak for the dead? Despite Jim Reason's remark, I have always been encouraged by those who use a process-tracing method and who are careful about the reach of the claims this method makes, and who are willing to expose and acknowledge the constraints that make these claims possible. A dismissal about "crawling into a dead man's skull" has long felt unrepresentative of what we try to do.

Speaking for the Dead, or Speaking about the Dead?

The alternative of speaking for the dead, after all, is to speak about the dead. Lots of people do this in the aftermath of accidents—from media to colleagues to investigation boards. Speaking about the dead is apparently easy; much easier than speaking for the dead. And it can quickly turn into speaking *badly* about the dead. Here is the voice of an investigation board, involved because the aircraft had been built in their country. It spoke about a crash whose investigation I assisted a few years ago. The board concluded that the crash resulted from a *poorly managed*, non-stabilized approach that resulted in a high crew workload during a critical phase of flight; the crew's subsequent *failure* to detect and respond to the decreasing airspeed and impending onset of the warnings during the final approach despite numerous indications; the crew's *failure* to abort the unstabilized approach and initiate a go-around; and their *failure* to respond to warnings in a timely manner which resulted in a stall and subsequent ground impact. It is all about failures and poor work.

The Old View really has efficiency and expediency going for it. You don't have to do a lot of analytic work, or put in the hard investigative yards. Just say where other people went wrong.

Linearity and hindsight offer those who speak about the dead plenty to go on. Find the series of mistakes, line them up and show the way to the bad outcome. It is so easy. The Old View really has efficiency and expediency going for it. You don't have to do a lot of analytic work, you don't have to put in the hard investigative yards. Just say where you judge other people to have gone wrong (because you now know what the right thing was, given hindsight). The vocabulary that remains is organized around human deficits. It is about

their shortcomings, their delays and failures, their poor management and poor decisions. All this highlights, from the human remains, is what they did not do. In the crash I mentioned above, this was about three human lives, rich with experiences, aspirations, intentions, hopes and dreams (and the assumption that they were going to turn around the airplane and fly back, just like any other work day). Their lives were formally reduced to a short paragraph of deficiencies. This is speaking badly about the dead.

There are at least two strong ethical reasons why we should deploy all that our field has to offer to fight this Old View. The first reason is that the dead are not just dead. They were, and in a sense always will be, colleagues, sons, mentors, daughters, fathers, bosses, mothers, friends. Recall the runway accident mentioned in the Preface. The crew, the investigation board wrote, had lost positional awareness, had failed to follow procedures, had violated regulations, had failed to maintain operational discipline. A student of mine was flying as commander for the same airline at the time, and lost one of his colleagues in that terrible accident. He was also detached as investigator into the accident. In the months and years following the crash, he fought to stop the board from speaking badly *about* the dead. He did all he could to get them to use a New View to build candidate explanations that could speak *for* the dead. The board would not listen.

The alternative of speaking for the dead is speaking about the dead. Investigations often speak badly about the dead.

My student's colleague had left behind a widow and two little children—who, he hoped, would never read the ugly, demeaning official report about their dead father who had gone off to work one day and had not come back. "Go meet the widow," I said to my student. "I am convinced she'd really appreciate it." He eventually did. And she did. The response was emotional and unremittingly positive. Today, a feature film is showing that covers the experience.[3] It, too, speaks for the dead—and the few remaining living.

The second reason is that when we speak for the dead, we do not just speak for the dead or their loved ones. The New View tries to speak for practitioners—dead *or* alive—because it is able to illuminate the constraints and obligations of their daily existence, the patterns of cognition and performance, the regularities of their interaction with devices, systems, organizations. If what the dead did made sense to them at the time, it will likely make sense to the living as well. That means that the ethical commitment to speak for the dead irradiates beyond the dead. And it applies to the future at least as much as it might to the past. Learning from the past is a matter of abstracting away from its contextual mess, ambiguities and indeterminacies, to begin to

see conceptual regularities. These we can take into the future. Automation surprises, error-intolerance—such concepts allow us to explain, predict, to help guide design, to perhaps prevent. Indeed, the very strength of our field lies in its ability to conceptualize, to see regularities at levels of resolution where others see only 'human error.' Speaking for the dead speaks as loudly for them as it does for everyone who could have been in their shoes. It speaks for the past as much as for the future.

Of course, we might throw up our hands and admit defeat; say that we cannot crawl into the skull of a dead man; concede that time is irreversible and that reconstructing somebody's mindset is impossible. But that does not mean that there is nothing systematic about what people do in interaction with complex systems, about why they do what they do. It does not mean that everything about their mind and the world in which its understanding unfolded is unintelligible, inexplicable. It does not mean that there is nothing predictable about how things go right or wrong. And it does not release us from the ethical responsibility to at least try. If all there is in the past is dead people, whose intentions and machinations are forever closed off to our analysis and attempts at understanding, then we have little to say about the future too.

Notes

1 Woods, D.D. Process-tracing methods for the study of cognition outside of the experimental laboratory. In: Klein, G.A., Orasanu, J.M., Calderwood, R., Zsambok, C.E., editors. *Decision making in action: Models and methods* Norwood, NJ: Ablex, 1993:228–51.

2 Sarter, N.B., Woods, D.D., Billings C. Automation surprises. In: Salvendy, G., editor. *Handbook of human factors/ergonomics*. New York: Wiley, 1997:1926–1943.

3 Dickens, K. *Sole Survivor*. USA: Chicago Filmmakers, 2012.

Index

Bold page numbers indicate figures, *italic* numbers indicate tables.