2

FEATURES OF DRIFT

❖

In the early afternoon of January 31, 2000, Alaska Airlines flight 261, a McDonnell Douglas MD-80 took off from Puerto Vallarta in Mexico, bound for Seattle. The pilots had just taken over the airplane from the incoming crew, who had nothing special to report about the status of the airplane.[1]

A bit into the flight, the pilots contacted the airline's dispatch and maintenance control facilities in Seattle on the radio. This was a shared company radio frequency between Alaska Airlines' dispatch and maintenance facilities at Seattle and its operations and maintenance facilities at Los Angeles International Airport (LAX). They had run into a pretty serious problem: the horizontal stabilizer, which helps control the aircraft's nose attitude while in flight, appeared to be jammed.

"Understand you're requesting diversion to LA?" Seattle maintenance asked the pilots at 3.50 p.m. "Is there a specific reason you prefer LA over San Francisco?"

"Well a lot of times it's windy and rainy and wet in San Francisco and uh, it seemed to me that a dry runway, where the wind is usually right down the runway seemed a little more reasonable."

A few minutes later, a dispatcher from Seattle provided the flight crew with the current San Francisco weather. The wind was light, out of the south (180 degrees), and the visibility was good (9 miles). But, the dispatcher added, "If you want to land at LA of course for safety reasons we will do that, we'll tell you though that if we land in LA, we'll be looking at probably an hour to an hour and a half we have a major flow program going right now," referring to air traffic control restrictions that would make it hard to get the aircraft out of Los Angeles again.

"I really didn't want to hear about the flow being the reason you're calling us, because I'm concerned about over flying suitable airports," the captain replied. He did, however, discuss with his first officer potential landing runways at SFO, and finding a discrepancy: "One eight zero at six … so that's runway one six what we need is runway one nine, and they're not landing runway one nine."

"I don't think so," the first officer replied.

The captain then asked Seattle dispatch if they could "get some support" or "any ideas" from an instructor to troubleshoot the problem.

He received no response.

"It just blows me away," the captain then said to his first officer, "they think we're gonna land, they're gonna fix it, now they're worried about the flow. I'm sorry, this airplane's not gonna go anywhere for a while. So you know."

"So they're trying to put the pressure on you," a flight attendant replied.

"Well, no. Yeah."

The Seattle dispatcher had not given up on San Francisco. He informed the flight crew a few minutes later that the landing runways in use at SFO were 28R and 28L and that "it hasn't rained there in hours so I'm looking at … probably a dry runway."

The captain replied that he was waiting for a requested center of gravity (CG) update (for landing), and then he requested information on wind conditions at Los Angeles. The dispatcher replied that the wind at LAX was out of the west (260°) at 9 knots.

Nine seconds later, the captain, comparing SFO and LAX wind conditions, told the SEA dispatcher, "versus a direct crosswind which is effectively no change in groundspeed … I got to tell you, when I look at it from a safety point I think that something that lowers my groundspeed makes sense."

"That'll mean LAX then for you," the dispatcher replied. He then asked the captain to provide LAX operations with the information needed to re-compute the airplane's CG because "they can probably whip out that CG for you real quick."

"We're going to LAX," the captain then said. "We're gonna stay up here and burn a little more gas, get all our ducks in a row, and then we'll be talking to LAX when we start down to go in there."

It was now almost four in the afternoon.

Turning to LAX operations, the captain asked if they could "compute [the airplane's] current CG based on the information we had at takeoff," and asked them once again for the latest weather information at San Francisco. The weather was basically unchanged.

"That's what I needed. We are coming in to see you," the captain then told LAX operations. The first officer began giving LAX operations the information it needed to re-compute the airplane's CG for landing.

Then, at 4.07 p.m., a mechanic at Alaska Airlines' LAX maintenance facility contacted the flight crew on the company radio frequency and asked, "Are you [the] guys with the uh, horizontal stabilizer situation?"

"Affirmative," the captain replied.

The mechanic, referring to the stabilizer's primary trim system, asked, "Did you try the suitcase handles and the pickle switches?"

"Yeah. We tried everything together," the captain said, adding: "We've run just about everything. If you've got any hidden circuit breakers we'd love to know about them."

"OK, I'll look at the circuit breaker guide," the mechanic replied, "just as a double check." He then asked the flight crew about the status of the alternate trim system.

"It appears to be jammed … the whole thing. The AC load meter spikes out when we use the primary, we get AC [electrical] load that tells me the motor's trying to run but the brake won't move it. When we use the alternate, nothing happens."

"You say you get a spike … on the meter up there in the cockpit when you try to move it with the primary, right?"

In the meantime, the captain had turned to the first officer in the cockpit. "I'm going to click it off," he said, talking about the autopilot. "You got it."

"Okay," the first officer said.

Then the captain reiterated to the LAX mechanic that the spike occurred "when we do the primary trim but there's no appreciable change in the electrical when we do the alternate."

The mechanic, not able to figure things out from these reports from a distance, replied that he would see them when they arrived at the LAX maintenance facility.

"Lets do that," the captain said.

The autopilot was still on at this time, and the captain said to the first officer: "This will click it off." According to the Flight Data Recorder (FDR) data, the autopilot was then disengaged at 1609:16.

At the same time, the cockpit voice recorder (CVR) recorded the sound of a clunk, followed by two faint thumps in short succession at 1609:16.9. The CVR recorded a sound similar to the horizontal stabilizer-in-motion tone.

"You got it?" the captain asked.

The autopilot was disengaged again and the airplane began to pitch nose down, starting a dive that lasted about 80 seconds as the airplane fell from about 31,000 to between 23,000 and 24,000 feet.

"It got worse," the captain managed to say. "You're stalled."

Airframe vibration was getting louder now.

"No, no, you've got to release it, you've got to release it."

A click sounded.

"Help me back, help me back."

"Okay," the first officer responded.

"Center, Alaska two sixty one we are, ah, in a dive here," the captain told the air traffic controller. "And I've lost control, vertical pitch."

The overspeed warning came on, and continued for the next 33 seconds.

"Alaska 261, say again?" the controller asked in disbelief.

"Yeah, we're out of twenty-six thousand feet, we are in a vertical dive … not a dive yet … but uh we've lost vertical control of our airplane." Not much later, the captain added, "Just help me."

For a moment, things seemed to get better. "We're at twenty-three seven, request, uh … " And then the captain added, "Yeah, we got it back under control here."

The first officer didn't think so. "No we don't," he immediately told the controller. Turning to the captain, he suggested: "Let's take the speedbrakes off."

"No, no leave them there. It seems to be helping." The speedbrakes, which come out of the top of the wings to help get rid of airspeed, were still out as a response to the overspeed from the first dive.

"Okay, it really wants to pitch down," the captain said. Transmitting to the controller, the captain said that they were at "Twenty-four thousand feet, kind of stabilized."

Three seconds later, he added, "We're slowin' here, and uh, we're gonna uh do a little troubleshooting, can you give me a block [altitude] between uh, twenty and twenty-five?"

It was 4.11 p.m … The airplane's airspeed had decreased to 262 knots, and the airplane was maintaining an altitude of approximately 24,400 feet with a pitch angle of 4.4°. The controller assigned flight 261 a block altitude of between FL 200 and 250. Between about 130 and 140 pounds of pulling force was required to recover from the dive and keep the aircraft level.

"Whatever we did is no good," the first officer said. "Don't do that again."

"Yeah, no, it went down. It went to full nose down."

"It's a lot worse than it was?"

"Yeah, we're in much worse shape now. I think it's at the stop, full stop. And I'm thinking, can it go any worse? But it probably can, but when we slowed down … Lets slow it, lets get down to two hundred knots and see what happens."

The captain turned to LAX maintenance again. "We did both the pickle switch and the suitcase handles and it ran away full nose trim down. And now we're in a pinch. So we're holding, we're worse than we were." The captain explained that he was reluctant to try troubleshooting the trim system again because the trim might "go in the other direction."

"Okay, well, your discretion. If you want to try it, that's okay with me if not that's fine. Um, we'll see you at the gate."

"I went tab down … right?" the captain explained to the mechanic, "and it should have come back instead it went the other way."

"You want to try it or not?" the captain then asked the first officer.

"Uh no. Boy, I don't know."

About 120 pounds of pulling force was being applied to the pilots' control columns at this point, just to keep the airplane level.

At a quarter past four, the Los Angeles controller instructed the flight crew to contact another controller on a different radio frequency, which the flight crew acknowledged.

"We're with you, we're at twenty-two five," the first officer radioed to the new controller, "we have a jammed stabilizer and we're maintaining altitude with difficulty. Uh, but, uh we can maintain altitude we think … Our intention is to land at Los Angeles."

The controller cleared the airplane direct to LAX and asked, "You want a lower altitude now, or what do you want to do sir?"

The captain replied this time. "I need to get down about ten , change my configuration, make sure I can control the jet and I'd like to do that out here over the bay if I may."

The Los Angeles ARTCC controller issued flight 261 a heading of 280° and cleared the flight to descend to 17,000 feet.

"Two eight zero and one seven, seventeen thousand, Alaska two sixty-one," the captain said. "And we generally need a block altitude."

"Okay," the controller said, "and just, um, I tell you what – do that for now sir, and contact LA center on one three five point five. They'll have further instructions for you sir."

"Okay," the first officer said to the new controller, "thirty-five five, say the altimeter setting."

"The LA altimeter is three zero one eight."

"Thank you," the first officer said.

This would be the last radio call made from flight 261.

After the radio transmission, the captain told a flight attendant that he needed "everything picked up" and "everybody strapped down." He told the first officer that he was going to stop pulling so hard on the control column: "I'm going to unload the airplane and see if we can … we can regain control of it that way."

The flight attendant came back with bad news. "Okay, we had like a big bang back there."

"Yeah," the captain said. "I heard."

He turned to his first officer again. "I think the stab trim thing is broke."

He then repeated to the flight attendant to make sure the passengers were "strapped in now," adding "because I'm going to release the back pressure and see if I can get it … back."

"Give me slats extend," he asked the first officer.

It was eighteen minutes past four.

"I'm test flying now," the captain said.

More flaps were ordered, down to 11 degrees.

"Its pretty stable right here … see? But we got to get down to a hundred and eighty knots."

Not much later, he asked the first officer, "Okay, bring the flaps and slats back up for me," The first officer complied.

"What I want to do … is get the nose up … and then let the nose fall through and see if we can stab it when it's unloaded."

"You mean, use this again?" the first officer asked. "I don't think we should … if it can fly."

"It's on the stop now, it's on the stop," the captain answered.

"Well, not according to that, it's not," the first officer said, adding, "the trim might be, and then it might be uh, if something's popped back there … it might be mechanical damage too. I think if it's controllable, we ought to just try to land it."

"You think so? Okay, lets head for LA."

About 5 seconds later, the CVR recorded four distinct thumps.

"You feel that?" the first officer asked.

"Yeah. Okay, give me slats."

Slats were extended again.

Then an extremely loud noise sounded, background noise started increasing and loose articles started moving all through the cockpit.

The airplane pitched down violently, at a rate of nearly 25° per second. The airplane started rolling left wing down, and the rudder was deflected back to the right.

The nose was now pointing almost straight down, at 70° and the aircraft kept rolling.

"Mayday," the first officer said, but did not manage to push the transmit button.

"Push and roll, push and roll," the captain said.

The dive diminished, but the airplane was now upside down, and had descended to 16,420 feet. The airspeed had decreased to 208 knots.

"Okay," the captain said, "we are inverted … and now we got to get it."

Using rudder and ailerons, he tried to get the airplane upright again.

"Push, push, push … push the blue side up," the captain counseled. "Okay, now lets kick rudder … left rudder, left rudder."

"I can't reach it," the first officer said.

"Okay, right rudder … right rudder. Got to get it over again." Then, acknowledging that they were now not diving as badly, he added, "at least upside down we're flying."

But not for much longer. With the airflow to them considerably disturbed, both engines now stalled and spooled down.

"Speedbrakes," the captain said.

"Got it," the first officer answered.

There was no more space for any recovery, and the aircraft would not have allowed it in any case.

"Ah, here we go," the captain concluded.

Twenty-one minutes past four, the airplane hit the Pacific Ocean near Port Hueneme, California. It came apart on impact and most of it disappeared below the surface. Pieces of the airplane wreckage were found floating on and beneath the surface of the ocean. The two pilots, three cabin crewmembers, and 83 passengers on board were all killed, and the airplane was destroyed by impact forces.[2]

The pilots had been confronted with one of the scariest experiences possible: an airplane that is no longer controllable. Whatever they did with their controls, it was to no avail. After a period of uncertainty and trouble-shooting, the airplane had suddenly plunged down, straight toward the ocean.

In the two weeks following the accident, remote-operated vehicles and sonar equipment was used to search the debris and load up and tow the large pieces of wreckage to the surface. A commercial trawler was used to drag the ocean bottom and gather the final smaller pieces. Eighty-five percent was recovered. It was all brought to Port Hueneme.

THE BROKEN PART

As the investigation examined airplane parts from the sea floor and matched the wreckage with data traces from the cockpit voice recorder and the flight data recorder, the culprit became obvious: the jackscrew-nut assembly that holds the horizontal stabilizer had failed, rendering the aircraft uncontrollable. The broken part had been found.

Like (almost) all aircraft, the MD-80 has a horizontal stabilizer (or tailplane, or small wing) at the back that helps direct the lift created by wings. It is this little tailplane that controls the aircraft's nose attitude: without it, controlled flight is not possible. The tailplane of the MD-80 controls nose attitude in two ways. At the back end of it, there's a control surface (the elevator) that connects directly to the pilots' control yokes in the cockpit. Pull the yoke in the cockpit back, the elevator angles up and the airplane's nose moves up in return.

The story of Alaska 261, however, is about another part of the horizontal stabilizer. The whole stabilizer itself can angle up or down in order to trim the nose up or down. As fuel is used up during the flight, or wing flaps are extended and retracted, or as a catering trolley moves up and down the aisle inside the aircraft during flight, the flight characteristics of the airplane change (for example, its center of gravity or the center of its lifting force shifts). To prevent the airplane from pitching down or up with changes in the center of gravity and center of lift, it needs to be able to trim. That's the role of the moving stabilizer. The stabilizer, that is, the entire horizontal tail, is hinged at the back, and the front end arcs up or down.

Here's how this is accomplished mechanically. Pushing the front end of the horizontal stabilizer up or down is done through a rotating jackscrew and a nut. The whole assembly works a bit like a carjack used to lift a vehicle, for example when changing a tire. You swivel, and the jackscrew rotates, pulling the so-called acme nuts inward and pushing the car up. In the MD-80 trim system, the front part of the horizontal stabilizer is connected to a nut that drives up and down a vertical jackscrew. An electrical trim motor rotates the jackscrew, which in turn drives the nut up or down. The nut then pushes the whole horizontal tail up or down.

On the 31st of January, 12 minutes after take-off, passing through 23,400 feet, the horizontal stabilizer had moved for the last time until the airplane's initial dive 2 hours and 20 minutes later.

Adequate lubrication is critical for the continued functioning of a jackscrew and nut assembly. Without enough grease, the constant grinding will wear out the thread on either the nut or the screw (in this case the screw is deliberately made of harder material, wearing the nut out first). The thread actually carries the entire load that is imposed on the vertical tail during flight. This is a load of around 5,000 pounds, similar to the weight of a whole family van hanging by the thread of a jackscrew and nut assembly. Were the thread to wear out on an MD-80, the nut would fail to catch the threads of the jackscrew. Aerodynamic forces then push the horizontal tailplane (and the nut) to its stop way out of the normal

range, rendering the aircraft uncontrollable in the pitch axis, which is essentially what happened to Alaska 261. Even the stop failed because of the pressure. A so-called torque tube runs through the jackscrew in order to provide redundancy (instead of having two jackscrews, like in the preceding DC-8 model). But even the torque tube failed in Alaska 261.

On the surface, the accident seemed to fit a simple category: mechanical failure as a result of poor maintenance. A single component failed because people did not maintain it well. It had not been lubricated sufficiently. This led to the catastrophic failure of a single component. The break instantly rendered the aircraft uncontrollable and sent it plummeting into the Pacific.

But such accidents do not happen just because something suddenly breaks. There is supposed to be too much built-in protection against the effects of single failures. Other things have to fail too. More has to go wrong. And indeed, the investigation found more broken components in the system.

Closest to the flight crew, it found that there was no suggestion in any checklist that the flight crew should divert to the nearest possible airport when they got the first indications of horizontal stabilizer trouble. In fact, they found that the use of the autopilot with a jammed stabilizer was "inappropriate" and that a lack of guidance on how to fly an airplane with a jammed stabilizer could lead crews to experimenting and improvising, possibly making the situation worse.

As for the lubrication of the jackscrew, the investigation determined that the access panel in the tail plane of this aircraft type was really too small to adequately perform the lubrication task. Also, there had been widespread deficiencies in Alaska Airlines' maintenance program, leading, for example to a lack of adequate technical data to demonstrate that extensions of the lubrication interval would not present a hazard. There was also a lack of a task-by-task engineering analysis and justification in the process by which manufacturers revise recommended maintenance task intervals and by which airlines establish and revise these intervals.

Coupled to this, the investigation concluded that the end-play check interval (which measures how much play or slack there is in the screw/nut assembly) was inadequate. A restraining fixture for end-play checks was used even though it did not meet aircraft manufacturer specifications. In addition, the so-called on-wing end-play check procedure was never validated and was known to have low reliability. There was no requirement to record or inform customers that allowed overhauled jackscrew assemblies back onto airplanes that had higher end play than expected.

Finally, the investigation noted shortcomings in regulatory oversight by the Federal Aviation Administration, and an aircraft design that did not account for the loss of the acme nut threads as a catastrophic single-point failure mode.

This is the logic that has animated our understanding of failure for a long time now. Failure leads to failure. In order to explain a broken component (the jackscrew), we need to look for other broken components (the checklist, the access panel, the stress tests, the maintenance program, the company's oversight: all of them broken in their own way). In fact, any upbringing in Western science

and engineering hardly allows people to think any other way. The first part of this book will explore the basis for this in Western thinking: the way in which our understanding of the world is rooted in the ideas of René Descartes and Isaac Newton, as well as many others who extended and popularized their ideas.

UNANSWERED QUESTIONS

But the idea that failure can be explained as looking for other failures leaves important questions unanswered. Questions, actually, that may stand between us and a firmer, further understanding of safety in complex systems. This is where the Newtonian–Cartesian basis of traditional human factors and systems safety is most limited, and limiting. Take, for example, the observation by the manufacturer that they knew of no reported cases of fatigue failures or fatigue damage on this torque tube design for their four basic models, the DC-9, MD-80, MD-90 and 717. No fatigue problems reported in over 95 million accumulated flight hours, and in over 2,300 airplanes delivered.[3] That would suggest that nothing was broken – neither in the parts, nor in the organizational processes surrounding their maintenance and approval.

A first question is quite simple. *Why*, in hindsight, do all these other parts (in the regulator, the manufacturer, the airline, the maintenance facility, the technician, the pilots) appear suddenly "broken" now? How is it that a maintenance program which, in concert with other programs like it never revealed any fatigue failures or fatigue damage after 95 million accumulated flight hours, suddenly became "deficient?" The broken parts are easy to discover once the rubble is strewn before our feet. But what exactly does that tell us about processes of erosion, of attrition of safety norms, of drift toward margins? What does the finding of broken parts tell us about what happened before all those broken parts were laid out before us? Our metaphors in safety, mostly driven by Newtonian–Cartesian thinking, can show us how things in an organization ended up. Not how they got there. The metaphors are mostly metaphors for resulting forms (for example, broken links between organizational compartments, layers of defense with holes), not models that capture processes of formation. There are few workable models that can even begin to capture the organic, jumbled, living interiors of the socio-technical organizations that do our risky business.

For more than 20 years now, safety researchers and practitioners have been talking about systems approaches, or systemic approaches. As a result, the discussions that ensued, the reports that got written, and models that were developed, have generally taken the context surrounding component failures more seriously. We now regularly look for sources of trouble in the organizational, administrative, and regulatory layers of the system, not just the operational or engineered sharp-end. The shift from sharp-end failures and operator error to the blunt end and organizational factors is certainly not complete, but at least legitimate today.

The Alaska investigation is testimony to this. It doesn't stop at the sharp end (either the cockpit or the maintenance shop). It goes all the way up through

the blunt organizational end and beyond, to the regulatory authority. But it still returns with broken components. Which is the problem. That is not system thinking. System thinking is about relationships, not parts. System thinking is about the complexity of the whole, not the simplicity of carved-out bits. Systems thinking is about non-linearity and dynamics, not about linear cause-effect-cause sequences. Systems thinking is about accidents that are more than the sum of the broken parts. It is about understanding how accidents can happen when no parts are broken, or no parts are seen as broken.

Which produces a second question, perhaps even more fascinating. Why did none of these deficiencies strike anybody as deficiencies at the time? Or, if somebody did note them as deficiencies, then why was that voice apparently not sufficiently persuasive? If things really were as bad as we can make them look post-mortem, then why was everybody, including the regulator – tasked with public money to protect safety – happy with what was going on? Happy enough, in any case, to not intervene? You see, we create a huge problem for ourselves when we call these after-accident discoveries by all kinds of normative names, like deficiency, or inadequacy, or shortcoming. Inadequate or deficient or short relative to *what*? Clearly, people must have seen the norms that ruled their assessments and their decisions at the time as quite acceptable, otherwise pressure would have built for changing those norms. Which didn't happen.

These are questions at the heart of the future of safety. Behind the broken part onboard Alaska 261 lay a vast landscape of things that weren't all that broken, or not seen as broken at the time. There had been organizational trade-offs and decisions that all seemed quite normal, there was deregulation and increasing competitive pressure that was normal because it operated on every company. There were changes in regulations and oversight regimes. Which there always are. There were underspecified technical procedures and continuous quality developments in aircraft upkeep, which are normal because procedures are always underspecified and changes in how to conduct maintenance are always ongoing – including the routine extension of maintenance intervals. There had been collective international shifts in aircraft maintenance practices, following the kind of cross-national public-private initiatives that mark a global industry. And all this interacted with what one company, and its regulator, saw as sensible and safe.

One way to look at these influences, and the accident they eventually helped fashion, is as a story of drift: drift into failure. The story of drift is just that: a *story*, a way of telling things, of weaving them into a coherent narrative. The reason for telling it is not to claim a true account of what happened. The reason is to encourage a language, or a perspective on accidents, that is more open to the complexity, the dynamics, erosion and adaptation that marks socio-technical systems, than our thinking about accidents has been so far.

That is what this book invites us to do: to read a story of drift into the period before the accident or incident. To read a story of drift into what is going on inside an organization even without an incident or accident having taken place. Not because it is the correct angle to take – in fact, reading "drift" into what we see means making all kinds of analytical sacrifices. But we make analytical

sacrifices no matter what perspective we take. No, the reason for looking for a story of drift is that it may help us stretch our imagination about safety and risk – beyond the components, beyond the now, beyond what's broken. Let's look at what such a story might tell us.

THE OUTLINES OF DRIFT

One of the greatest conundrums in safety work is to explain why a slide into disaster, so easy to see and depict in retrospect, was missed by those who inflicted it on themselves. Judging after the fact that people suffered a failure of foresight is easy. All we need to do is plot the numbers, trace out the trajectory, and spot the slide into disaster. But that is when we are standing amid the rubble and looking back. Then it is easy to marvel at how misguided or misinformed people must have been. But why was it that the conditions conducive to an accident were never acknowledged or acted on by those on the inside the system – those whose job it was to not have such accidents happen?

Foresight is not hindsight. There is a profound revision of insight that turns on the present. It converts a once vague, unlikely future into an immediate, certain past.[4] When we look back and we see drift into failure, it is, of course a byproduct of that hindsight. It seems as if the organization was inescapably heading toward the failure, and we almost automatically re-interpret the availability of signs pointing in that direction as really hard to miss. "People who know the outcome of a complex prior history of tangled, indeterminate events, remember that history as being much more determinant, leading 'inevitably' to the outcome they already knew"[5] But this certainty of trajectory and outcome is just ours. There was no trajectory and no outcome for the people involved at the time, or they would have done something about it. This is why I said that we can read a story of drift *into* the data, not out of the data. If we take a particular vantage point, that is. And standing in the aftermath of the failure definitely helps. The vantage point of the retrospective outsider is what helps us see that drift occurred.

But people's situations at the time would have looked very, very different. Sure, it may have offered multiple possible pathways. Some of these pathways or options might have been proven, some recognizable and familiar, and many of them plausible. But they would all have stretched out into the ever denser fog of futures not yet known. The decisions, trade-offs, preferences and priorities that people made, even if seemingly out of the ordinary and immoral after the failure, were once normal and common sense. Just as we like to believe that ours are now.

Drift, then, may be very difficult to recognize as drift *per se* without a bad outcome, without the benefit of hindsight. But there is probably another reason. It has to do with our thinking (and educating students) that complex systems are collections of components that all have different reliabilities. The "human component" is of course on the bottom reliability rung. It has to do with our

belief that redundancy (through automation, procedures, back-up systems, monitoring) is the best way to protect against such unreliability.

Even before a bad outcome has occurred, indications of trouble are often chalked up to unreliable components that need revision, double-checking, more procedural padding, engineered redundancy. And this might all help. But it also blinds us to alternative readings. It deflects us from the longitudinal story behind those indications, it downplays our own role in rationalizing and normalizing signs of trouble. This focus on components gives us the soothing assurance that we can fix the broken parts, shore up the unreliable ones, or sit down together in a legitimate-sounding or expert-seeming committee that decides that the unreliability of one particular part is not all that critical because others are packed around to provide redundancy. And that all will be fine.

So not recognizing drift while it still happening, without an outcome to prove us right, could partly be a result of the language with which we have learned to look at our world. From whatever vantage point – retrospective or not – a different language will help us see different things. That is the point here. A lot has happened in various sciences since Descartes and Newton. There are interesting ideas and concepts that can offer us better insight into the dynamics, relationships, habituation, adaptive capacity, and complexity of the systems we want to understand and control. The previous chapter introduced five concepts that together may characterize drift:

- Scarcity and competition
- Decrementalism, or small steps
- Sensitive dependence on initial conditions
- Unruly technology
- Contribution of the protective structure.

SCARCITY AND COMPETITION

The fuel, the energy, for adaptive organizational behavior (and, eventually, drift) is the simple fact that the organization is never alone. No organization operates in a vacuum. This is a fundamental feature of complexity and of complex systems: they are open to influences from the outside and are in constant transaction with what happens around them. Any organization operates and must try to survive in an environment that has real constraints, such as the amount of capital available, the number of customers reachable, and the qualifications of available employees. There are also hard constraints on how fast things can be built, developed, driven. And, most importantly, there are other organizations that try to do the same thing.

Many industries, aviation among them, are characterized by a dog-eats-dog business mentality in which any moment of weakness or inattention gets exploited immediately by a competitor. Deregulation set this in serious motion in the airline industry. Before the mid-1970s, the airline industry in most Western countries was essentially cartelized. Regulators were granted the authority by governments

to assign certain routes to particular airlines, and to control the fares on these routes. There was a rationing of routes and players, resulting in strict limits on competition. New players had a very difficult time getting into the business, and rarely succeeded for long if they did. This arrangement kept airfares well above those that would be attained in a free market.

The 1978 Airline Deregulation Act in the U.S.A. changed all that, and was soon followed by similar initiatives in other Western countries. Market forces could now determine who could enter the industry, who could fly which routes and for what prices. The regulator at the time, the Civil Aeronautics Board (or CAB) was disbanded and replaced by the Federal Aviation Administration. Interestingly, deregulation became an asynchronous process in most countries: even though airlines were now allowed to compete under market forces, the infrastructure that they used remained in the hands of governments for a long time (airports, airways and the air traffic control system to run it), and even transnational arrangements often constrained heavily who could fly from which country to which. In the wake of deregulation, however, the airline industry expanded its employment by 32 percent, and passenger travel increased by 55 percent. The real cost of travel dropped by about 17 percent in the first decade after deregulation alone, and dropped even further in the ensuing decades.[6]

Deregulation cannot in itself be construed as a safety culprit, of course. Heavily regulated industries do not necessarily have a better safety record (if anything, it may encourage collusion between regulator and industry, particularly if part of the regulator's role is to encourage business development), and the period after deregulation actually saw a steady increase in airline safety.[7] But changing the rules of the game does change what goes on inside a complex system. And complexity theory can easily predict that changing the number of agents will change the dynamics of any complex system; it will affect the speed at which feedback about agents' actions travels and the patterns along which it reverberates. It might even change the way in which success is defined and assured, and also the way in which failure is bred and perhaps no longer recognized.

Jens Rasmussen suggested that work in complex systems is bounded by three types of constraints. There is an economic boundary, beyond which the system cannot sustain itself financially. Then there is a workload boundary, beyond which people or technologies can no longer perform the tasks they are supposed to. And there is a safety boundary, beyond which the system will functionally fail. For Rasmussen, these three constraints, or boundaries, surround the operation of any safety-critical system from three sides. There is no way out, there is only some room for maneuvering inside the space delineated by these three boundaries. Managerial and economic pressure for efficiency will push the system's operations closer to the workload and safety boundaries. Indeed, the likely result of increasing competitive pressure on a system, and of resource scarcity, will be a systematic migration toward workload and safety boundaries.[8]

Noncommercial enterprises experience resource scarcity and Rasmussen's three constraints too. With respect to the Alaska 261 accident, for example, a new

regulatory inspection program, called the Air Transportation Oversight System (ATOS), was put into use in 1998 (two years prior to the accident). It drastically reduced the amount of time inspectors had for actual surveillance activities. A 1999 memo by a regulator field-office supervisor in Seattle suggested how this squeezed their oversight work into a corner where the safety and workload boundaries met:

> We are not able to properly meet the workload demands. Alaska Airlines has expressed continued concern over our inability to serve it in a timely manner. Some program approvals have been delayed or accomplished in a rushed manner at the "eleventh hour," and we anticipate this problem will intensify with time. Also, many enforcement investigations ... have been delayed as a result of resource shortages. [If the regulator] continues to operate with the existing limited number of airworthiness inspectors ... diminished surveillance is imminent and the risk of incidents or accidents at Alaska Airlines is heightened.[9]

Adapting to resource pressure, approvals were delayed or rushed, surveillance was reduced. Yet doing business under pressures of resource scarcity is normal: Scarcity and competition are part and parcel even of doing inspection work. Few regulators anywhere will ever claim that they have adequate time and personnel resources to carry out their mandates. Yet the fact that resource pressure is normal does not mean that it has no consequences. Of course the pressure finds ways out. Supervisors write memos, for example. Battles over resources are fought. Trade-offs are made. The pressure expresses itself in the common organizational, political wrangles over resources and primacy, in managerial preferences for certain activities and investments over others, and in almost all engineering and operational trade-offs between strength and cost, between efficiency and diligence. In fact, working successfully under pressures and resource constraints is a source of professional pride. Being able to create a program that putatively allows better inspections with fewer inspectors may win a civil servant compliments and chances at promotion, while the negative side effects of the program are felt primarily in some far-away field office.

Yet the major engine of drift hides somewhere in this conflict, in this tension between operating safely and operating at all, between building safely and building at all. This tension provides the energy behind the slow, steady disengagement of practice from earlier established norms or design constraints. This disengagement can eventually become drift into failure. As a system is taken into use, it learns, and as it learns, it adapts. A critical ingredient of this learning is the apparent insensitivity to mounting evidence that, from the position of retrospective outsider, could have shown how bad the judgments and decisions actually are. This is how it looks from the position of the retrospective outsider: the retrospective outsider sees a failure of foresight. From the inside, however, the abnormal is pretty normal, and making trade-offs in the direction of greater efficiency is nothing unusual. In making these trade-offs, however, there is a feedback imbalance. Information on whether a decision is cost-effective or efficient can be relatively easy to get. An

early arrival time is measurable and has immediate, tangible benefits. How much is or was borrowed from safety in order to achieve that goal, however, is much more difficult to quantify and compare. If it was followed by a safe landing, apparently it must have been a safe decision. Extending a lubrication interval similarly saves immediately measurable time and money, while borrowing from the future of an apparently problem-free jackscrew assembly. Remember, the manufacturer knew of no reported fatigue problems or failures in 95 million flight hours accumulated by 2,300 airplanes.

Evidence from a feedback imbalance, then, suggests strongly that the system can operate equally safely, yet more efficiently. From the outside, such fine-tuning constitutes incremental experimentation in uncontrolled settings. On the inside, incremental nonconformity is an adaptive response to scarce resources and production goals. This means that departures from the norm become the norm. Seen from the inside people's own work, deviations 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 or clinical requirements, but also deadlines).

As Weick pointed out, however, safety in those cases may not at all be the result of the decisions that were or were not made, but rather an underlying stochastic variation that hinges on a host of other factors, many not easily within the control of those who engage in the fine-tuning process.[10] Empirical success, in other words, is no proof of safety. Past success does not guarantee future safety. Borrowing more and more from safety may go well for a while, but we never know when we are going to hit. This moved Langewiesche to say that Murphy's law is wrong: everything that can go wrong usually goes right, and then we draw the wrong conclusion.[11]

DECREMENTALISM, OR SMALL STEPS

When it first launched the aircraft in the mid 1960s, Douglas recommended that operators lubricate the trim jackscrew assembly every 300 to 350 flight hours. For typical commercial usage, that could mean grounding the airplane for such maintenance every few weeks. Immediately, the sociotechnical, organizational systems surrounding the operation of the technology began to adapt, and set the system on its course to drift. Through a variety of changes and developments in maintenance guidance for the DC-9/MD-80 series aircraft, the lubrication interval was extended. These extensions were hardly the product of manufacturer recommendations alone, if at all. A much more complex and constantly evolving web of committees with representatives from regulators, manufacturers, subcontractors, and operators was at the heart of a fragmented, discontinuous development of maintenance standards, documents, and specifications. Rationality for maintenance-interval decisions was produced relatively locally, relying on incomplete, emerging information about what proved to be, for all its deceiving simplicity, unruly technology. Although each decision was locally rational, making

sense for decision-makers in their time and place, the global picture became one of drift into failure.

Starting from a lubrication interval of 300 hours, the interval at the time of the Alaska 261 accident had moved up to 2,550 hours, almost an order of magnitude more. As is typical in the drift toward failure, this distance was not bridged in one leap. The slide was incremental: step by step, decision by decision that weaved through the aviation system from different angles and directions, alternating cross-industry initiatives with airline-driven ones, but all interacting to push the system into one direction: greater intervals.

In 1985, as an accompaniment of deregulation in the airline industry, jackscrew lubrication was to be accomplished every 700 hours, at every other so-called maintenance B check (which occurs every 350 flight hours).

In 1987, the B-check interval itself was increased to 500 flight hours for the entire industry, pushing lubrication intervals to 1,000 hours.

In 1988, B checks were eliminated altogether, and tasks to be accomplished were redistributed over A and C checks. The jackscrew assembly lubrication was to be done each eighth 125-hour A check: still every 1,000 flight hours.

But in 1991, A-check intervals were extended across the entire industry to 150 flight hours, leaving a lubrication every 1,200 hours. Three years later, the A-check interval was extended again, this time to 200 hours. Lubrication would now happen every 1,600 flight hours.

In 1996, Alaska airlines removed the jackscrew-assembly lubrication task from the A check and moved instead to a so-called task card that specified lubrication every 8 months. There was no longer an accompanying flight-hour limit. For Alaska Airlines, eight months translated to about 2,550 flight hours.

The jackscrew recovered from the ocean floor, however, revealed no evidence that there had been adequate lubrication at the previous interval at all. It might have been more than 5,000 hours since it had last received a coat of fresh grease. Through this drift, the effect of small things could suddenly explode into something huge. Miss one lubrication if you service the part every 350 hours, and you have no worries. Miss one when you lubricate every 2,550 hours, and you could get in deep trouble. This is where decrementalism, or small changes, can lead to big events. And where the way the system is organized (for example, around set lubrication intervals rather than continuous checks and lubrication-as-necessary) can make it highly sensitively dependent on such small changes.

SENSITIVE DEPENDENCY ON INITIAL CONDITIONS

This, in complexity and systems thinking, is called a sensitive dependency on initial conditions (or butterfly effect). Whether technology takes an unruly trajectory toward failure may depend on seemingly innocuous features or infinitesimally small differences in how it all got started.

The original DC-9 was certified in 1965 under Civil Aeronautics Regulations (CAR) 4b from three years prior (1962). The MD-80 was certified in 1980. More recent models of the DC-9, MD-80, MD-90, and Boeing 717 were certified under

14 CFR Part 25 and applicable amendments. However, systems that were similar to or that did not change significantly from the earlier DC-9 models, such as the longitudinal trim control system, were not required to be recertified. CAR 4b remained the certification basis for those parts of the MD-80, MD-90, and 717. This meant that even the most modern of the lineage, the Boeing 717, today has parts in it that are certified in 1965 under rules from 1962, and never needed to be re-certified. Just a time check: these rules were laid down a year before Kennedy was assassinated. When the first DC-9 was certified, we still had four years to go before landing on the moon. These were slide-rule times. No computer-aided design. The effects of, and sensitive dependency upon, such initial conditions are with us in lots of airplanes flying around today.

Understanding how a system may sensitively depend on initial conditions is something that certification processes are supposed to be able to do. But certification does not typically take *lifetime* wear of parts into account when judging (in this case) an aircraft airworthy, even if such wear will render an aircraft, like Alaska 261, quite unworthy of flying. Certification processes certainly do not know how to take socio-technical adaptation of new equipment, and the consequent potential for drift into failure, into account when looking at nascent technologies. Systemic adaptation or wear is not a criterion in certification decisions, nor is there a requirement to put in place an organization to prevent or cover for anticipated wear rates or pragmatic adaptation, or fine-tuning.

As a certification engineer from the regulator testified, "Wear is not considered as a mode of failure for either a system safety analysis or for structural considerations" (NTSB, 2002, p. 24). Because how do you take wear into account? How can you even predict with any accuracy how much wear will occur? McDonnell-Douglas surely had it wrong when it anticipated wear rates on the trim jackscrew assembly of its DC-9. Originally, the assembly was designed for a service life of 30,000 flight hours without any periodic inspections for wear. But within a year, excessive wear had been discovered nonetheless, prompting a reconsideration.

The problem of certifying a system as safe to use can become even more complicated if the system to be certified is sociotechnical and thereby even less calculable. What does wear mean when the system is sociotechnical rather than consisting of pieces of hardware? In both cases, safety certification should be a lifetime effort, not a still assessment of decomposed system status at the dawn of a nascent technology. Safety certification should be sensitive to the co-evolution of technology and its use, its adaptation. Using the growing knowledge base on technology and organizational failure, safety certification could aim for a better understanding of the ecology in which technology is released – the pressures, resource constraints, uncertainties, emerging uses, fine-tuning, and indeed lifetime wear.

Safety certification is not just about seeing whether components meet criteria, even if that is what it often practically boils down to. Safety certification is about anticipating the future. Safety certification is about bridging the gap between a piece of gleaming new technology in the hand now, and its adapted, coevolved, grimy, greased-down wear and use further down the line. But we are not very

good at anticipating the future. Certification practices and techniques oriented toward assessing the standard of current components do not translate well into understanding total system behavior in the future. Making claims about the future, then, often hangs on things other than proving the worthiness of individual parts. Take the trim system of the DC-9 again.

The jackscrew in the trim assembly had been classified as a "structure" in the 1960s, leading to different certification requirements from when it would have been seen as a system. The same piece of hardware, in other words, could be looked at as two entirely different things: a system, or a structure. In being judged a structure, it did not have to undergo the required system safety analysis (which may, in the end, still not have picked up on the problem of wear and the risks it implied). The distinction, this partition of a single piece of hardware into different lexical labels, however, shows that airworthiness is not a rational product of engineering calculation. Certification can have much more to do with localized engineering judgments, with argument and persuasion, with discourse and renaming, with the translation of numbers into opinion, and opinion into numbers – all of it based on uncertain knowledge.

As a result, airworthiness is an artificially binary black-or-white verdict (a jet is either airworthy or it is not) that gets imposed on a very grey, vague, uncertain world – a world where the effects of releasing a new technology into actual operational life are surprisingly unpredictable and incalculable. Dichotomous, hard yes or no meets squishy reality and never quite gets a genuine grip. A jet that was judged airworthy, or certified as safe, may or may not be in actual fact. It may be a little bit unairworthy. Is it still airworthy with an end-play check of 0.0042 inches, the set limit? But "set" on the basis of what? Engineering judgment? Argument? Best guess? Calculations? What if a following end-play check is more favorable? The end-play check itself is not very reliable. The jet may be airworthy today, but no longer tomorrow (when the jackscrew snaps). But who would know? Nobody, really. Because the technology is unruly.

UNRULY TECHNOLOGY

Unruly means that something is disorderly and not amenable to discipline or control. It is a better word than uncertain. Very little to none of the technology we put into our complex systems is believed to be uncertain. Uncertain means unknown. And the whole point of engineering and validation and verification is to calculate our way out of such uncertainty, to not have such unknowns. We run the numbers, do the equations, build the computer simulations, so that we know under what load something will break, what the mean time between failures will be.

But despite all these efforts, there will always be unknowns, even if the technology is considered or judged or believed to be certain. The term "unruly technology" was introduced by Brian Wynne in 1988 to capture the gap between our image of tidiness and control over technology through design, certification, regulation, procedures, and maintenance on the one hand and the messy, not-so-governable interior of that technology as it behaves when released into a field of

practice. Technology and safety is about how things behave in context, not on the drawing board. Universal proclamations or assurances about reliability figures are mute when ideas or designs are put into working solutions in this or that situation. A crucial skill involves finding a practical balance between universality of safety assumptions and their contextualization.[12]

This is a balance (and a possible gap) that not only operates between those on the outside (the public, or consumers of the technology) and insiders (engineers, managers, regulators), but applies even to insiders themselves: to practitioners very close to the technology. If the operational system is not itself following the rules by which it was predicted or supposed to operate, insiders can also reconcile such data with their beliefs – not by changing beliefs, but by looking differently at the data. This works particularly when (1) it is a relatively common problem, among a mass of other relatively common problems, (2) there are routine operational ways of compensating for it (providing more redundancy), and (3) alternative approaches would severely disrupt the economic or operational viability of the system.[13] The stretching of maintenance intervals, even against the background of a technology that behaves different in practice than what was predicted in the drawing, is a relatively common issue. Through these processes, insiders can keep their beliefs and they can retain their image of tidiness and controllability. But the technology may remain unruly, no matter what.

With as much lubrication of the jackscrew assembly as it originally recommended, Douglas thought it had no reason to worry about thread wear. So before 1967, the manufacturer provided or recommended no check of the wear of the jackscrew assembly. The trim system was supposed to accumulate 30,000 flight hours before it would need replacement. But operational experience revealed a different picture. After only a year of DC-9 flying, Douglas received reports of thread wear significantly in excess of what had been predicted. The technology, in other words, refused to play by the manufacturer's rules.

In response, the manufacturer recommended that operators perform a so-called end-play check on the jackscrew assembly at every maintenance C check, or every 3,600 flight hours. The end-play check uses a restraining fixture that puts pressure on the jackscrew assembly, simulating the aerodynamic load during normal flight. The amount of play between nut and screw, gauged in thousandths of an inch, can then be read off an instrument. The play is a direct measure of the amount of thread wear.

From 1985 onward, coinciding with the deregulation of the airline industry, end-play checks at Alaska Airlines became subject to the same kind of drift as the lubrication intervals. In 1985, end-play checks were scheduled every other C check, as the required C checks consistently came in around 2,500 hours, which was rather ahead of the recommended 3,600 flight hours. This would unnecessarily ground the aircraft for maintenance that was not officially due yet.

By scheduling an end-play test every other C check, though, the interval was extended to 5,000 hours. By 1988, C-check intervals themselves were extended to 13 months, with no accompanying flight-hour limit. End-play checks were now performed every 26 months, or about every 6,400 flight hours. In 1996, C-check

intervals were extended once again, this time to 15 months. This stretched the flight hours between end-play tests to about 9,550.

The last end-play check of the accident airplane was conducted at the airline maintenance facility in Oakland, California in 1997. At that time, play between nut and screw was found to be exactly at the allowable limit of 0.040 inches. This introduced considerable uncertainty. With play at the allowable limit, what to do? Release the airplane and replace parts the next time, or replace the parts now? The rules were not clear. The so-called AOL 9–48A said that "jackscrew assemblies could remain in service as long as the end-play measurement remained within the tolerances (between 0.003 and 0.040 inch)."[14] It was still 0.040 inches, so the aircraft could technically remain in service. Or could it? How quickly would the thread wear from there on? Six days, several shift changes and another, more favorable end-play check later, the airplane was released. No parts were replaced: they were not even in stock in Oakland. The airplane "departed 0300 local time. So far so good," the graveyard shift turnover plan noted.[15]

Three years later, the trim system snapped and the aircraft disappeared into the ocean not far away. Between 2,500 and 9,550 hours there had been more drift toward failure. Again, each extension made local sense, and was only an increment away from the previously established norm. No rules were violated, no laws broken. Even the regulator concurred with the changes in end-play check intervals. These were normal people doing normal work around seemingly normal, simple, stable technology. Even the manufacturer had expressed no interest in seeing these numbers or the slow, steady degeneration they may have revealed. If there was drift, in other words, no institutional or organizational memory would know it. But isn't that exactly what we build protective structures for? Regulators, maintenance programs, accountable managers, nominated postholders?

CONTRIBUTION OF THE PROTECTIVE STRUCTURE

The operation of risky technology is surrounded by structures meant to keep it safe. The protection offered by these structures takes a lot of forms. In addition to the organization necessary to actually run the technology, there are regulatory arrangements, quality review boards, nominated post holders, safety departments, and a lot more. These deal with all kinds of information through meetings, committees, international or cross-industry collaborations, data monitoring technology, incident reporting systems, and a host of informal and formal contacts. All of which produces vast arrays of rules, routines, procedures, guidance material, prescriptions, expert opinions, engineering assessments and managerial decisions. And the people who populate this system change jobs and might even move from industry to regulator and back again. This is one of the features that makes complex systems complex: their boundaries are fuzzy. It is almost impossible to say where the system that makes the maintenance rules begins or ends.

There is, in other words, an astounding web of innumerable relationships in which the operation of risky technology is suspended – a web, moreover, that has no clear boundaries, no obvious end or beginning. This creates a wonderful

paradox. The whole structure that is designed (and has evolved) to keep the technology safe, can make the functioning and malfunctioning of that technology more opaque. The meaning of signals about the technology (for example, that it is not behaving according to original manufacturer specifications, or that people are not abiding by the latest procedure) get constructed, negotiated, and transacted through the web of relationships that is strung throughout this structure. The weak signals that are left over trigger only weak organizational responses, if any at all.[16] What if the protective structure itself contributes to the construction and treatment of weak signals, and by extension to drift – in ways that are inadvertent, unforeseen, and hard to detect? The organized social complexity surrounding the technological operation, all the maintenance committees, working groups, regulatory interventions, approvals, and manufacturer inputs, that all intended to protect the system from breakdown, could actually help set its course to, and over, the edge of the envelope.

Take the lengthy, multiple processes by which maintenance guidance was produced for the DC-9 and later the MD-80 series aircraft. At first glance, the creation of maintenance guidance would seem a solved problem. You build a product, you get the regulator to certify it as safe to use, and then you tell the user how to maintain it in order to keep it safe. It was nowhere near a solved problem.

Alaska 261 illustrates the large gap between the production of a system and its operation. Inklings of that gap appeared in observations of jackscrew wear that was higher than what the manufacturer expected. Not long after the certification of the DC-9, people began work to try to bridge the gap. Assembling people from across the industry, a Maintenance Guidance Steering Group (MSG) was set up to develop guidance documentation for maintaining large transport aircraft, particularly the Boeing 747. Using this experience, another MSG developed a new guidance document in 1970, called MSG-2, which was intended to present a means for developing a maintenance program acceptable to the regulator, the operator, and the manufacturer.

The many discussions, negotiations, and inter-organizational collaborations underlying the development of an "acceptable maintenance program" showed that how to maintain a once-certified piece of complex technology was not at all a solved problem. In fact, it was very much an emerging understanding, open to constant revision and negotiation. Technology that appeared simple and certain on the drawing board, proved more unruly. It was not before it hit the field of practice that deficiencies became apparent, if one knew where to look.

In 1980, through combined efforts of the regulator, trade and industry groups and manufacturers of both aircraft and engines in the U.S.A. as well as Europe, a third guidance document was produced, called MSG-3. This document had to deconfound earlier confusions, for example, between "hard-time" maintenance, "on-condition" maintenance, "condition-monitoring" maintenance, and "overhaul" maintenance. Revisions to MSG-3 were issued in 1988 and 1993. The MSG guidance documents and their revisions were accepted by the regulators, and used by so-called Maintenance Review Boards (MRB) that convene to develop guidance for specific aircraft models.

A Maintenance Review Board, or MRB, does not write guidance itself, however; this is done by industry steering committees, often headed by a regulator. These committees in turn direct various working groups. Through all of this, so-called on-aircraft maintenance planning (OAMP) documents get produced, as well as generic task cards that outline specific maintenance jobs.

Both the lubrication interval and the end-play check for MD-80 trim jackscrews were the constantly changing products of these evolving webs of relationships between manufacturers, regulators, trade groups, and operators, who were operating off continuously renewed operational experience, and a perpetually incomplete knowledge base about the still-uncertain technology (remember, end-play test results were not recorded or tracked).

So what are the rules? What should the standards be? The introduction of a new piece of technology is followed by negotiation, by discovery, by the creation of new relationships and rationalities. "Technical systems turn into models for themselves," said Weingart, "the observation of their functioning, and especially their malfunctioning, on a real scale is required as a basis for further technical development."[17] Rules and standards do not exist as unequivocal, aboriginal markers against a tide of incoming operational data (and if they do, they are quickly proven useless or out of date). Rather, rules and standards are the constantly updated products of the processes of conciliation, of give and take, of the detection and rationalization of new data. As Brian Wynne said:

> Beneath a public image of rule-following behavior and the associated belief that accidents are due to deviation from those clear rules, experts are operating with far greater levels of ambiguity, needing to make expert judgments in less than clearly structured situations. The key point is that their judgments are not normally of a kind – how do we design, operate and maintain the system according to 'the' rules? Practices do not follow rules, rather, rules follow evolving practices.[18]

Nor is there a one-way and unproblematic relationship between the original rules or requirements and subsequent operational data. Even if the data, in one reading, may prove the original requirements or rules wrong, this doesn't mean that complex systems, under the various normal pressures of operating economically, reject the requirements and rules and come up with new ones. Instead, the meaning of the data can get renegotiated. People may want to wait for more data. The data can be denied. The data can be said to belong to a category that has nothing to do with safety but everything with normal operational variance (as happened in the *Columbia* Space Shuttle).

A STORY OF DRIFT

Setting up the various teams, working groups, and committees was a way of bridging the gap between building and maintaining a system, between producing it and operating it. Bridging the gap is about adaptation – adaptation to newly

emerging data (for example, surprising wear rates) about unruly technology. But adaptation can mean drift. And drift can mean failure. Alaska 261 is about unruly technology, about gradual adaptations, about drift into failure. It is about the inseparable, mutual influences of mechanical and social worlds, and it highlights the inadequacy of our current models in human factors and system safety. Structuralist models are limited. Of course, we could claim that the lengthy lubrication interval and the unreliable end-play check were structural deficiencies. Were they holes in layers of defense? Sure, if we want to use that language. But such metaphors do not help us look for where the hole occurred, or why.

What kind of model could capture the gradual adaptation of a system like that surrounding jack screws, and predict its eventual collapse? There is something complexly organic about MSGs, something ecological, that is lost when we model them as a layer of defense with a hole in it; when we see them as a mere deficiency or a latent failure. When we see systems instead as internally plastic, as flexible, as organic, their functioning is controlled by dynamic relations and ecological adaptation, rather than by rigid mechanical structures. They also exhibit self-organization (from year to year, the makeup of MSGs was different) in response to environmental changes, and self-transcendence: the ability to reach out beyond currently known boundaries and learn, develop and perhaps improve.

Such an account needs to capture what happens within an organization, with the gathering of knowledge and creation of rationality within workgroups, once a technology gets fielded. A functional account could cover the organic organization of maintenance steering groups and committees, whose makeup, focus, problem definition, and understanding coevolved with emerging anomalies and growing knowledge about an uncertain technology. The journey that ended with the demise of Alaska 261 was marked by smalls steps, by a succession of decrements or downshifts in what people considered normal and acceptable, including those people (both internal and external to the company) tasked with safety oversight. The journey took place in a landscape of increasing competitive pressure and resource scarcity, constantly changing and evolving guidance on how to manage the unruly technology, and with normal people who came to work each day to do a normal job.

Can incident reporting not reveal a drift into failure? This would seem to be a natural role of incident reporting, but it is not so easy. The normalization that accompanies drift into failure (an end-play check every 9,550 hours is "normal," even approved by the regulator, no matter that the original interval was 2,500 hours) severely challenges the ability of insiders to define incidents. What is an incident? Before 1985, failing to perform an end-play check every 2,500 hours could be considered an incident, and given that the organization had a means for reporting it, it may even have been considered as such. But by 1996, the same deviance was normal, regulated even. By 1996, the same failure was no longer an incident.

And there was much more. Why report that lubricating the jackscrew assembly often has to be done at night, in the dark, outside the hanger, standing in the little basket of a lift truck at a soaring height above the ground, even in the rain? Why

report that you, as a maintenance mechanic you have to fumble your way through two tiny access panels that hardly allow room for one human hand – let alone space for eyes to see what is going on inside and what needs to be lubricated – if that is what you have to do all the time? In maintenance, this is normal work, it is the type of activity required to get the job done. The mechanic responsible for the last lubrication of the accident airplane told investigators that he had taken to wearing a battery-operated head lamp during night lubrication tasks, so that he had his hands free and could see at least something. These things, such improvisations, are normal, they are not report-worthy. They do not qualify as incidents. Why report that the end-play checks are performed with one restraining fixture (the only one in the entire airline, fabricated in-house, nowhere near the manufacturer's specifications), if that is what you use every time you do an end-play check? Why report that end-play checks, either on the airplane or on the bench, generate widely varying measures, if that is what they do all the time, and if that is what maintenance work is often about? It is normal, it is not an incident.

Even if the airline had had a reporting culture, even if it had a learning culture, even if it had had a just culture so that people would feel secure in sending in their reports without fear of retribution, these would not be incidents that would turn up in the system. This is the banality of accidents thesis. These are not incidents. Incidents do not precede accidents. Normal work does. In these systems:

> accidents are different in nature from those occurring in safe systems: in this case accidents usually occur in the absence of any serious breakdown or even of any serious error. They result from a combination of factors, none of which can alone cause an accident, or even a serious incident; therefore these combinations remain difficult to detect and to recover using traditional safety analysis logic. For the same reason, reporting becomes less relevant in predicting major disasters.[19]

The failure to adequately see the part to be lubricated (that non-redundant, single-point, safety-critical part), the failure to adequately and reliably perform an end-play check – none of this appears in incident reports. But it is deemed "causal" or "contributory" in the accident report. The etiology of system accidents, then, may well be fundamentally different from that of incidents, hidden instead in the residual risks of doing normal business under normal pressures of scarcity and competition. This means that the so-called common-cause hypothesis (which holds that accidents and incidents have common causes and that incidents are qualitatively identical to accidents except for being just one step short of a true or complete failure) is probably wrong for complex systems.

The signals that are now seen as worthy or reporting, or the organizational decisions that are now seen as "bad decisions" (even though they seemed like perfectly good or reasonable proposals at the time) are seldom big, risky events or order-of-magnitude steps. Rather, there is a succession of weak signals and

decisions, a long and steady progression of small, decremental steps of accepting them that unwittingly take an organization toward disaster. Each step away from the original norm that meets with empirical success (and with no obvious sacrifice of safety) is used as the next basis from which to depart just that *little* bit more again. It is this decrementalism that makes distinguishing the abnormal from the normal so difficult. Knowing what is worth reporting and what is not becomes lost in the fog of uncertainty and lack of differentiation. If the difference between what "should be done" (or what was done successfully yesterday) and what is done successfully today is minute, then this slight departure from an earlier established norm is not worth remarking or reporting on. Decrementalism is about continued normalization: It allows normalization and rationalizes it.

What we need, then, is not yet another structural account of the end-result of organizational deficiency. Not another story of broken components up and down the organizational ladder. What we need is a language that can help us get to a more functional account. A story of living processes, run by people with real jobs and real constraints but no *ill* intentions, whose constructions of meaning co-evolve relative to a set of environmental conditions, and who try to maintain a dynamic and reciprocal relation with their understanding of those conditions. This book intends to pry our logic away from hunting the broken components, and explore a language that might help us better understand the complex system behind the generation of success and failure.

REFERENCES

1 National Transportation Safety Board. (2000). *Factual Report: Aviation (DCA00MA023), Douglas MD-83, N963AS, Port Hueneme, CA, 31 January 2000.* Washington, DC: NTSB.

2 National Transportation Safety Board (2000). Ibid.

3 National Transportation Safety Board (2000). Ibid., p. 23.

4 Fischhoff, B. (1975). Hindsight is not foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3), 288–303.

5 Weick, K. (1995). *Sensemaking in organizations.* London: Sage, p. 28.

6 Poole, R.W. Jr., and Butler V. (1999). Airline deregulation: The unfinished revolution. *Regulation*, 22(1), 8.

7 Poole, R.W. Jr., and Butler V. (1999). Ibid.

8 Rasmussen, J. (1997). Risk management in a dynamic society: A modeling problem. *Safety Science*, 27(2/3), 183–213.

9 National Transportation Safety Board (2000). Ibid., p. 175.

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

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

12 Wynne, B. (1988). Unruly technology: Practical rules, impractical discourses and public understanding. *Social studies of science*, 18, 147–67.
13 Wynne, B. (1988). Ibid.
14 National Transportation Safety Board (2000). Ibid., p. 29
15 National Transportation Safety Board (2000). Ibid., p. 53.
16 Weick, K.E., and Sutcliffe, K.M. (2007). *Managing the unexpected: Resilient performance in an age of uncertainty*. San Francisco: Jossey-Bass.
17 Weingart, P. (1991). Large technical systems, real life experiments, and the legitimation trap of technology assessment: The contribution of science and technology to constituting risk perception. In T.R. LaPorte (ed.), *Social responses to large technial systems: Control or anticipation*, pp. 8–9. Amsterdam: Kluwer, p. 8
18 Wynne, B. (1988). Ibid., p. 153.
19 Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems. *Safety Science*, 37, 109–26, p. 112.

3
THE LEGACY OF NEWTON AND DESCARTES

❖

Not long ago, a nurse from Wisconsin was charged with criminal "neglect of a patient causing great bodily harm" in the medication error-related death of a 16-year-old woman during labor. The nurse accidentally administered a bag of epidural analgesia instead of the intended penicillin by the intravenous route. The criminal complaint concluded that:[1]

- The child's attending physician and the defendant's nurse supervisor reported that the nurse failed to obtain authorization to remove the lethal chemicals that caused the child's death from a locked storage system.
- The nurse disregarded hospital protocol by failing to scan the bar code on the medication, a process of which the nurse had been fully trained and was cognizant. Had the lethal chemicals been scanned, medical professionals would have been forewarned of its lethality and the death would have been prevented.
- The nurse disregarded a bright, clearly written warning on the bag containing the lethal chemicals prior to injecting them directly into the child's bloodstream.
- The nurse injected the lethal chemicals into the bloodstream in a rapid fashion, failing to follow the approved rate for any medications that may have been prescribed for the child, in an apparent effort to save time. The rapid introduction of these chemicals dramatically hastened the death of the child, effectively thwarting any ability to save her life.
- The nurse disregarded hospital protocol and failed to follow professional nursing procedures by not considering the five rights of patients prior to the administration of the lethal chemicals (right patient, right route, right dose, right time and right medication). The practice at her hospital requires the consideration of these five factors at least three times prior to the administration of any medication; the most important procedure established to prevent putting a patient's life in jeopardy through medication errors.

4

THE SEARCH FOR THE BROKEN COMPONENT

❖

If there is one idea that Newton and Descartes have taught us when things go wrong, it is to look for the broken component. Nothing characterizes the Newtonian–Cartesian worldview as much as its emphasis on analytic reduction, or decomposition. The story in Chapter 2, of course, is about a broken part – the jackscrew/acme nut on an MD-80 airliner. Though perhaps it is more correct to say that the story starts with a broken part. In itself the part isn't that interesting. If we want to understand *why* it ended up broken, analytic reduction doesn't get us very far. Instead, we need to go up and out, rather than down and in. We have to begin to probe the hugely intertwined webs of relationships that spring out and away from the broken part, into the organizational, the institutional, the social. Yet often our quest to understand why parts are broken simply leads us to other broken parts. The decompositional logic is almost everywhere.

Consider the practice, very common in all kinds of accident investigations, to examine the 24-hour and 72-hour histories of the persons involved at the sharp end – the train driver, the pilots, the first mate. Hardly ever do these examinations reveal anything interesting. Sure, the pilot may have been playing tennis with her husband the day before, got 7 hours of sleep before getting up to report to the airport. The train driver had a 10-hour day prior to the day of the accident. So? Well, this is to rule out anything to do with the operators, investigators have told me. Fair enough.

But this seemingly very innocent practice of excluding possible causes serves to reconfirm the image of an investigation as a search for the broken component in the rubble. It legitimates as well as reproduces a kind of naïve Newtonian physics about how systems work and fail. How do we find broken components? The 24- and 72-hour histories suggest that we first start with the people at the sharp end. We rummage through their lives, we probe more and more hours backwards into their professional and private doings, hoping for the "Aha!" part to pop up. A pill bottle on the night stand. A fight with the spouse. A lousy hotel room on an overnight with too much noise and too little sleep. That last beer leaking over into the 8-hour window from bottle to throttle.

If this search doesn't turn up anything; if we fail to find the broken component there, we redirect our attention to the next system, or the next level. We may end up by looking at the Safety Management System of the organization, a fashionable thing to do nowadays. And we may find that it is broken in various places. Reporting doesn't work as well, for example, or the quality management of a training program is deficient in how it tracks people's progress along various levels of documentation and bureaucratic accountability. The resulting investigation, a catalogue of broken or bent components, is all too common. The "findings" section of accident reports are often precisely such an engineering catalogue, listing the components that were found broken, and the components that were found not broken.

BROKEN COMPONENTS AFTER A HAILSTORM

After its investigation of an incident in which an airliner flew into a hailstorm and sustained severe damage, the report came up with the following in its "findings" section (among others):

1. The pilots' licenses and ratings were valid.
OK, component not broken.

2. The aircraft had a valid certificate of airworthiness.
OK, component not broken.

3. The pilots reported for duty well before the required time.
OK, component not broken.

4. The pilots did not study or take with them the Nordic Severe Weather Chart for the time of their flight. Moreover, they did not study the Severe Weather Chart prepared by the World Area Forecast Centre.
Not OK – two broken components found.

5. The Aeronautical Weather Services was preparing a SIGMET (a forecast for significant weather) but it was published too late with regard to the development of weather conditions.
Not OK – one broken component found.

6. Engine anti-ice and ignition systems were used as instructed.
OK, component not broken.
7. The galleys in the cabin had been secured for landing.
OK, component not broken.

8. The pilot flying did not disengage the autothrottle when the aircraft flew into turbulence, as recommended by the Aircraft Operating Manual.
Not OK, one broken component found.

9. The co-pilot tried to switch the weather radar on, but no radar image was shown.
Not OK – one broken component found.

10. After the windshields had been cracked, the pilot flying turned the windshield heating switch off. This action is not included in the malfunction check list flight deck window cracks in flight. The flight crew did not take the actions mentioned in the checklist.
Not OK – one broken component found.

11. The cabin crew paid attention to the abnormal, rattling sound of the engines. However, they did not inform the cockpit crew of their observation during the flight.
Not OK – one broken component found.

And so it goes on. A catalogue of components, both broken and not broken.

An accident report that lists findings as its major product invokes comparison with the way in which we used to train doctors (and still do, in many places). Give the students a bunch of components (both broken and not broken), divvy them up across classes ranging from anatomy to physiology to neurology, and then let the students themselves piece it all together; let them figure out how all these components work together to account for the functioning or malfunctioning of a human body. In case of a typical accident or incident report, the writers make a similar leap of faith. The reader is trusted to make a meaningful connection between a list of broken and not-broken components and a coherent narrative of the event. This connection only works in a Newtonian–Cartesian world, where there is a straightforward relationship between the way in which components work or fail, and how the system works or fails as a whole.

The perpetual search for broken components is in part a heritage of the engineering origins of accident investigations. An engineering understanding of the world, at least if it is driven by Newtonian–Cartesian logic, is to break the large thing (the accident or incident) apart so as to reveal its inner components. Ferret about among those components, and find the ones that are broken.

Note the really interesting (and, as you will see later, really problematic) assumption that there is a direct relationship between the broken component(s) inside the system, and the behavior of the system as a whole. In the example above, the supposed mishandling of the encounter with hail by the flight crew, and the resulting damage and danger to the aircraft and its occupants, can be reduced to the pilots not conducting some of the procedures by the book, and making other procedures up where the book offered none. The whole, in other words, can be explained by reference to the behavior of the parts. The functioning or non-functioning of the whole can be understood through the functioning or non-functioning of constituent components. In the Newtonian–Cartesian world,

their relationship is unproblematic, direct, and in no need of further proof or synthetic work. This, in effect, is where many investigations close the book. Broken component(s) identified, fixes suggested. Finished.

On the other hand, we could also consider our system as something that is alive, as something that has all kinds of processes running through it at a variety of levels, which connect the various components in many complex ways. In that case, pulling the system apart and lifting components out of it to examine them for their individual performance basically kills the phenomenon of interest. Broken components may not really explain a broken system. But we are getting ahead of ourselves. For now, let's try to understand how the hunt for broken components is something that we equate with good analysis, with finding the cause, with knowing what went wrong, with knowing what to fix. So what if we can't find a broken component?

In the summer of 1996, Trans World Airlines flight 800 from New York to Paris, a Boeing 747–100 with 230 people onboard, exploded suddenly when climbing out over the sea alongside Long Island. It was only 11 minutes after takeoff. The huge amount of kerosene for its Atlantic crossing vaporized and ignited, creating a fireball that was seen along the coastline of Long Island. The fragmented aircraft disappeared into the ocean, forming an expanding bubble of debris so pulverized that a weather radar interpreted it as a rain cloud.

What followed was one of the most expensive and most vexing accident investigations to date. Yet, despite its length and the reconstruction of the entire wreckage in a hangar close by, it failed to yield the "eureka part," that single bit of wreckage, or equipment, that contained a critical clue, a clue that could point to the broken component. It gave rise to a large number of conspiracy theories. A light streaking towards the jet was seen by some witnesses just prior to the explosion, metal in one place seemed bent inwards and then outwards, pointing to a hole created by an external projectile, and microscopic traces of PETN (Pentaerythritol Tetranitrate), a compound used in plastic explosives, were found on a piece of flooring from the passenger compartment. A conspiracy involving the U.S. Navy, a test missile, the FBI and a whole host of other agencies, would be one reasonable candidate cause. Later, however, the FBI stated that these traces were consistent with explosives allegedly spilled during a training exercise aboard the aircraft a few weeks before the crash.

After more than four years, the National Transportation Safety Board (NTSB) concluded that the probable cause of the accident was an explosion of the center wing fuel tank, resulting from ignition of the flammable fuel/air mixture in the tank. The source of ignition energy for the explosion could not be determined with certainty, but, of the sources evaluated by the investigation, the most likely was a short-circuit outside of the tank that allowed excessive voltage to enter it through electrical wiring associated with the fuel quantity indication system. Researchers from across the U.S.A., dissatisfied with the many anomalies and unexplained pieces of evidence in the official investigation, have gathered in a

group called the "Flight 800 Independent Researchers Organization," (or FIRO) dedicated to assisting official investigators determine "the cause" of this yet unresolved disaster.

Interestingly, the chairman of the NTSB at the time, Jim Hall, raised the specter of his agency not being able to find the eureka part, which would challenge its entire reputation: "What you're dealing with here is much more than an aviation accident because of the profile of the crash. What you have at stake here is the credibility of this agency and the credibility of the government to run an investigation," he said.

In other words, the credibility of the U.S. government to run an investigation could hinge on its ability to return with a eureka part in its hands. A failure to produce a eureka part, conversely, would mean that the government's approach, methods, techniques and expertise were insufficient or inadequate. The eureka part was out there, but it would take, as FIRO aimed to do, a different kind of investigation, less political, less cowed by other agencies, and following up on other clues, ironing out inconsistencies and ambiguities that had been left in place.

In a sense, the credibility of the government to run an investigation hinged on its ability to find the cause for the effect, to find the "eureka part," to locate the part that would have everybody in the investigation declare that they'd found the broken component, the trigger, the original culprit, that could carry the explanatory load of the loss of the Boeing 747. But for this crash, the so-called "eureka part" was never found. The FIRO (Flight 800 Independent Researchers Organization) position represented the culmination of dissatisfaction with this inability. Theirs represented a very Newtonian commitment, just like the NTSB's. For FIRO, the eureka part was out there; all it would take was a better method, or a different method. The idea that accidents could be produced without any part failing, or without anything external interfering, but rather as a result of normally functioning components interacting in unforeseen ways,[1] was not up for debate here. It would have fallen outside the ruling paradigm: an effect must have a cause, and that cause should be a broken component, somewhere, somehow, that can be found. Even if it took belief in a conspiracy to come up with one.

BROKEN COMPONENTS TO EXPLAIN A BROKEN SYSTEM

Componential explanations that condense accounts of failure down to the failure of a part still reign supreme. And the human is often that part. A recent analysis by Richard Holden showed that between 1999 and 2006, 96 percent of investigated U.S. aviation accidents were attributed in large part to the flight crew. In 81 percent, people were the *sole* reported cause. The language used in these analyses has judgmental or even moralistic overtones too. "Crew failure" or a similar term appears in 74 percent of probable causes and the remaining

cases contain language such as "inadequate planning, judgment and airmanship," "inexperience" and "unnecessary and excessive … inputs." "Violation" of written guidance was implicated as cause or contributing factor in a third of all cases.[2] Single-factor, judgmental explanations for complex system failures are not unique to aviation – they are prevalent in fields ranging from medicine, shipping, military operations, to road traffic.[3]

The problem of safety analysis reverting to condensed and individual/componential explanations rather than diffuse and system-level ones was one driver behind the fields of human factors and system safety. A charter of these fields has been to take safety scientists and practitioners behind the label "human error," to more complex stories of how normal system factors contribute to things going wrong (including factors associated with new technology and with the organization and its multiple interacting goals). But common discourse about failure in complex systems remains tethered to language such as "chain-of-events," "human error" and questions such as "what was *the* cause?" and "*who* was to blame?"

NEWTON AND THE SIMPLICITY OF FAILURE

The search for broken or underperforming components that can carry the explanatory and moral load of accidents in complex systems has deep roots in what Western society finds rational, logical and just. As shown in the previous chapter, throughout the past three and a half centuries, the West has equated scientific thinking with a Newtonian–Cartesian worldview which prizes decomposition, linearity, and the pursuit of complete knowledge. Let's look at how this has given rise to what I call a Newtonian ethic of failure – which makes particular assumptions about the relationship between cause and effect, foreseeability of harm, time-reversibility and the ability to come up with the "true story" (that is, "the" cause) of a particular event. These assumptions animate much of our safety work today, including accident investigations, societal expectations, managerial mandates, technical tools and artifacts, and judicial responses to system failures.

The logic behind Newtonian science is easy to formulate, although its implications for how we think about the ethics of failure are subtle, as well as pervasive. In this section we review aspects of the Newtonian–Cartesian worldview that influence how we understand (and consider the ethics around) failure in complex systems, even today. Until the early twentieth century, classical mechanics, as formulated by Newton and further developed by Laplace and others, was seen as the foundation for science as a whole. It was expected that observations made by other sciences would sooner or later be reduced to the laws of mechanics. Although that never happened, other disciplines, including psychology and law, did adopt a reductionist, mechanistic or Newtonian methodology and worldview. This influence was so great that most people still implicitly equate "scientific thinking" with "Newtonian thinking." The mechanistic paradigm is compelling in its simplicity, coherence and apparent completeness. It was not only successful

in scientific applications, but also largely consistent with intuition and common sense.

REDUCTIONISM AND THE EUREKA PART

As said, the best known principle of Newtonian science, formulated well before Newton by the philosopher-scientist Descartes, is that of analysis or reductionism. To understand any complex phenomenon, you need to take it apart, that is, reduce it to its individual components. This is recursive: if constituent components are still complex, you need to take your analysis a step further, and look at their components. In other words, the functioning or non-functioning of the whole can be explained by the functioning or non-functioning of constituent components. Attempts to understand the failure of a complex system in terms of failures or breakages of individual components in it – whether those components are human or machine – are very common.

Linear, componential thinking permeates the investigation of accidents and organizational failures as well as reliability engineering methods. The defenses-in-depth metaphor, popularized as the "Swiss Cheese Model" and used in event classification schemes relies on the componential, linear parsing-up of a system, so as to locate the layer or part that was broken. The analytic recursion in these methods ends up in categories such as "unsafe supervision" or "poor managerial decision-making." Indeed, in technically increasingly reliable systems, the "eureka part" has become more and more the human and there is a cottage industry of methods that try to locate and classify which errors by which people lay behind a particular problem. Most or all implicitly presume a linear relationship between the supposed error and the parts or processes that were broken or otherwise affected.

Our understanding of the psychological sources of failure is subject to reductive Newtonian recursion as well. In cases where the component failure of "human error" remains incomprehensible, we take "human error" apart too. Methods that subdivide human error up into further component categories, such as perceptual failure, attention failure, memory failure or inaction are now in use in air traffic control and similar linear, reductionist understandings of human error dominate the field of human factors. The classically mechanistic idea of psychology that forms the theoretical bedrock for such reductionist thinking of course predates human factors. Analytic reduction sponsors an atomistic view of complex psychological phenomena: understanding them comes from revealing the functioning or breakdown of their constituent components.

Even sociological or cultural phenomena are often explained using a reductionist approach. Much safety culture work, for example, breaks "culture" down into what individual employees experience or believe (mostly through attitude surveys about, for example, work pressure, management behaviors in relation to safety, risk perceptions), as well as the presence or absence of particular components of, for example, a safety management system, safety investments, worker qualifications or other safety arrangements. It measures those, adds them up and

gets a number for a safety culture. Together, these components are assumed to constitute a "culture," yet it is never made clear how the parts become the whole. Such approaches meet with skepticism from those who see culture as something much more complex and incompressible:

> Culture cannot be managed; it emerges. Leaders don't create cultures; members of the culture do. Culture is an expression of people's deepest needs, a means of endowing their experiences with meaning. Even if culture in this sense could be managed, it shouldn't be (...) it is naive and perhaps unethical to speak of managing culture.[4]

This is not to say that measuring these things is not meaningful. It may well identify interesting areas for safety intervention. But to say that it measures "culture" is making a particular set of Newtonian assumptions about the relationship between part and whole, between individual and society. This relationship is called into question by anthropologists like Hutchins, who argue that the regularities that are seen as characteristic for a culture (the whole) cannot be easily found inside the members (the parts):

> A given group of individuals may enact different distributed cognitive processes depending on institutional arrangements. Observed patterns of behavior emerge from the interactions of the internal processes with structures and processes that are present in the environment for action. This means that the regularities that are often identified as being characteristic of a culture may not be entirely "inside" the individual members of the society in question and may not generalize across activity settings. From the contextual point of view, the term culture can be read as a shorthand label for an emergent uneven distribution of a variety of material, social, and behavioral patterns that result from a universal human process.[5]

The atomistic view of culture or community or society sees the relationship between parts and whole as unproblematic. The parts simply add up to make the whole. If parts are imperfect or even missing (for example, if particular components of a safety management system are incomplete, or access to the boss with safety concerns is impossible), then this will add up to a culture in a straightforward way, to a measurable safety culture (or lack thereof). This logic (of a simple, additive relationship between parts and the whole) is so pervasive that Margaret Thatcher even denied the existence of community or society altogether in an interview in 1987: " ... there is no such thing as society. There are individual men and women, and there are families." From that extreme point of view, there are only atoms, parts, components. And somehow, they make up a whole. But for that point of view, the whole isn't the point. It's the parts and the easily added numbers they represent.

It is interesting to note that such a position is quite consistent with Western, and often Protestant societies. Through the Enlightenment, and the intellectual tradition since the Scientific Revolution, it has seemed self-evident to evaluate

ourselves as individuals, bordered by the limits of our minds and bodies, and that we get assessed in terms of our own personal achievements. From the renaissance onwards, the individual became a central focus, fueled in part by Descartes' psychology that created self-contained individuals. The rugged individualism developed on the back of mass European immigration into North America in the late nineteenth and early 20th centuries accelerated the image of independent, free heroes accomplishing greatness against all odds, and anti-heroes responsible for disproportionate evildoing. The notion that it takes teamwork, or an entire organization, or an entire industry (think of Alaska 261) to break a system is just too eccentric relative to this cultural prejudice.

The philosophy of Newtonian science is one of simplicity: the complexity of the world is only apparent and to deal with it we need to analyse phenomena into their basic components. The way in which legal reasoning in the wake of accidents separates out one or a few actions (or inactions) on the part of individual people follows such reductive logic. For example, the Swedish Supreme Court ruled that if one nurse had more carefully double-checked a particular medication order before preparing it (mistakenly at 10 times the intended dose) a three-month old baby would not have died. Such condensed, highly focused accounts that converge on one (in)action by one person (the "eureka part") give componential models of failure a societal legitimacy that keeps reproducing and instantiating Newtonian physics.

CAUSES FOR EFFECTS CAN BE FOUND

In the Newtonian vision of the world, all that happens has a definitive cause and a definitive effect. In fact, there is a symmetry between cause and effect (they are equal but opposite). The determination of the "cause" or "causes" is, of course, seen as the most important function of accident investigation today, and assumes that physical effects (a crashed airliner, a dead patient) can be traced back to physical causes (or a chain of causes-effects). The notion that effects cannot occur without causes makes it into legal reasoning in the wake of accidents too. For example, "to raise a question of negligence, harm must be caused by the negligent action."[6] It is assumed that a causal relationship (that negligent action caused harm) is indeed demonstrable, provable beyond reasonable doubt.

The Newtonian view of the world that holds all this up is materialistic: all phenomena, whether physical, psychological or social, can be reduced to matter, that is, to the movement of physical components inside three-dimensional Euclidean space. The only property that distinguishes particles is where they are in that space. Change, evolution, and indeed accidents, can be reduced to the geometrical arrangement (or misalignment) of fundamentally equivalent pieces of matter, whose interactive movements are governed exhaustively by linear laws of motion, of cause and effect. A visible effect (for example, a baby dead of lidocaine poisoning) cannot occur without a cause (a nurse blending too much of the drug). The Newtonian assumption of proportionality between cause and effect can in fact make us believe that really bad effects (the dead baby) have really

bad causes (a hugely negligent action by an incompetent nurse). The worse the outcome, the more "negligent" its preceding actions are thought to have been. In road traffic, talking on a cell phone is not considered illegal by many, until it leads to a (fatal) accident. It is the effect that makes the cause bad.

The Newtonian model has been so pervasive and coincident with "scientific" thinking, that if analytic reduction to determinate cause-effect relationships (and their material basis) cannot be achieved, then either the method or the phenomenon isn't considered worthy of the label "science." This problem of scientific self-confidence has plagued the social sciences since their inception, inspiring not only Durkheim to view the social order in terms of an essentialist naïve Newtonian physics, but also for example to have Freud aim "to furnish a psychology that shall be a natural science: that is, to represent psychical processes as quantitatively determinate states of specifiable material particles, thus making those processes perspicuous and free from contradiction."[7] Behaviorists like Watson reduced psychological functioning to mechanistic cause-effect relationships in a similar attempt to protect social science from accusations of being unscientific.[8]

THE FORESEEABILITY OF HARM

According to Newton's image of the universe, the future of any part of it can be predicted with absolute certainty if its state at any time was known in all details. With enough knowledge of the initial conditions of the particles and the laws that govern their motion, all subsequent events can be foreseen. In other words, if somebody can be shown to have known (or should have known) the initial positions and velocities of the components constituting a system, as well as the forces acting on those components (which in turn are determined by the positions of these and other particles), then this person could, in principle, have predicted the further evolution of the system with complete certainty and accuracy. A system that combines the physiology of a three-month old baby with the chemical particles diethylamino-dimethylphenylacetamide that constitute lidocaine will follow such lawful evolution, where a therapeutic dose is less than 6 mg lidocaine per gram serum, and a dose almost 10 times that much will kill the baby.

If such knowledge is in principle attainable, then the harm that may occur if particles are lined up wrongly is foreseeable, too. Where people have a duty of care (like nurses and other healthcare workers do) to apply such knowledge in the prediction of the effects of their interventions, it is consistent with the Newtonian model to ask how they failed to foresee the effects. Did they not know the laws governing their part of the universe (that is, were they incompetent, unknowledgeable)? Were they not conscientious or assiduous in plotting out the possible effects of their actions? Indeed, legal rationality in the determination of negligence follows this feature of the Newtonian model almost to the letter: "Where there is a duty to exercise care, reasonable care must be taken to avoid acts or omissions which can reasonably be foreseen to be likely to cause harm. If, as a result of a failure to act in this reasonably skillful way, harm is caused, the person whose action caused the harm, is negligent."[9]

In other words, people can be construed as negligent if the person did not avoid actions that could be foreseen to lead to effects – effects that would have been predictable and thereby avoidable if the person had sunk more effort into understanding the starting conditions and the laws governing the subsequent motions of the elements in that Newtonian sub-universe. Most road traffic legislation is founded on this Newtonian commitment to foreseeability too. For example, a road traffic law in a typical Western country might read how a motorist should adjust speed so as to be able stop the vehicle before any hinder that might be foreseeable, and remain aware of the circumstances that could influence such selection of speed.[10] Both the foreseeability of all possible hinders and the awareness of circumstances (initial conditions) as critical for determining speed are steeped in Newtonian epistemology. Both are also heavily subject to outcome bias: if an accident suggests that a hinder or particular circumstance was not foreseen, then speed was surely too high. The system's user, as a consequence, is always wrong. And any search for the cause of failure will therefore always turn up a broken part.

TIME-REVERSIBILITY

The trajectory of a Newtonian system is not only determined towards the future, but also towards the past. Given its present state, we can in principle reverse the evolution to reconstruct any earlier state that it has gone through. The Newtonian universe, in other words, is time-reversible. Because the movement of, and the resulting interactions between, its constituent components are governed by deterministic laws of cause and effect, it does not matter what direction in time such movements and interactions are plotted. Such assumptions, for example, give accident and forensic investigators the confidence that an event sequence can be reconstructed by starting with the outcome and then tracing its causal chain back into time. The notion of *re*construction reaffirms and instantiates Newtonian physics: our knowledge about past events is nothing original or creative or new, but merely the result of uncovering a pre-existing order. The only thing between us and a good reconstruction are the limits on the accuracy of our representation of what happened. We then assume that this accuracy can be improved by "better" methods of investigation, for example.

COMPLETENESS OF KNOWLEDGE

The traditional Western belief in science is that its facts have an independent existence outside of people's minds: they are naturally occurring phenomena "out there," in the world. The more facts a scientist or analyst or investigator collects, the more it leads, inevitably, to more, or better, science: a better representation of "what happened." The belief is that people create representations or models of the "real" out there, models or representations that mimic or map this reality. Knowledge is basically that representation. When these copies, or facsimiles, do not match "reality," it is due to limitations of perception, rationality, or cognitive

resources, or, particularly for investigators or researchers, due to limitations to methods of observation. More, or more refined methods and more data collection, can compensate for such limitations.

Newton argued that the laws of the world are discoverable and ultimately completely knowable. God created the natural order (though kept the rulebook hidden from man; instead God gave man intelligence to go figure it out for himself) and it was the task of science to discover this hidden order underneath the apparent disorder. The Newtonian view is based on the reflection-correspondence view of knowledge: our knowledge is an (imperfect) mirror of the particular arrangements of matter outside of us.[11] The task of investigations, or science, is to make the mapping (or correspondence) between the external, material objects and the internal, cognitive representation (for example, language, or some mental model) as accurate as possible. The starting point is observation, where information about external phenomena is collected and registered (for example, the gathering of "facts" in an accident investigation), and then gradually completing the internal picture that is taking shape. In the limit, this can lead to a perfect, objective representation of the world outside.[12] The world is already there, pre-formed and pre-existing. All we need to do is uncover and then describe the order that the world already possesses, that it was given by its creator.

The Newtonian position can be recognized in what society generally sees as high science today. The leader of the International Human Genome Sequencing Consortium and currently director of the U.S. National Institutes of Health recently wrote a book entitled *The Language of God*.[13] For him, the sequencing of the human genome (our genetic information, stored on 23 chromosomes and occupied by more than three billion DNA base pairs) was ultimately a discovery of God's language: God's code in which the essence of humanity had been written. The language, the order, had been there for thousands of millennia already. But at the closing of the second millennium CE, science had finally arrived at a method to lay it bare. It is a position that inspires awe and respect: people can marvel in the amazing creation.

But the human genome project revealed that the supposedly pre-existing order of the building blocks of life (DNA) severely underdetermines how the complex system looks, or works, or fails. Only about 1 percent of the 3 billion base pairs actually code for human features that we know about. Humans have fewer than double the number genes of very basic organisms such as fruit flies and roundworms, and have more than 99 percent genes in common with chimpanzees. Genetically, we are even closer to Neanderthals. This is typical for complex, as opposed to Newtonian, systems. Complexity implies an ultimately intractable relationship between the parts and the whole.

What is more, the language that the Human Genome Sequencing Consortium uncovered was actually hardly orderly, and hardly entirely pre-existing. Human cells make extensive use of alternative splicing by which a single gene can code for multiple proteins, which already explodes any straightforward relationship between parts and whole. There also seem to be nonrandom patterns of gene density along chromosomes which are not well understood, and regions of coding and non-coding DNA. Then there is gene-switching: the ability to turn genes on or off through the organism's interaction with its changing environment.

The human genome, in other words, can at best be described as a set of hardware, which in turn can run all kinds of versions of software and thousands of different and partially overlapping and even contradictory programs at the same time. Complex systems are open systems, and so is the human genome. It is this openness to the environment and the system's ability to recognize, adapt, change, and respond that renders any project to describe the pre-existing order quite hopeless. When you think you've described the system once, you'll find that it will have morphed away from your description before you're even done. And even any temporary or tentative description of the arrangement of parts spectacularly underdetermines what can be observed at the system level.

Founding sociologist Emile Durkheim took the same Newtonian position for social science in the nineteenth century. Underneath a seemingly disordered, chaotic appearance of the social world, he argued, there is a social order governed by discoverable laws (obligations and constraints) and categories of human organization (institutions). As with the theological reading of the human genome project a century-and-a-half later, Durkheim wanted to pursue the essential properties of social systems: those features that are enduring, unchanging, and that can be discovered and described independent on who does the discovering or describing, from what angle or perspective, and at what point in time. Such "essentialism" is typical for Newtonian science: the idea that behind our initial befuddlement and confusion of the world as it meets us, lies an unchanging, pre-existing order of hard facts that we can lay bare in due time.

The consequence of this Newtonian position for understanding failure is that there can be only one true story of what happened. For Newton and Descartes, the "true" story is the one in which there is no more gap between external events and their internal representation. Those equipped with better methods, and particularly those who enjoy greater "objectivity," (that is, those who, without any bias that distorts their perception of the world, will consider *all* the facts) are better poised to achieve such a true story. Formal, government-sponsored accident investigations enjoy this aura of objectivity and truth – if not in the substance of the story they produce, then at least in the institutional arrangements surrounding its production. First, their supposed objectivity is deliberately engineered into the investigation as a sum of subjectivities. All interested parties (for example, vendors, the industry, the operator, the legal system, unions, professional associations) can officially contribute (though some voices are easily silenced or sidelined or ignored). Second, those other parties often wait until a formal report is produced before publicly taking either position or action, legitimating the accident investigation as arbitrator between fact and fiction, between truth and lie. It supplies the story of "what really happened." Without first getting that "true" story, no other party can credibly move forward.

A NEWTONIAN ETHIC OF FAILURE

Together, taken-for-granted assumptions about decomposition, cause-effect symmetry, foreseeability of harm, time reversibility, and completeness of knowledge give rise to a Newtonian ethic in the wake of failure. It can be summed up as follows:

- To understand a failure of a system, we need to search for the failure or malfunctioning of one or more of its components. The relationship between component behavior and system behavior is analytically non-problematic.
- Causes for effects can always be found, because there are no effects without causes. In fact, the larger the effect, the larger (for example, the more egregious) the cause must have been.
- If they put in more effort, people can better foresee outcomes. After all, they would better understand the starting conditions and are already supposed to know the laws by which the system behaves (otherwise they wouldn't be allowed to work in it). With those two in hand, all future system states can be predicted. If they can be predicted, then harmful states can be foreseen and should be avoided.
- An event sequence can be reconstructed by starting with the outcome and tracing its causal chain back into time. Knowledge thus produced about past events is the result of uncovering a pre-existing order.
- There can be only one true story of what happened. Not just because there is only one pre-existing order to be discovered, but also because knowledge (or the story) is the mental representation or mirror of that order. The *truest* story is the one in which the gap between external events and internal representation is the smallest. The *true* story is the one in which there is no gap.

As in many other fields of inquiry, these assumptions remain largely transparent and closed to inquiry in safety work precisely because they are so self-evident and common-sensical. The way they get retained and reproduced is perhaps akin to what the sociologist Althusser called "interpellation."[14] People involved in safety work are expected to explain themselves in terms of the dominant assumptions; they will make sense of events using those assumptions; they will then reproduce the existing order in their words and actions. Organizational, institutional and technological arrangements surrounding their work don't leave plausible alternatives (in fact, they implicitly silence them). For instance, investigators are mandated to find the probable cause(s) and turn out enumerations of broken components as their findings. Technological-analytical support (incident databases, error analysis tools) emphasizes linear arrangements and the identification of malfunctioning components. Also, organizations and those held accountable for failure inside them, need something to "fix," which further valorizes condensed accounts, focuses on localization of a few single problems and re-affirms a pre-occupation

with components. If these processes fail to satisfy societal accountability requirements, then courts deem certain practitioners criminal, lifting uniquely bad components out of the system.

Newtonian hegemony, then, is maintained not by imposition but by interpellation, by the confluences of shared relationships, shared discourses, institutions, and knowledge. Foucault called the practices that produce knowledge and keep knowledge in circulation an epistemé: a set of rules and conceptual tools for what counts as factual. Such practices are exclusionary. They function in part to establish distinctions between those statements that will be considered true and those that will be considered false. Or factual rather than speculative. Or just rather than unjust.[15] The true statement will be circulated through society, reproduced in accident reports, for example, and in books and lectures about accidents. These true statements will underpin what is taken to be common-sensical knowledge in a society – that is, the Newtonian physical order. The false statement will quickly fade from view as it not only contradicts common sense, but also because it is not authorized by people legitimated and trusted by society to furnish the truth.

A naïve socio-technical Newtonian physics is thus continuously read into events that could yield much more complexly patterned interpretations. Newtonian assumptions not only support but also reproduce their own ideas, values, sentiments, images, and symbols about failure, its origin and its appropriate ethical consequences. Collectively, they assert that action in the world can be described as a set of casual laws, with time reversibility, symmetry between cause and effect, and a preservation of the total amount of energy in the system. The only limiting points of such an analysis are met when laws are not sufficiently rigorous or exhaustive, but this merely represents a problem of further methodological refinement in the pursuit of greater epistemological rigor and more empirical data produced because of it.

On April 27, 1900, Lord Kelvin, the eminent physicist and president of Great Britain's Royal Society, confidently told his members that Newton's approach had been successfully extended to embrace all of physics, including both heat and light. In fact, the turn of the century was accompanied by an optimism about Newton's theories of particles and motion, which had been confirmed by generations of scientists in fields ranging from chemistry to biology to sociology.[16] Newton appeared to be able to explain every phenomenon in the universe. Everything was, in principle, explicable, knowable. The great order of the creation would be uncloaked. The outbreak of colossal epidemics had been tamed, the industrial revolution was moving into overdrive. People's infatuation with science and technology gave them supreme confidence in their ability to manipulate and control the world around them.

Today, we have inherited a Newtonian commitment that all but excludes even a common awareness of alternatives. It is not that more complex readings would be "truer" in the sense of corresponding more closely to some objective state of affairs. But they could hold greater or at least a different potential for safety improvement, and could help people reconsider what is useful and ethical in the

aftermath of failure. The next chapter studies the scientific literature on safety and accidents for how it has attempted to theorize drift. As we will see, Newtonian assumptions keep cropping up, even in narratives that try to take history and complexity seriously.

REFERENCES

1 Leveson, N., Cutcher-Gershenfeld, J., Barrett, B., Brown, A., Carroll, J., Dulac, N., Fraile, L., and Marais, K. (2004). Effectively Addressing NASA's Organizational and Safety Culture: Insights from Systems Safety and Engineering Systems. *Paper presented at the Engineering Systems Division Symposium, MIT*, Cambridge, MA March 29–31.

2 Holden, R. (2009). People or Systems? To blame is human. The fix is to engineer. *Professional Safety*, 12, 34–41.

3 Catino, M. (2008). A review of literature: Individual blame vs. organizational function logics in accident analysis. *Journal of Contingencies and Crisis Management*, 16(1), 53–62.

4 Martin, J. (1985). Can organizational culture be managed? In Frost, P.J. (ed.), *Organizational Culture*. Thousand Oaks, CA: Sage Publications, p. 95.

5 Hutchins, E., Holder, B., and Pérez, R.A. (2002). *Culture and flight deck operations (Research Agreement 22–5003)*. San Diego: University of California, UCSD, p. 3.

6 GAIN (2004). *Roadmap to a just culture: Enhancing the safety environment*. Global Aviation Information Network (Group E: Flight Ops/ATC Ops Safety Information Sharing Working Group).

7 Freud, S. (1950). Project for a scientific psychology. In *The standard edition of the complete psychological works of Sigmund Freud, vol. I*. London: Hogarth Press, p. 295.

8 Dekker, S.W.A. (2005). *Ten questions about human error: A new view of human factors and system safety*. Mahwah, NJ: Lawrence Erlbaum Associates.

9 GAIN (2004). Ibid., p. 6.

10 This comes from the Swedish *Trafikförordning* 1996:1276, kap.3, §14 and 15. See also Tingvall, C. and Lie, A. (2010). The concept of responsibility in road traffic (Ansvarsbegreppet i vägtrafiken). Paper presented at *Transportforum*, Linköping, Sweden, 13–14 January.

11 Heylighen, F. (1989). Causality as distinction conservation: A theory of predictability, reversibility and time order. *Cybernetics and Systems*, 20, 361–84.

12 Heylighen, F., Cilliers, P., and Gershenson, C. (2005). *Complexity and philosophy*. Vrije Universiteit Brussel: Evolution, Complexity and Cognition.

13 Collins, F.S. (2006). *The language of God*. New York: Free Press.

14 Althusser, L. (1984). *Essays on ideology*. London: Verso.

15 Foucault, M. (1980). Truth and power. In C. Gordon (ed.), *Power/Knowledge*, pp. 80–105. Brighton: Harvester.

16 Peat, F.D. (2002). *From certainty to uncertainty. The story of science and ideas in the twentieth century*. Washington, DC: Joseph Henry Press.

5
THEORIZING DRIFT

❖

MAN-MADE DISASTERS

In 1966, a portion of a coal mine tip (unusable material dug up in the process of mining coal) on a mountainside near Aberfan, South Wales, slid down into the village and engulfed its school. 144 people were killed, including 116 children. The post-disaster inquiry waded into a morass of commissions and bodies and agencies and parties responsible (or not) for the various aspects of running a coal mine, including the National Coal Board, the National Union of Mineworkers, the local Borough Council, the local Planning Committee, the Borough's engineering office, the Commission on Safety in Mines, Her Majesty's Inspector of Mines and Quarries and a local Member of Parliament. Collectively, there was a belief that tips posed no danger; that mining was dangerous for other reasons (which were generally believed to be well-controlled through these multifarious administrative and regulatory arrangements). While danger in the tip grew, the belief that everything was okay remained in place. Until the tip slid into the village.

A few years later, at the University of Exeter, a researcher called Barry Turner found that the inquiry reports of this and other disasters contained a wealth of data about the administrative shortcomings and failures that precipitated those accidents. Yet there was no coherent framework, no *theory*, that could do something sensible with all this data. His book *Man-Made Disasters* changed that. It came out in 1978.[1] It had begun as a doctoral dissertation, based on a qualitative analysis of 84 British accident inquiry reports (which bore the title *Failure of Foresight* – considered as subtitle of the subsequent book).

Man-Made Disasters was an opening shot in theorizing drift. This chapter will run through ideas in the literature that have tried to theorize drift one way or other. In these ideas, drift is virtually always associated with a shift in norms, a shift in what is considered acceptable. It has to do with a gradual, virtually unnoticed erosion of safety constraints, with the complexity and lack of transparency of

6

WHAT IS COMPLEXITY AND SYSTEMS THINKING?

❖

In trying to understand why things go wrong in complex systems, the safety specialist has traditionally looked at the functioning or breakdown of carefully isolated parts or phenomena. How is it that the pilot did not notice a speed decay? How can it be that the nurse did not see the bold-printed warning EPIDURAL ONLY on the infusion bag? Why did a piece of insulating foam separate from the left bipod ramp section of the Space Shuttle *Columbia* external tank at 81.7 seconds into the launch?

These are important questions, and often hold the key to unlocking at least part of a story on how things went wrong. These questions make it possible to research and systematize various detailed processes, or to find out about the state of a part or a component at or before the time of some failure. But there is something really important that such knowledge cannot tell us, and that is how a number of different things and processes act together when exposed to a number of different influences at the same time.[1] This is the question that complexity and systems thinking tries to answer.

MORE REDUNDANCY AND BARRIERS, MORE COMPLEXITY

The Newtonian focus on parts as the cause of accidents has sponsored a belief that redundancy is the best way to protect against hazard. Safety-critical systems usually have multiple redundant mechanisms, safety systems, elaborate policies and procedures, command and reporting structures as well as professional specialization. The results of such organizational, operational and engineering measures makes these systems relatively safe from single point failures. It protects them against the failure of a single component or part that could directly lead to a bad outcome. But there is a cost to this.

Barriers, as well as professional specialization, policies, procedures, protocols, redundant mechanisms and structures add to a system's complexity. With the

introduction of each new part or layer of defense, there is an explosion of new relationships (between parts and layers and components) that spreads through the system. Complexity and systems thinking says that accidents emerge from these relationships, even from perfectly "normal" relationships, where nothing (not even a part) is broken. The drive to make these systems reliable, in other words, also makes them very complex. Which, paradoxically, can make them less safe. Redundancy, or putting in extra barriers, or fixing them up better, does not provide protection against this safety threat. In fact, it helps create it, by making the systems more interactively complex.[2]

Interactive complexity refers to component interactions that are non-linear, unfamiliar, unexpected or unplanned, and either not visible or not immediately comprehensible for people running the system. Linear interactions among components, in contrast, are those in expected and familiar production or maintenance sequences, and those that are visible and understandable even if they were unplanned. But complex interactions produce unfamiliar sequences, or unplanned and unexpected sequences, that are either not visible or not immediately comprehensible. An electrical power grid is an example of an interactively complex system. Failures, when they do occur, can cascade through these systems in ways that may confound the people managing them, making it difficult to stop a progression of failure. Increasing complexity, particularly interactive complexity, has thus given rise to a new kind of accident over the last half century: the system accident.[3] In a system accident, no component needs to be broken. The accident results from the relationships between components (or the software and people running them), not from the workings or dysfunction of any component part.

This insight has grown out of what became known as systems engineering and system safety, pioneered in part by 1950s aerospace engineers who were confronted with increasing complexity in aircraft and ballistic missile systems at that time. One particular aspect of complexity fascinated and concerned safety researchers: the interconnectivity and interactivity between system components. Their relationships, in other words. The problem with this was that greater complexity led to vastly more possible interactions than could be planned, understood, anticipated or guarded against. Rather than being the result of a few or a number of component failures, accidents involve the unanticipated interaction of a multitude of events in a complex system – events and interactions, often very normal, whose combinatorial explosion can quickly outwit people's best efforts at predicting and mitigating trouble. Interactive complexity refers to component interactions that are non-linear, unfamiliar, unexpected or unplanned, and either not visible or not immediately comprehensible for people running the system.[4] The idea behind system accidents is that our ability to intellectually manage interactively complex systems has now been overtaken by our ability to build them and let them grow (like a NASA-contractor bureaucratic-organizational complex).

Just think of the introduction a new procedure to double-check something that was implicated as a broken part in some previous accident. The new procedure

relates to the old procedure and its remnants in people's memories and rehearsed action sequences. It relates to how people need to carry it out in context, to people who have to train the new procedure, to the regulator who may need to approve it. It may take time and attention away from other tasks, which in turn can create a host of reverberations throughout the system.

"Closed loop communication" was introduced as a safety practice on ships to avoid misunderstandings. Closed loop means that both the receipt and one's understanding of a message are explicitly acknowledged, for example by repeating the message. A maritime pilot, however, had observed that under certain conditions in his pilotage area, the changes of course and speed came so fast, that the helmsman's repeating of the pilot's instructions resulted in them talking at the same time (producing overlapping speech). This resulted in more communication, with more overlap. So he abandoned closed loop communication.[5]

The introduction of automation as a layer of redundancy in many systems has produced the same sorts of ironies. It has created new work, a kind of work that people generally are not good at (monitoring very reliable machinery), and introduced a host of new pathways to breakdown.[6] In fact, the explosive growth of software has added greatly to systems' interactive complexity. With software, the possible states that a system can end up in become mind-boggling. Nancy Leveson and her colleagues at MIT estimated that TCAS II, the second-generation airborne collision prediction and avoidance system now onboard all large airplanes (with more than a million lines of software code), could generate 1,040 states.[7]

How software and redundancy don't really fit together in the classical mechanical sense is apparent from a NASA study of an experimental aircraft that was executing two versions of the control system. It found that all the software problems occurring during flight testing resulted from errors in the redundancy management system and not in the much simpler control software itself, which worked perfectly.[8] Also, software removes many of the natural constraints that held engineers back in introducing even greater complexity to their systems.[9] It doesn't weigh anything, it is easy to modify, and the only capacity constraints on its functioning are system memory and processor speed, two things that have grown explosively too.

An example of a software-based system that led to multiple deaths without any broken components is the Therac-25 radiation therapy machine (technically called medical linear accelerators (linacs). These accelerate electrons to create high-energy beams that can destroy tumors with minimal impact on the surrounding healthy tissue).[10] Radiation-treatment machines were in enormous demand in hospitals throughout North America, and the Therac equipment, principally controlled by software, was widely considered the best in a growing field.

After a number of suspicious cases of overdosing by Therac-25 machines, a physicist and technician became fascinated by a "Malfunction 54" message that had flashed on the screen during a treatment (the patient fell into a coma and died three weeks later). They typed and retyped the prescription into the computer console, determined to re-create Malfunction 54. They went to the bottom of the screen and then moved the cursor up to change the treatment mode from x-ray mode to electron mode, over and over, for hours. Finally they got the malfunction.

What made the difference was the speed with which the instructions were entered. The computer would not accept new information on a particular phase of treatment (in this case changing the x-ray mode to electron mode) if the technician made the changes within eight seconds after reaching the end of the prescription data. That's what Malfunction 54 meant (but didn't say). If the changes were made so soon, all the new screen data would look correct to the technician, but inside the computer, the software would already have encoded the old information.

That meant the beam on the Therac-25 was set for the much stronger dose needed for an x-ray beam while the turntable was in the electron position. The software running inside the computer included no check to verify that various parts of the prescription data agreed with one another.[11]

Accidents in software-rich systems, such as the Therac-25 but also the Ariane 5 rocket or the NASA Mars Polar Lander, show that none of the components in these systems had to fail in order for things to go terribly wrong. All components can meet their specified design requirements, and still a failure can occur.[12] Better component design, or better protection against failures of individual components (for example, barriers, redundancy) is not the answer to such failures. It is not the answer to systems accidents. The answer lies in understanding relationships. Which is what complexity and systems thinking try to do, particularly when applied to socio-technical organizations.

UP AND OUT, NOT DOWN AND IN

Not long ago, a colleague and I were asked to help a maintenance company find out what factors had "caused" a recent increase in safety events and incidents. For a few months we held our ear to the ground – walked around, interviewed people throughout the organization, even sent out a survey which people answered with enthusiasm. Then the day came to report back to management. "So," my colleague asked me, "what do we say the causes are?"

I was quiet for a moment. Because what could we say in response to that question? I had no idea what the "causes" were, even after doing all the analytic work together. It was fascinating, in fact, that management wanted the "causes"

of their problems pinpointed, just like in a broken machine. They wanted us to tell them which screws to tighten, which lines to replace, which areas to lubricate. This was not surprising, probably, for a company whose business consists of doing just that.

But such a Newtonian–Cartesian, or mechanistic understanding of organizations is not limited to people who are in the mechanics business. My undergraduate professor in industrial psychology, decades ago, illustrated the same thinking by putting his hands beside his ear, as if he were holding a small box in the air. "What you want to do is shake that organization around, to find out what's loose in there, what's rattling," he said. That was the recommendation to a large group of wannabe consultants and organizational psychologists. Shake the box, open the box, find the loose part, fix it, close the box, hand it back.

In the case of the maintenance organization that my colleague and I tried to help more recently, we never found what was rattling loose inside. We did identify various broad influences that seemed to shape how people related to their work, what they found important, what they felt encouraged to pay attention to. We began to trace how a number of different things and processes acted together when exposed to a number of different influences at the same time.

The recent spin-off of the maintenance organization, that aimed to have it stand on its own commercial legs, had led to a tension between intensive and extensive growth – should the company nurse the depth and quality of its core processes for its only large customer (the organization of which it had, until recently, been part?) or should it aggressively court outside business and broaden its offerings? Organizations like this produce means-ends oriented social action through its formal structures and processes. Those aim to achieve certainty, conformity and goal attainment. But if you suddenly expose the organization to a new competitive environment, which offers a host of previously unencountered and unexplored threats and opportunities, it can quickly wash away the basis of certainty for what goals need attaining, for what social expectations there are and which ones are considered legitimate by whom, for what norms should be conformed with.

People in core business units stopped reporting incidents: aspects of the new environment made the entire notion of incident negotiable, not just for those facing them on the work floor, but also for those hearing about them higher up. Even when things were reported, little to nothing was done. How exactly such larger influences of a new competitive environment travel down to the heads of individual employees to help determine what they will see as rational at any moment in their work is still anybody's guess, even that of current theorizing. So we could never "pinpoint" the organizational spin-off as a cause.

To make things easy, our story could have ended with a description of the limits on cognitive performance by individuals (the "parts" in the system), how fatigue, language difficulties in the communication with contractors, lack of motivation, goal displacement and role ambiguity all constrained the ability of individuals to do their jobs conscientiously, thus leading to incidents.

But we tried to lift the story out of a fixation with parts. We refused to just go down and into the box to find what was rattling. We went up and out, trying to sketch the influences of the new environment on the organization's structure and relationships (management, individuals, complexity), its processes (patterns of interaction that produce action, including deviance and mistake) and people's tasks (their technical uncertainty and inexactitude, the imperfect knowledge and interpretive flexibility necessary to go from written guidance to action in context). We never enumerated certainties, only possibilities. We never pinpointed, only sketched. That is what complexity and systems thinking does.

Mechanistic thinking about failures, that is, the Newtonian–Cartesian approach, means going down and in. Understanding why things went wrong comes from breaking open the system, diving down, finding the parts, and identifying which ones were broken. This approach is taken even if the parts are located in different areas of the system, such as procedural control, supervisory layers, managerial levels, regulatory oversight. In contrast, systems thinking about failures means going up and out. Understanding comes from seeing how the system is configured in a larger network of other systems, of tracing the relationships with those, and how those spread out to affect, and be affected by, factors that lie far away in time and space from the moment things went wrong.

Take a simple question: How does a cell phone work? To understand systems thinking, or the difference between going down and in versus up and out, this is actually a useful question. The mechanistic answer would begin by screwing the back plate off the phone. What are the parts that are inside? How do they get energized, how do they interact, how do electrical signals flow from input to output? And then go deeper still. What is the software that governs these flows, where are the bits and bytes that hold it all up? The behavior of the phone can be brought down to the behavior of its individual components. It is just the aggregate of all of those component behaviors.

The systemic answer would do something radically different. It would go up and out, locating the cell phone in the complex ecological web of markets and the flows of goods and people that supply its parts, and create the demand for it. One place where the answer could end up is the massive Kahuzi-Biega National Park in eastern Congo in Africa, where the elephant population has been virtually wiped out, and a previously fairly healthy gorilla population has been reduced 50 percent to 10 percent of its original size.[13]

The prize to be found there is Coltan, short for Columbite-Tantalite, a metallic ore. When refined, Coltan becomes a heat resistant powder with unique properties for storing electrical charge (tantalum capacitors control current flow in cell phone circuit boards). Eighty percent of it, worldwide, comes from the eastern Congo. Coltan is manually mined by groups of men. They don't need much more than a shovel. They dig basins, or craters, in streams by scraping off the surface mud. They then slosh the water around the crater, which causes

the Coltan ore to settle to the bottom where it is then retrieved by the miners. A group can harvest one kilogram (about two pounds) of Coltan per day. When economic times are good, and cell phones are in demand, a Coltan miner can earn as much as US$200 per month, compared to a salary of US$10 per month for the average Congolese worker.

The cell phone, through the mining of Coltan, links with local civil and cross-national wars. In less than 18 months, the Rwandan army made a quarter of a billion dollars from selling Coltan. The strange thing is that Rwanda has no Coltan. But it borders on the eastern Congo. From there, links go up to Brussels and Antwerp. Vendors in Belgium, the colonial master of Zaire (Congo's predecessor) play a role in smuggling and reselling Coltan. The cell phone also links to gorillas and elephants. In areas where Coltan is mined, the ground is cleared to make mining easier, destroying and contaminating the animals' habitat. The cell phone also links to poverty, exploitation and child labor, which in turn has its basis in the displacement of the local populations by the miners and roving armies. Which again links back to the gorillas. Gorillas are killed and their meat is sold as bush meat to the miners and the varying and fluctuating rebel armies that control the area. Alternative sources of Coltan are difficult to find on this planet, and assurances of it being ecologically responsible, humanitarian or "gorilla-safe" are hard to substantiate.

So how does a cell phone work? The systemic answer says little about how a cell phone works as a self-contained, engineered piece of micro-technology. But it could argue that a cell phone works in part by destroying livelihoods while briefly benefiting a few others, that it works by fueling and sustaining civil and regional wars, and by wiping out local populations of elephants and gorillas. That's not a value statement or moral indictment. It is systems thinking: going up and out, trying to understand how influences from many things relate to and impact on many other things; how the system in question (the cell phone) is configured in networks of other systems – social, natural, political and post-colonial, humanitarian, ecological.

SYSTEMS THINKING

The disenchantment with going down-and-in to explain the whole became acute at the turn of the twentieth century. Ironically, it was the same time that the success of the Newtonian model was being celebrated, and its continuing dominance was predicted: 1900, the year that Lord Kelvin honored the Newtonian method in front of the assembled Royal Society in London. But in that same year, Albert Einstein graduated from the Zurich Polytechnic Academy. Max Planck published his first paper on the quantum, while Henri Poincaré started running into abstruse difficulties with Newtonian mechanics. This would lead to chaos theory more than half a century later.[14]

In the first decades of the twentieth century, several scientific disciplines started to show serious engagement with what we now call systems thinking. The emphasis on parts, as Boyle had pursued, Descartes had theorized and Newton had made lawful, began to show its limits in the explanatory power and development potential of various sciences at that time. Physics was, of course, one of them, and, particularly given its status as Descartes' "stem" in the system of sciences, quite a prominent one.

Relativity theory, developed almost entirely by Einstein himself, and quantum theory, which had a larger gaggle of physicists behind it, both served to change the view of physics forever – and with it our understanding of the universe. The theory of relativity captured why the world looks different to observers moving at different speeds and that there is no absolute frame of reference against which all speeds can be measured. For Einstein, though, fundamental laws of an objective nature underlie such relative appearances, a position that would later turn him into the "last classical physicist." Quantum theory, which grew out of classical experimental investigations into the make-up of atoms, revealed sensational and totally unexpected results. Far from yielding hard, basic, minimalist particles, the subatomic world was not populated by the solid objects predicted by Newton and classical physics:

> At the subatomic level the interrelations and interactions between the parts of the whole are more fundamental than the parts themselves. There is motion, but, ultimately, no moving objects; there is activity but there are no actors; there are no dancers, only the dance.[15]

Quantum physics showed that the Newtonian–Cartesian strategy of decomposition would, quite literally, lead to nothing. Decompose matter far enough, down to the subatomic level, and you are left not with basic particles, but with relationships, with mere probability patterns. Matter is not made out of smaller matter. It is made, really, out of "nothing" when you apply the terms of classical Newtonian physics. The parts cannot explain the whole, because the parts are basically not there. "At the subatomic level, the solid material objects of classical physics dissolve into wave-like patterns of probabilities. These patterns, furthermore, do not represent probabilities of things, but rather probabilities of relationships."[16]

Quantum physics showed that the universe could never be understood as a collection of isolated objects, and the entire notion of separate particles as the bearer of a fundamental explanation began to crumble. The most basic matter in the universe could not be explained as individual entities, but must rather be understood as complex webs of interrelationships, which were not mathematically predictable. A particle was not an independently existing, unanalysable entity. It was a set of relationships that reach outward to other things.[17] It began to dawn on scientists that laws and predictions could be expressed only as probabilities, not certainties. The outcomes of a complex system could not be predicted. The only thing that could be predicted was the approximate likelihood of something happening. Its probability, in other words.

One of my students recently argued that we really shouldn't be too pessimistic in our outlook for safety. Aviation, as a world-wide industry, is not drifting into failure. After all, he said, haven't we made enormous progress? Look at the accident statistics and the number of fatalities, he said. Of course, there are spikes and variations. But year over year, we seem to get better and better, not drifting into something worse.

Well, another said, this depends on how you define progress, and where you draw the boundary of the fatalities that the system may be thought to be responsible for (and, indeed, how you define such "responsibility"). We generate energy in aviation the very same way as we did in pre-history. We burn carbon-based material, just as we did when we were dwelling in caves. Doing the same thing for 10,000 years? We should not really count that as progress. What's more, how do we know that this is producing fewer fatalities? Doesn't that depend on where we look for system interconnections and where we start and stop counting dead bodies? The traditional way of assessing fatalities in aviation is to count the victims in the crashed airplane and any additional victims on the ground. That seems to make good sense, of course. But perhaps it is also a quite limited way of looking at the fatal consequences of flying. And, in a sense, a very parochial and self-indulgent one. We count only those in the world who were already fortunate (and affluent) enough to take direct part in the activity (which in itself is an increasingly sizable portion of humanity). But what about other people who are affected?

For example, we might want to look for connections between aviation activity, global warming, and increasing wars in Africa. The body count would go up dramatically, changing the notion of "progress" if only we are willing to include a bigger slice of humanity in the picture. Since 1990, emissions of carbon dioxide by aviation have risen 90 percent. According to some reports, aviation's share of CO_2 emissions could rise from about 2 percent globally today, to as much as 20 percent in 2050.[18] Of course, enormous economic opportunities are created by such global interconnections and their increase, in ways that are not very dependent on land-based infrastructure. So that could count as progress along a number of dimensions, for sure.

But there are other effects. The release of CO_2 into the atmosphere (each kilogram of aviation fuel that is burned basically adds two kilograms of CO_2 to the atmosphere) has been linked not only to a global increase in temperature, but to changes in regional climates as well.[19] The earth, of course, has always known temperature cycles and fluctuations in atmospheric gas composition. And only some of the CO_2 in the atmosphere comes from human activity. Since the dawn of the Industrial Revolution, however, this is an increasing (and, lately,

accelerating) proportion. Today the coupling between anthropogenic CO_2 and climate change (or global warming) is seen as proven and lawful by many; others dismiss it as amusing and as junk science.

But changes in CO_2 concentrations (whatever their source) do have effects on regional climates. Using what they called a zonally symmetric synchronously coupled biosphere-atmosphere model, NASA and MIT scientists showed a connection between CO_2 concentrations and changes in the biosphere-atmosphere system in Africa. Changes in CO_2 concentration make certain regions wetter,[20] and, by extension, certain regions drier.[21] Other scientists have found a strong link between drier years and the prevalence of conflict in Africa. Long-term fluctuations of war frequency follow cycles of temperature change: there is more conflict during warmer years. Rain-fed agriculture accounts for more than 50 percent of gross domestic product, and up to 90 percent of employment, across much of the continent. A drier, warmer climate decreases agricultural output (a 10–30 percent drop per °C of warming), which in turn degrades economic performance, which in turn increases conflict incidence. This holds even after correcting for per capita income or democratization (increases in both tend to reduce war risk).

Warming, then, has a role in shaping conflict risk, with projected climate changes increasing armed conflict by 54 percent and bringing an additional 393,000 African deaths by 2030.[22] That would be like 20,000 deaths each year. If aviation's share of CO_2 production indeed rises from 2 percent to 20 percent over the next 40 years, its share in those deaths would drift from about 400 per year today to about 2,000 fatalities per year in 2030. In aviation, such figures would not count as progress; it could count as another slow, steady drift into failure. Also, it would outpace any beneficial effects of African economic growth and democratization (some of which, paradoxically, are also brought by aviation with its role in stimulating economic development and bringing people and ideas and goods and services together).

Of course, there are a lot of numbers in the example above. But the example is not about the numbers, it is about the possible relationships, and about where we believe we should draw the boundaries and call something a "system." Such boundaries are always exclusionary. Wherever we draw them, something will fall outside of it. Systems thinking is not saying that there should be no boundaries (because they are an essential part of defining what a system is), but that we can at least be more flexible, more imaginary, and perhaps more democratic in where we draw those boundaries.

And, as in all complex systems, other factors could play a plausible role too. All explanations for how system behavior is the result of relationships are tentative, and open for change when better arguments come in. It is not necessary, for example, that additional conflict deaths due to temperature increases are the result of regional warming. There is a known correlation between violent crime and higher temperatures, which could account for at least some of those

additional deaths,[23] and even non-farm labor productivity can decline with higher temperatures, meaning that declines in agricultural output do not have to act as an intermediary variable between more heat and more deaths.[24] It is consistent with complexity thinking, then, when researchers say "We interpret our results as evidence of the strength of the temperature effect rather than as a documentation of the precise future contribution of economic progress or democratization to conflict risk. Similarly, we do not explicitly account for any adaptations that might occur within or outside agriculture that could lessen these countries' sensitivities to high temperatures, and thus our 2030 results should be viewed as projections rather than predictions."[25] Being modest about the reach of one's results, and explicitly open to their being wrong or open for revision, was once seen as a weakness in heavily naturalized science. But in complexity and systems thinking, it is a strength.

The example shows how physics was by no means left alone in this drift away from classical Newtonian science. In fact, systems thinking was in part pioneered by biologists, who began to look at living systems not as collections of components but as irreducible, integrated wholes. Social sciences and humanities have similarly put up ever stronger resistance over the past few decades against the wholesale naturalization of their research work. The only way to be "scientific" was once to be like physics (in fact, for centuries being more scientific meant having to be like physics). That meant that research had to be done in the form of carefully controlled experiments, in which the scientist had clear control over one or a few variables, tweaking them so that she or she or he could exactly understand what had caused what. Preferably, the results were expressed as, or converted into, numbers, so that the claim that something fundamental had been found about the world looked even more credible. This is the way of doing science that Newton and Descartes proposed. Ironically, with a shift of natural sciences into systems thinking and complexity, and social sciences and humanities doing the same, it seems that they once again seem to come closer to each other. Not because the social sciences and humanities are becoming "harder" and more quantifiable, but because natural sciences are becoming "softer" with an emphasis on unpredictability, irreducibility, non-linearity, time-irreversibility, adaptivity, self-organization, emergence – the sort of things that may always have been better suited to capture the social order.

It is interesting to see that complexity and systems thinking predates what we regard as the Scientific Revolution in the sixteenth and seventeenth centuries – a revolution that gave rise to the Newtonian–Cartesian hegemony. Leonardo (born Leonardo di Ser Piero, Vinci, Florence, 1452–1519) is regarded as science's first system thinker and complexity theorist. An artist, engineer, architect, relentless empiricist, experimentalist and inventor, Leonardo embodied the humanist fusion of art and science. He embraced a profound sense of the interrelatedness of things, interconnecting observations and ideas from different disciplines (for example, physiology and mechanical engineering; aerodynamics and music) so as to see problems in a completely new light and come up with nifty solutions.

Leonardo embraced his science as complexity – it had no boundaries since his goal was to combine, advance, investigate and understand processes in the natural world through an inter-disciplinarian view. He was a committed systems thinker, eschewing mechanistic explanations of phenomena and instead giving primacy to ecological ones – thinking up and out, rather than only down and in, and always seeing (and, in his art, depicting) linkages among living organisms (people, groups, organizations) so as to reveal new solutions. Leonardo was fascinated by the phenomenon of flight and produced many studies on the flight of birds (for example, his 1505 Codex). He designed a variety of aircraft – of which one, a hang glider, was recently demonstrated to actually work.

COMPLEX SYSTEMS THEORY

With his formulation of General Systems Theory, Von Bertalanffy helped establish a serious scientific foundation for the alternative to Newtonian–Cartesian thought in the early 1970s.[26] Today that alternative is known as complexity and systems theory. Recently, Cilliers summarized the characteristics of complex, as opposed to Newtonian, systems nicely.[27] I sum the points up as follows, and will explain them in more detail below:

- Complex systems are open systems – open to influences from the environment in which they operate and influencing that environment in return. Such openness means that it is difficult to frame the boundaries around a system of interest.
- In a complex system, each component is ignorant of the behavior of the system as a whole, and doesn't know the full effects of its actions either. Components respond locally to information presented by them there and then. Complexity arises from the huge, multiplied webs of relationships and interactions that result from these local actions.
- Complexity is a feature of the system, not of components inside it. The knowledge of each component is limited and local, and there is no component that possesses enough capacity to represent the complexity of the entire system in that component itself. This is why the behavior of the system cannot be reduced to the behavior of the constituent components, but only characterized on the basis of the multitude of ever-changing relationships between them.
- Complex systems operate under conditions far from equilibrium. Inputs need to be made the whole time by its components in order to keep it functioning. Without that constant flow of actions, of inputs, it cannot survive in a changing environment. The performance of complex systems is typically optimized at the edge of chaos, just before system behavior will become unrecognizably turbulent.
- Complex systems have a history, a path-dependence. Their past is co-responsible for their present behavior, and descriptions of complexity have to take history into account.

- Interactions in complex systems are non-linear. That means that there is an asymmetry between, for example, input and output, and that small events can produce large results. The existence of feedback loops means that complex systems can contain multipliers (where more of one means more of the other, in turn leading to more of one, and so forth) and butterfly effects.

Complex systems are not closed. They don't act in a vacuum like Newton's planetary system. Socio-technical systems or organizations are open systems. They are in constant interplay with their changing environment, buffeted and influenced by what goes on in there, and influencing it in turn. Whereas the systems studied by high reliability theory (see Chapter 4) were relatively closed (an aircraft carrier at sea, power grids before deregulation, air traffic control run by a single government entity, the Federal Aviation Administration), this is not true for many other safety critical systems. NASA, for example, has operated in a (geo-)political and societal force field for decades, suffused with budgets and operational priorities and expectations that never were her own but seeped in from a number of directions and players (congress, the defense department, other commercial space operators) and affected what were seen as rational trade-offs between acute production pressures and chronic safety concerns at the time. Of course the influence doesn't just go one way. NASA itself also affected what politicians saw as reasonable and doable, in part because of its own production, and its actions would have affected the actions of other commercial space operators too (for example, the choice by the French Ariane consortium not to use manned vehicles for the launching of satellites). This same openness goes for healthcare too: it often is a handmaiden of local or national politics and funding battles – hay gets made from promises to bring down surgical waiting lists, for instance. This can have effects on how funding is allocated, on where resources are added and where they are taken away.

Open systems mean that it can be quite difficult to define the border of a system. What belongs to the system, and what doesn't? This is known as the frame problem. It is very difficult to explicitly specify which conditions are not affected by an action. When you trace the ever-changing webs of relationships in the mining of Coltan for the production of cell phones, for example (see above), you will find that actions by local people reverberate across an almost infinite range of economies and ecologies, affecting not only wildlife in the eastern Congo, but everything from inflation of food prices in local villages surrounding Kahuzi-Biega, to the allure of discarded Russian cargo planes and those who remember how to fly them, to the volatility of commodity markets and demands of just-in-time production of electronics in Taiwan, to the prosperity of particular factions in the underworld of Belgium. And all of these reverberations produce further reverberations, in Russia, Belgium, Taiwan, affecting people and markets and all kinds of relationships there too.

One solution of complexity and systems theory is to determine the scope of the system by the purpose of the description of the system, not by the system itself. If the purpose is to describe the effects of Coltan mining on local gorilla

populations, then you might draw the geographic system boundaries not far from Kahuzi-Biega. Then you can trace ecological contamination and physical habitat destruction, and you might restrict functional boundaries to bush-meat prices, hunting methods, and numbers of miner mouths to feed. Never mind Russia or Belgium or Taiwan. But such boundaries are a choice: a choice that is governed by what you want to find out. System theory itself provides no answer to the frame problem. Where you place the frame is up to you, and up to the question you wish to examine. The adage of forensic science to "follow the money" is the same commitment. On the one hand, it leaves the observer or investigator entirely open to where the trail may take her or him and how it branches out into multiple directions, organizations, countries. That is where the system of interest is open. On the other hand, the commitment frames the system of interest as that which can be expressed monetarily.

This highlights a very important aspect of the post-Newtonian perspective: the world we can observe isn't just there, completely and perfectly ready-formed and waiting for us to discover. We ourselves play a very active role in creating the world we observe, precisely because of our observations. If we don't have the language or the knowledge to see something, we won't ever see it. What our observations come up with is in large part up to us, the observers. In post-Newtonian science, it is very hard to separate the observer from the observed; to say where one ends and the other begins. In post-Newtonian science, reality does not contain the sort of hard, immutable facts that have a life entirely independent of a human observer. Human observation cannot be the neutral arbiter or producer of knowledge. After all, the observers themselves impose a particular language, interest, and they come with imaginations and a background that actively bring certain aspects of a problem into view, while leaving others obscured and unexamined.

In a complex system, each component is ignorant of the behavior of the system as a whole. This is a very important point.[28] If each component "knew" what effects its actions had on the entire rest of the system, then all of the system's complexity would have to be present in that component. It isn't. This is the whole point of complexity and systems theory. Single elements do not contain all the complexity of the system. If they did, then reductionism could work as an analytic strategy: we could explain the whole simply by looking at the part. But in complex systems, we can't, and analytic reduction doesn't work to enhance anybody's understanding of the system.

Complexity is the result of a rich interaction, of constantly evolving relationships between components and the information and other exchanges that they produce. Rich interaction takes a lot of components, which is indeed what complex systems consist of. Of course a beach has a lot of components (sand kernels) too, but that alone doesn't make it complex. The components have to interact. They have to give each other things like energy, information, goods. Those interactions, however, are limited, or local. Each component responds only to the limited information it is presented with, and presented with locally. The Coltan miner has no idea about the price of retired Russian cargo planes, and might not even recognize one if he saw it. What he mines is taken away

through the bush on the backs of other men, perhaps with the use of animals or motorbikes, then onto trucks if they can make it that far into the jungle. He may not even know what he is mining, what the stuff is for, or why it is worth so much. The pilot flying the plane to a nearby airstrip is probably ignorant about gorillas, and may not know who really offloads his cargo at the receiving end. He makes the trip not because he works for a major cell phone manufacturer, but because he gets U.S. dollars in bundles of cash at the end of it.

Components in the system only act on and respond to information that is available to them locally. This locality principle again confirms the post-Newtonian view on knowledge and reality, as it makes access to some pre-existing, objective external reality not only impossible but also irrelevant. In the words of Heylighen, Cilliers and Gershenson:

> According to cybernetics, knowledge is intrinsically subjective; it is merely an imperfect tool used by an intelligent agent to help it achieve its personal goals. Such an agent not only does not need an objective reflection of reality, it can never achieve one. Indeed, the agent does not have access to any 'external reality:' it can merely sense its inputs, note its outputs (actions) and from the correlations between them induce certain rules or regularities that seem to hold within its environment. Different agents, experiencing different inputs and outputs, will in general induce different correlations, and therefore develop a different knowledge of the environment in which they live. There is no objective way to determine whose view is right and whose is wrong, since the agents effectively live in different environments ('Umwelts') – although they may find that some of the regularities they infer appear to be similar.[29]

That interactions between components in a complex system play out over fairly short ranges, and information is received primarily from immediate neighbors, does of course not mean that the ultimate influence of local agents is limited to those ranges. First, some agents physically travel through the system (like that pilots flying the cargo plan with Coltan), and will thus interact with multiple local settings along the way. Second, actions reverberate through further relationships, and thus affect what goes on in another part of the system. But because these reverberations pass through other components and their relationships, they can be modulated, suppressed, amplified.

THE BUTTERFLY EFFECT

One reason why this happens is because the interaction in complex systems is typically non-linear. It is non-linear not just in a mathematical sense (of equations whose terms are not of the first degree), and not in a spatial geometric sense (not arranged in a straight line), but rather in a physics sense, where non-linear means a lack of linearity or symmetry between two related qualities such as input and output. Non-linearity guarantees that small actions can eventually have large results. The amplification, or multiplication of effects in a complex, open system

(a kingdom doing battle) was noted in a poem that dates back to at least the fourteenth century):

> For want of a nail, the shoe was lost.
> For want of a shoe the horse was lost.
> For want of a horse the rider was lost.
> For want of a rider the battle was lost.
> For want of a battle the kingdom was lost.
> And all for the want of a horseshoe nail.

As the interactions unfurled through the system, from nail to shoe to horse to rider to battle to kingdom, their effects spread and multiplied. Recognize how this contrasts with Newtonian science, which holds (according to the third law of motion) that the size of the effect is equal (and opposite) to the cause. Such symmetry is typically not the case in complex systems. This is not only related to nonlinearity, but to the nature of the feedback loops themselves, which in a complex system are typically not unidirectional. The effect of any action can feed back onto itself, sometimes directly, and sometimes through multiple intervening stages. When a favorite rider is out of action it affects surrounding riders, then surrounding units. The crumbling of morale through the ranks, then, could be one of those loops. The notion of multipliers has since then become a formal one in complexity theory.

Man-made CO_2 production could once be compensated by natural processes, but these may now no longer be able to keep up. As the oceans absorb CO_2, for example, they become more acidic. This, combined with increasing ocean temperatures, diminishes their ability to continue absorbing CO_2, and more CO_2 stays in the atmosphere. In 1960, a metric tonnne (1,000 kilograms) of CO2 emissions resulted in around 400 kilograms of CO_2 that remained in the atmosphere. In 2006, 450 kilograms of the same tonne remain in the atmosphere.[30] Hence a tonne of CO2 emissions today results in more heat-trapping capacity in the atmosphere than the same tonne emitted decades ago.

In a similar development, published in *Science* in 2010, it appears that methane coming out of the East Siberian Arctic Shelf (in addition to that coming from nearby Siberian wetlands) is comparable to the amount coming out of the entire world's oceans. The release of long-stored methane (a greenhouse gas more powerful than CO_2) is itself a result of global warming, which melts ice and permafrost and allows methane to bubble free from its ground storage (whether sub-sea or subterranean). In recent measurements, more than 80 percent of deep water in the East Siberian Arctic Shelf had methane levels between 8 and 250 times greater than background levels. That said, the water in the East Siberian Arctic Shelf isn't actually that deep, which means the methane coming up doesn't have enough time to oxidize, which means more of it escapes into the atmosphere. As it does, global warming could increase even more quickly, which in turn speeds up the release of more methane. This is a feedback loop, or a multiplier, where more means more. It is a feature of a complex system that is very hard to stop.[31]

The idea that small events can produce large results because of non-linearity is popularly known as the butterfly effect. The technical term, which was introduced in Chapters 1 and 2, is sensitive dependence on initial conditions. The phrase "butterfly effect" is related to the work by Edward Lorenz, a mathematician and meteorologist, and comes from the idea that a butterfly's wings might create tiny changes in the atmosphere that can ultimately balloon into a hurricane, or alter the path of a brewing hurricane, or delay, accelerate or even prevent the occurrence of a hurricane in a certain location. The flapping wing represents a small change in the initial condition of the system. This small change triggers a non-linear succession of events, leading to large-scale alterations of the state of the system. Had the butterfly not flapped its wings, the trajectory of the system might have been vastly different. Of course the butterfly does not "cause" the tornado in the sense of providing the energy for the tornado (indeed, that would be a Newtonian reading of "cause" with symmetry between cause and effect and a preservation of the total amount of energy through the causal chain). But it does "cause" it in the sense that the flap of its wings is an essential part of the initial conditions eventually resulting in a hurricane, and without that flap that particular hurricane would not have existed.

How small events can contribute to large results is a popular theme in the story of Anthony Eden, Foreign Secretary of the U.K. during the Suez crisis, and it is reminiscent of the loss of a kingdom (or empire) for want of a nail.[32] In 1952, a year after becoming Foreign Secretary in the new Churchill government, Anthony Eden suffered several attacks of upper abdominal pain, followed by more severe pain a year later. Sir Horace Evans, physician to the Queen, was called in and advised urgent surgery. On 12 April 1953 at the age of 55, surgery was performed in London by Mr John Basil Hume, surgeon at St Bartholomew's Hospital. While a very senior surgeon, Hume was not one with large amounts of recent experience. It is not entirely certain what happened during this operative procedure. But during cholecystectomy, Eden likely sustained an injury to the proximal common hepatic duct at its bifurcation and possibly, also, an injury of the right hepatic artery. In common terms, people would call it a botched gallbladder operation, which in itself is not uncommon.[33]

On 26 July 1956 Colonel Nasser seized the Suez Canal. This followed a long period of Egyptian opposition to what they regarded as a Canal Zone occupation, starting with King Farouk's demand for total and immediate withdrawal of British troops from the Suez Canal in 1950.

Ten weeks later, on 5 October 1956, Eden collapsed with a high temperature of 106°. This, more than three years after his bile duct repair, was the first of several major attacks of pain and fever.

Three weeks after his collapse on 27 October 1956, Israel invaded Egypt and a further 4 days later, Eden ordered the Anglo-French forces to occupy the Suez Canal Zone on the pretext of separating Egypt and Israel. In reality, the British government wanted the canal returned to international control to safeguard the oil supply of Western Europe, which passed through Suez. Eden also believed that Nasser might fall from power if he was forced to back down on his nationalization of the canal, and thus the Egyptian's brand of radical pan-Arab nationalism would be marginalized. But Eden's action caused an uproar in the British parliament, in the U.S.A and also at the United Nations. In response, all parties agreed to an early ceasefire on 6 November 1956, and 3 weeks after this occupation, on 23 November 1956, the Anglo-French troops began to withdraw from the Suez Canal Zone. Eden's action eventually had the geopolitical effect opposite to its intentions: he lost British control over Suez.

Prior to his collapse in July 1956, Eden was widely acknowledged by public servants working with him and by his many biographers, as a cool composed man, an expert in the use of diplomacy even under the most difficult circumstances. He believed in the rule of the law and in the supremacy and effectiveness of the United Nations. As Foreign Minister during Churchill's peacetime ministry, Eden had followed a global strategy that understood the post-war limitations on British power. Pursuing a military option by calling in the Anglo-French forces was, therefore, most uncharacteristic. At that time he was irritable, quick tempered, often tired, and most uncharacteristically, conspiratorial with France and possibly with Israel (although the latter has been denied). Uncharacteristically also, he upset the U.S.A., especially John Foster Dulles. Eden resigned 10 weeks later. He remained in poor health for the rest of his life.

ADAPTATION AND HISTORY

Complexity is a characteristic of the system, not of the parts in it. Recall that the Newtonian reflex – indeed, that of science in general – has been to see localized structures (for example, humans, or technical parts) as the prime causal agents. Without those, nothing gets organized, nothing happens. But complex systems do not rely on an external designer, and they do not rely for their continued functioning on some central internal coordinating agency or nerve center to specify how all the parts are going to work. Complex systems can keep on working in changing environments because the relationships among their parts can adjust themselves. These reconfigurations are a response to what goes on in the system, in the environment and in the multitude of interactions between them. Thus, complex systems are adaptive, and they can be resilient precisely because of their complexity. All the patterns of interaction between components make for a complex system. Webs of relationships wax and wane and adapt to what happens in the environment around it.

Drug cartels from South America have moved into West Africa, establishing themselves all over the region and opening up a new route for transporting cocaine from the plantations of Colombia, Peru and Bolivia to the consumers of Europe, particularly Britain and Spain. In recent years effective patrolling of traditional trans-Caribbean and transatlantic smuggling routes has forced the cartels to seek out new paths, which is where West Africa comes in: the shortest line of latitude westwards from the ports and airstrips of South America reaches land in Guinea-Bissau, which received the dubious honor of recently becoming West Africa's first "narco-state." Five years ago the amount of cocaine shipped to Europe via West Africa was negligible. Today 50 tonnes a year worth up to 2 billion dollars pass through the region. Interpol estimates as much as two thirds of the cocaine sold in Europe in 2009 alone will have reached the Continent via West Africa.

"We are seeing multi-ton shipments transiting West Africa. We have recorded arrests of Latin Americans all over West Africa," said Antonio Mazzitelli, at the regional office of the United Nations Office for Drugs and Crime in Dakar, Senegal. "They are using ships, speedboats, small and large airplanes, 4WDs. There is really no limit to the imagination of traffickers."

The cartels work with local criminal gangs and officials. Some consignments are for corrupt armies, customs and police forces. Plastic-wrapped cartons, each containing up to fifty 1-kg bars of cocaine, are divided up and distributed by gangs who use speedboats, trans-Sahara trucks, aircraft or human "mules," who board commercial flights to traffic the drugs to Europe. Spain is the usual entry point, with the northeastern Galicia region being the favorite place owing to the maze of inlets along its coastline. There is evidence that Irish gangsters, Balkan mobsters and Italy's Calabrian mafia are also involved.

The sheer value of drugs transported through West Africa – one of the world's poorest regions – dwarfs entire economies and corrupts security forces and politicians. Guinea is just one of more than a dozen countries in the region in varying states of disarray and poverty, many only recently emerging from years of bloody civil war fought over the control of "blood diamonds" and other resources. Cocaine, experts warn, is another resource that the gangs consider worth fighting for. "What we're seeing is the criminalization of the state as a result of drug trafficking," Corinne Dufka, a West Africa expert at Human Rights Watch, said. "The unlimited cash at the gangs' disposal risks toppling these desperately weak states."[34]

In the example above, individual traffickers seem very smart. They deflect routes away from patrols, and use a diversity of means of transportation. But it is unlikely that there is one single mastermind who has plotted it all out and who can keep up with plotting the right moves in response to a full knowledge of all the countermoves. Again, the mastermind would have to be as complex as the system

that is being run. This is computationally impossible: the capacity for all that information (and its dynamics) exceeds what individual people or even computers can handle. Nor do they need to – that is the beauty of complex systems.

Rather, the traffickers in western Africa respond to local opportunities and constraints, just like the Latin Americans did by finding them as points of contact and throughput when flying or boating across the Caribbean became too risky. Local conditions drive local responses. If land-based traffic by 4-wheel drive vehicles gets difficult, speed boats will be used more to pace along the coast of West Africa. If policing becomes a little too effective on the northern plains of Guinea because of a surge of Interpol funding, then neighboring Guinea-Bissau becomes a hub instead, where local smugglers can make use of connections into the power elite and their military arms (as they did with the Red Berets in Guinea, once commandeered by the late President Conte in Conakry). That these patterns of adaptation play out in similar ways but at different scales (across the continents of Latin America and Africa, between the countries Guinea and Guinea-Bissau, across the different regions, towns, means of transportation and even fields or roads to use or avoid) is another property of complex systems. It is known as recursion (or fractals when speaking in geometric terms), where stochastically self-similar processes operate at different relative sizes or orders of magnitude. Such patterns, recognizable at different levels of resolution when studying complex systems, are one way of finding some order in how complexity works, some way of mapping what goes on there.[35]

Complex systems operate under conditions far from equilibrium. There has to be a constant flow of inputs to keep the system and its goals on track. Not making those inputs, those changes, means that the system will disintegrate, will stop functioning. In biological systems, equilibrium means death. For the smugglers of West Africa this goes too. If routes and methods are kept constant or stable, then they would be quickly shut down. It is precisely because they are only dynamically stable – able to function because of making inputs and responding to the environment the whole time, as if riding a bicycle – that they are able to survive in a changing world. They actively compensate perturbations that originate in the environment. Of course, by doing so, the environment will change too. The greater the variety of perturbations that a system has to respond to, the greater its diversity of responses has to become. It has to perform a greater variety of actions to stay in business.

If we look at the language that is often still used to describe activities like cocaine smuggling and their countermeasures, we can quickly see how it remains problematic and limiting. Calling it "organized crime," for example, is a convenient label, but it imports significant Newtonian assumptions. Organization (see also below under "organized complexity") suggests a form of Weberian bureaucracy: charts with boxes and connector lines between them, and somebody at the top of it. This does not reflect the highly dynamic, adaptive nature of how localized agents work, and how they collectively constitute a Coltan-mining and exporting "organization" of some kind – without even knowing that they do. And then, proclaiming that such an organization (a crime

syndicate) has been dealt with by "chopping off its head" reduces the problem to what it isn't. Chopping off the head of a complex system doesn't work; it is even logically impossible. After all, its executive, its intelligence, its central nervous system, is distributed throughout the entire system. It is the system that is complex and smart and adaptive, not some omniscient governor that can be dealt with in isolation.

Of course, even in complex systems, not all components are equal. Some people are more involved in planning than others. And some people have a larger span of control or influence over other components than others. But that still doesn't mean that those people understand or control the whole system. Or that by taking them out of it, the system will collapse. This problem dogs the extrajudicial killings of terrorist leaders by remotely-operated drones too. Stopping terrorism by taking them out is really intractable, as others always step in to take their place, making up for their lack of experience with reinvigorated zeal. Announcing that one has interfered with "the machinery" of the criminal or terrorist organization and its activities also falsely suggests that componential solutions work. Taking one component out of a machine usually means it can no longer function. Taking one component out of a complex system does not mean it ceases to function. It adapts around the loss, making up for it, taking over functions and redistributing them. Importing Newtonian assumptions about the hunt for a culprit part by talking about the system in these terms, dramatically oversimplifies and mischaracterizes how these systems work; what makes them work, what makes them resilient.

Jim Nyce, an anthropologist colleague, and I have made this argument for the threat of improvised explosive devices (IED's) as well.[36] The U.S. Department of Defense (DoD) estimates that IED's are responsible for almost 50 percent of the casualties, mortal and injured, in Iraq to date and nearly 30 percent of the casualties in Afghanistan since the start of combat operations there.

Whenever we think about organizations and how to improve them, or try to understand how institutions in other cultures operate, we fall back on Weber, often without realizing it. This is not just because Weber's model informs much of our folk sociology; it is also because we take for granted that this model, based on what we take to be logic, rationality, and science, is the most effective and efficient organizational form. The taken-for-granted assumption becomes that if an adversary is effective in delivering IED's, in achieving its goals, then they have to have something like an efficient Weberian machine in operation. In other words, when we seek to understand IED financing, development, manufacturing, and distribution so as to stay ahead of the adversary's curve, we tend to look for some bureaucratic means of production, in terms of structure and hierarchy. Who is in charge where, of what, how are responsibilities managed, what are lines of supply, production? This tends to map directly on to questions of who to take out, which lines to disrupt.

But when something is not created bureaucratically, the means of organization and production, to say nothing of potential targets, tend to drop out of sight, and our means for determining appropriate action become less effective. All a Western analyst can do is register surprise that the original categories are not working, or that the adversary does not think like us, that he is less scrupulous, and just luckier.

The adversary's ability to adapt around our new technologies and better intelligence gathering does not hinge on a better bureaucracy on their part, even if some of their "coordination," may at first glance look very Weberian. What needs to be built into the equation is the possibility that these organizations work from and are informed by a set of principles fundamentally different from our own. This is often why these opponents, their intentions, and targets remain refractory, if not actually invisible, to us despite our best efforts. We do not have enough of the appropriate conceptual apparatus to see the groupings or entities that they do employ as organizations, let alone as threats.

Understanding how work gets done in the Middle East makes it clear that people there have refined the art of bricolage, of exploiting Western culture and its products, picking pieces from here and there, and then reassembling those in, for us, entirely incongruous, contradictory fashions. It is through this process of creative inspection, disassembly, and reassembly of Western technology and science that Iranian munitions become shaped charges, and garage door openers and cell phones become triggers for bombs. This also explains how our own media and analytic models, like network theory, have been turned against us. All this occurs through a set of kin relations, social institutions, and individual loyalties that defy any kind of Weberian description.

The West, however, consistent with the enlightenment and Newtonian science, tends to default to the individual. For this reason, terrorist acts are linked to something like demonic agency, as the ad hoc actions on the part of individuals that seemingly defy reason and morality. The IED threat, more than anything however, challenges the usefulness of applying such Newtonian–Cartesian concepts.

The example reveals another important point and that is that, in complex systems, history matters. Complex systems themselves have a history. They not only evolve over time, but their past is co-responsible for their present behavior. The IED threat faced by the various Western military forces in Iraq and Afghanistan is an example. U.S. funding and support for the Afghan Mujahideen (freedom fighters) during the 1979–1989 war against the Soviet Union there built the ideological and military basis for what they themselves have to battle against today. Known for their fierceness, and their stringent, extreme version of Islam, the Afghan Mujahideen attracted followers from a variety of countries who were interested in waging Jihad. Among those drawn to Afghanistan was a wealthy

ambitious and pious Saudi named Osama bin Laden. He built what would later become al Qaeda, the militant Islamic fundamentalist group whose goal is to establish a pan-Islamic caliphate by expelling Westerners and non-Muslims from Islamic countries. Similarly, the accident of Alaska 261 cannot be understood without going back to the 1960s and how, from then on, the certification and rationalities surrounding the jackscrew and acme nut evolved. Any study of a complex system, then, has to be diachronic, that is, concerned with how it has developed over time. Its current behavior, after all, is path-dependent. Synchronic descriptions (that is, only a snapshot in time) will ignore a very important aspect of complexity.

All of these features of complexity mean that a complex system is never entirely describable. No set of equations can be drawn up that captures the entire functioning of the system. This has to do with the computational intractability and the instability of a complex system, and with its openness to a changing environment. The brain, social systems, "organized" crime, ecological systems – these can only be "understood" (though not in the Newtonian sense of fully described) as a system. This is precisely where reductionist methods fail. We have no direct access to the complexity. Our knowledge of complex systems is therefore of necessity limited.

COMPLEX VERSUS COMPLICATED

Complex is not the same as complicated. A complicated system can have a huge number of parts and interactions between parts, but it is, in principle, exhaustively describable. We can, again in principle, develop all the mathematics to capture all the possible states of the system. This is true of the Boeing 737 I fly, for example. Well almost – we will soon see why this is not completely so. It is a system with hundreds of thousands of parts that all interact in many ways. But these parts, and their interactions, can be described, modeled. This, of course, is not the same as saying that they *have* been described exhaustively. Or that any one human or computer program or simulator or flight manual can represent a model of everything that all parts of a 737 will do at all times in all environments and still enable people to understand it in its entirety. This probably exceeds the practically available computational capacity of most systems, including all pilots who fly the plane.

Complicated systems often (if not always) do rely on an external designer, or group or company of designers. The designers may not beforehand know how all their parts are going to work together (this is why there are lengthy processes of flight testing and certification), but in due time, with ample resources, in the limit, it is possible to draw up all the equations for how the entire system works, always. Reductionism, then, is a useful strategy to understand at least large parts of complicated systems. We can break them down and see how parts function or malfunction and in turn contribute to the functioning or malfunctioning of super-ordinate parts or systems.

Or can we? Two similar and hugely vexing accidents with 737s have happened in which it has been impossible to conclusively nail down the broken part (though a so-called Power Control Unit (or PCU) that helped govern the pressure applied on the rudder in the aircraft's tail, was always a suspect and indeed later replaced over initial manufacturer objections). These cases, where the aircraft seemed to suddenly flip on its back and then tumble straight down into the ground from low height, have been impossible even to meaningfully characterize at a system level. Was it the wind? Was it a rare, virtually unpredictable and very subtle interaction between autopilot rudder actuation, pilot action, and the wind? Was it the pilot pushing right rudder when left rudder was called for, or was it the pilot pushing left rudder but the airplane going right because of a PCU reversal problem?[37] The pulverized remains of the two aircraft yielded no answers, and even if the remains hadn't been pulverized, they would have likely yielded that answer in a Newtonian universe only. A universe which the scenarios of these accidents seemingly defied.

Perhaps complicated systems are not just complicated, but increasingly also complex. One wild card in this (growing in importance all the time) is software. Few parts or sub-systems of an airplane today (including very much the human pilots) do not somehow interact with software. This can mean that a system is no longer just complicated, but that it becomes complex. Of course, there is a designer behind software, though this is becoming recursive: software is increasingly written by other software. But that there was a designer at one point in time may mean very little once the system is fielded in an open environment.

Recall the example of TCAS or traffic collision avoidance systems from above, which are now a mandatory part of all large airplanes. With more than a million lines of code, they warn aurally and visually of other traffic and tell the pilot to maneuver up or down to avoid it. Telling the pilot to go sideways is outside the pale of the system. This would take considerably more computational capacity, millions more lines of code, and create considerably more uncertainty. With so much code already, and a literally infinite number of vector-to-vector situations, there really is no way to exhaustively model the interaction between all the parts of the code (for example, objects, lines) and all possible situations in which it may be triggered into action. Reductionism begins to lose traction here.

The apparently simple jackscrew-acme nut combination of Alaska 261 (Chapter 1) showed this nicely. Even for a very straightforward system, once the design was done, bets were off. Thread wear on acme nuts should, in principle, have been predictable with sufficiently accurate models of the amalgam of materials used to build the screw and the nut, of predictions of particle concentrations in the air (for example, salt, sand) in which airplanes were going to fly, and of the behavior of the lubricant over time, and over the likely temperature, air density and moisture ranges (and rates of change) along which all of this was going to rub together. But even if this was possible, or done, what could not be predicted were other parts of the environment in which it was going to operate. The system turned out to be quite open for influences way beyond the control or even imagination of those who designed it. Consider deregulation, developments in maintenance

guidance and a philosophy towards more paper-based system-level oversight, evolutions of task distributions across different checks, operational ambiguities about on-airplane or off-airplane wear testing and the parts used to do it with, and how each individual operator of the airplane was going to interpret all of this and get it approved. None of this would have been easy to predict. It would have been impossible to model how it all might interact to produce legal lubrication intervals far beyond a 1960s engineer's worst nightmare.

Fielding an engineered system (like a PCU or jackscrew), that is, putting it to use in actual practical settings, opens a previously closed system. By opening it to the world, it will very likely start to display properties that are more akin to a complex, rather than complicated system. This, again, is the story of Alaska 261. Influences from regulatory relief, and the resulting scarcity and competition, and how this may have accelerated developments of system-level maintenance guidance and inspection recommendations start leaking into the system operating the jackscrew, and it responds in return. This once more shows the importance of diachronically understanding complexity: tracing the development of a system over time may reveal a lot about why it is behaving as it is synchronically, at any one point in time.

Complicated systems like a Boeing 737, or any other large aircraft (or, for that matter, almost any other complicated piece of technology) can also become complex when they are taken into cultural territories that lie at a far distance from the design and engineering assumptions that went into making it. Such implementation in unfamiliar or misunderstood social settings leads to unforeseen and uncontrollable explosions of relationships and interdependencies. Two seats in a cockpit and the artifacts (displays, switches, checklists) and procedures that connect them, are no longer animated by the manufacturer-envisioned, technical cooperation necessary to drive a piece of machinery alone. Putting people from cultures other than the manufacturer's in there breeds new relationships and interdependencies between people, procedures and technologies. It leads, in other words, to complexity.

Western assumptions about language competency, operational hierarchy and interpersonal communication and collaboration are woven into most technology produced in the Western world, even if implicitly. This has led to concerns, for example, that airliner cockpits and the way in which they are supposed to be operated collaboratively and relatively democratically today, are insufficiently sensitive to the cultural predispositions of some of the countries in which they end up being flown. Airliners that are designed and built in a culture that has no politeness differentiation in interpersonal address are flown in cultures that have such differentiation, like the two levels of many European languages, or like the six politeness levels of interpersonal address in some Asian languages. Western technology that is operated by multiple people also typically assumes a transmitter-orientation in interpersonal communication. That means that any lack of clarity (about content or urgency of the message) is mainly the responsibility of the sender, not the receiver. It is the sender who has to escalate, the sender who has to repeat, the sender who has to clarify. In some cultures, such

transmitter initiative and insistence will be seen as violating and deeply impolite. And, independent of culture, the collaborative operation of machinery in time-critical situations assumes the use of unmitigated language – language that doesn't hedge, qualify, weaken or soften, but that delivers unvarnished and unambiguous messages. People generally have a really hard time doing that, because they are not the machine they are operating. They are people, with feelings and social expectations about relationships. Only constant training and indoctrination, as well as positive reinforcement, can help build confidence and acceptance for the use of unmitigated language.

There is no way that a technological environment can accommodate such nuance and diversity without consequences. Assumptions about purely technical cross-checking and challenging, baked into features of displays, procedures and checklists, fall flat when subjected to a culture that takes a different view of command, power, politeness and seniority.[38] The result is that original designs, when opened up to such influences, get adapted in ways that are hard to foresee by those who built it, by those who put the original parts together. The ability to describe, beforehand, the possible states that the system may get into, is challenged severely by the ways in which it will be put to use in different parts of the world.

The importance of the diachronic study of complex systems comes out of a recent investigation into an accident with a Boeing 737 as well.[39] During approach, this aircraft suffered problems with its left radar altimeter. The radar altimeter is used in part to govern the autothrottles. If pilots have the intention to make an automatic landing, the radar altimeter will tell the autothrottles to start retarding 27 feet above the ground, and stay in retard. In this case, the right pilot (co-pilot) was flying the aircraft, and the right autopilot was switched on. The co-pilot might have thought that, as a result of this, the right radar altimeter was telling the autothrottles what to do, and that as such, he would suffer no consequences from the problems that the left radar altimeter was having. In all the books and training available to pilots on the 737, this is exactly the suggestion that is being made.

But it is not true. The right radar altimeter never talks to the autothrottle. It is always the left radar altimeter that talks to the autothrottle, and tells it when to retard or do other things. This is independent of which pilot is flying, and independent of which autopilot is engaged. In the case of the accident, the left radar altimeter erroneously told the aircraft that it had already landed, even while it was still at 2,000 feet and a few miles short of the runway. The result was a loss of airspeed and eventual crash.

The system design (originally certified in the 1960s) prioritized data integrity for the left pilot (commander), but sacrificed it for the first officer (second-in-command), consistent with societal and airline hierarchical arrangements at the time. In today's much flatter cockpits, however, it leaves a first officer with a seemingly autonomous system that is actually driven by "borrowed" data and auto-throttle inputs from the left pilot, which can be corrupted without his knowledge. As airline training footprints shrank over the next decades, this information disappeared from pilot education and manuals, leaving an unknown automation booby trap long into the twenty-first century. A once complicated system became complex when exposed to the diverse, changing world in which it was to operate and survive.

COMPLEXITY AND DRIFT

The paradox is of course this. Complexity can guarantee resilience. Because they consist of complex webs of relationships, and because a lot of control is distributed rather than centralized, complex systems can adapt to a changing world. They can survive in the world thanks to this ability to adapt. So how can it be that complexity contributes to failure, to accidents? What is the relationship between complexity and drift? Complexity opens up a way for a particular kind of brittleness. Their openness means unpredictable behavior. Releasing a jackscrew into a world of competition and scarcity, and bets about maintenance intervals are off. Release a complicated engineered system into a world of cultural nuance, diversity and societal maturation, and original design assumptions get adapted, forgotten, muffled. But that is just the start. Complexity and systems theory gives us a language, and some metaphors, to characterize what may happen during the journey into failure, during the trajectory toward an accident.

The path-dependence of complex systems (or of a transformative journey from complicated to complex system) is a great starting point. Drift into failure can never be seen synchronically; systems have to be studied diachronically to have any hope of being able to discern where they might be heading and why. The non-linearity of relationships between components offers opportunities for dampening and modulating risky influences (for example, an increase in lubrication intervals might be accompanied by better endplay measurement devices and checking, which could increase the reliability of that part non-linearly even with less lubrication). But the non-linearity can also turn small events into large ones. The same small event, of missing one lubrication opportunity, would have had small consequences in the 1960s, and huge ones in the late 1990s. Making it increasingly difficult for smugglers to get their drugs across the Caribbean by numerous small steps that improved monitoring and interdiction led to a large event: a wholesale deflection of smuggling routes through West Africa.

In the mechanistic worldview, it is enough to understand the functioning or breaking of parts to explain the behavior of the system as a whole. In complexity and systems thinking, where nothing really functions in an unbroken or strictly linear fashion, it is not. Recall, from the second chapter, the outlines of drift into failure. Here is how they interact with what complexity theory has to say:

- Resource scarcity and competition, which leads to a chronic need to balance cost pressures with safety. In a complex system, this means

that the thousands smaller and larger decisions and trade-offs that get made throughout the system each day can generate a joint preference without central coordination, and without apparent local consequences: production and efficiency get served in people's local goal pursuits while safety gets sacrificed – but not visibly so;

- Decrementalism, where constant organizational and operational adaptation around goal conflicts and uncertainty produces small, step-wise normalization where each next decrement is only a small deviation from the previously accepted norm, and continued operational success is relied upon as a guarantee of future safety;
- Sensitive dependence on initial conditions. Because of the lack of a central designer or any part that knows the entire complex system, conditions can be changed in one of its corners for a very good reason and without any apparent implications: it's simply no big deal. This may, however, generate reverberations through the interconnected webs of relationships; it can get amplified or suppressed as it modulates through the system;
- Unruly technology, which introduces and sustains uncertainties about how and when things may fail. Complexity can be a property of the technology-in-context. Even though parts or sub-systems can be modeled exhaustively in isolation (and therefore remain merely complicated), their operation with each other in a dynamic environment generates the unforeseeabilities and uncertainties of complexity;
- Contribution of the entire protective structure (the organization itself, but also the regulator, legislation, and other forms of oversight) that is set up and maintained to ensure safety (at least in principle: some regulators would stress that all they do is ensure regulatory compliance). Protective structures themselves can consist of complex webs of players and interactions, and are exposed to an environment that influences it with societal expectations, resource constraints, and goal interactions. This affects how it condones, regulates and helps rationalize or even legalizes definitions of "acceptable" system performance.

The concern behind complexity and drifting into failure is how a large number of things and processes interact, and generate organizational trajectories when exposed to different influences. Resource scarcity and goal oppositions form one such pervasive influence. They express themselves in thousands of smaller and larger trade-offs, sacrifices, budgetary decisions – some very obvious, others hardly noticed. The ripple effects of such decisions and trade-offs are sometimes easy to foresee, but often opaque and resistant to anything resembling deterministic prediction. Decrementalism shows up in all kinds of subtle ways as people in the organization adapt, rationalize and normalize their views, assessments and decisions.

The contribution of the protective structure (for example, a safety regulator) to such adaptation and normalization, as well as exposure to its own resource

constraints and goal interactions, is another influence on this. Such influences ebb and flow to different parts of the operational organization or even originate there, and are negotiated, dealt with, ignored or integrated. As Heisenberg put it, "The world thus appears as a complex tissue of events, in which connections of different kinds alternate or overlap or combine and thereby determine the texture of the whole."[40] Drift can be one property of all of these countless relationships and subtle interactions that emerges at the system level; as one aspect of Heisenberg's visible texture. Let's lift a few concepts out of that texture – emergence, phase shifts and the edge of chaos – to see how they can apply to drift into failure.

WHAT IS EMERGENCE?

As always in the tug between Newtonian and holistic systems of thought, the basic tension that animated this development was that between the parts and the whole. What was the relationship between parts and whole? Could the parts explain the whole, or did they fall short in accounting for the behavior of the whole? What could be considered as "parts" in the first place? In twentieth-century science, the holistic perspective, which rejects the idea that the whole can be understood as the sum of parts, has become known as systemic, and the way of reasoning it encouraged was called systems thinking.[41] Again, the basic reflex in systems thinking is to go up and out, not down and in. It is an understanding of relationships, not parts, that marks systems thinking.

The question of accidents in complex systems is most certainly a question about the relationship between parts and wholes. The Newtonian answer to the question of accident causation (and, by implication, the question of the relationship between parts and wholes) has always been simple: the parts fully account for the behavior of the whole. Hence our search for the broken part. Hence our satisfaction when we have found the "human error" by any other name or human, that can be held responsible for the accident. In complexity and systems thinking, the relationship between parts and wholes is – you guessed it – a bit more complex. The most common way to describe that relationship is by using the idea of emergence. A whole has emergent properties if what it produces cannot be explained through the properties of the parts that make up the whole. Here is an example:

What are the emergent properties of kitchen salt?

- It has a salty taste (no, really?)
- It is edible (well, in reasonable quantities)
- It forms crystals

So what are the components that make up kitchen salt? You might say: salt crystals. Well, no. That's the whole point. The crystals are an emergent property, as are the taste and their edibility. The "components" that kitchen salt is made up of are Na and Cl, or Sodium (Natrium) and Chlorine. Sodium is a poisonous

gas. Natrium is a violently reactive soft metal (Yes, you eat this. A lot). So kitchen salt displays properties that are completely different from the properties of its (chemical) component parts. And what's more, it doesn't display the properties of its component parts at all. You could argue that it only gets dangerous at quantities like those consumed by Morgan Spurlock in his 2004 documentary "Supersize Me," not because it's a poisonous gas or a violently reactive metal but because it does things with blood pressure and kidney function and such. You remember Morgan puking out of the car window after eating yet another Mackey D's breakfast, the camera gleefully tracking the yellow glop as it spluttered onto the pavement? Okay, that's not what I'm talking about.

We used to say that the whole is more than the sum of its parts. Today, we would say that the whole has emergent properties. But it means the same thing. Emergence comes from the Latin meaning of "bringing to light." The nineteenth century English science philosopher George Henry Lewes was probably the first who made a distinction between resultant and emergent phenomena.[42] Resultant phenomena can be predicted on the basis of the constituent parts. Emergent phenomena cannot. The heat of apple sauce is a resultant phenomenon: the faster the constituent molecules are moving around, the hotter the apple sauce. The taste of an apple, however, is emergent. Taste does not reside in each individual cell, it cannot be predicted on the basis of the cells that make up the apple. Wetness does not reside in individual H_2O molecules. They are not wet, they don't flow. Wetness is an emergent property, something that can only be observed and experienced at system level. Wetness cannot be reduced to (or found in) the molecular components that make up water.

Theorizing around emergence took off in earnest in the late sixties with the study of slime mold in New York City.[43] Why slime mold? Because as a collective, or a whole, it does some amazing stuff that the component parts couldn't dream of pulling off. Slime mold is like ordinary fungi, or mildew. It's like the stuff on the inside your basement walls. Mushrooms are fungi too, a kind of mold in another form. Not long ago, a Japanese scientist succeeded in getting slime mold to find its way through a maze, and have it find food. One clump of slime mold even stretched itself thin to gain access to two sources of food in different places in the maze simultaneously. In a later experiment, it was possible to have the slime mold stretch itself out in a mimic of the Tokyo subway map.

Sending organisms through mazes in the pursuit of food is something that scientists used to reserve for mice and rats and other rodents. What puts rodents apart from slime mold is the same thing that separates a rule-based system from a complex system: they have a central nervous system – a location for their supposed intelligence, a mechanism for them to learn about and map their world, to build some kind of memory, and to then act in accordance with it (for which rodents have four paws, to walk them through the maze). Slime mold doesn't have a central nervous system. It also doesn't have paws. In fact, slime mold doesn't have much of anything, other than basic biological material that can reproduce and that needs food to survive. Slime mold has no central organizing entity, like

a brain, that connects eyes on the one side with paws on the other. Nothing or nobody is in charge, nothing or nobody directs eyes where to look, paws where to walk. There is no specialization of parts that take on the various tasks. Every cell or protoplasm of slime mold is basically the same. One big egalitarian democracy of jelly, or slime.

In other words, slime mold is made up of very simple components, but it does very complex stuff, in interaction with its changing environment. Its emergent behavior would seem intelligent. The slime mold as a collective, amazingly enough, pursues goals. Like getting food. And collectively, it adjusts its behavior on the basis of what it finds in the environment. It does this really well. Slime mold actually walks. Well, in a sense. It moves across the forest floor, for example, in pursuit of food. And it gets even more uncanny: if there's a lot of food, it will actually break up into a flock, so as to capitalize on the various food sources. If there's not a lot of food, slime mold will pull together into one larger clump and stick together to ride out the lean times.

ORGANIZED COMPLEXITY

Such emergent behavior relies on a kind of distributed, bottom-up intelligence, not on a unified, top-down intelligence. Rather than running one single, smart program (as the rat in the maze supposedly would), there is a whole swarm of simple mini-programs, each running inside the cellular or a-cellular components that make up the mold. These programs really are sets of very simple rules. Anthills get built like this. In fact, much of ant colony life is organized around a few sets of simple rules. To find the shortest route to food, for example, (1) wander around randomly until finding food, (2) grab it and follow your pheromone trail back to the nest. Other ants will follow the same trail out because it is strong and fresh. If they don't find food at the end of it, or the trail begins to evaporate, the ants will go back to the first rule. Anthills, which are very sophisticated emergent structures, get built in the same way, though the rules are slightly different. Even the ant garbage dump and the location of their cemetery (maximal distance from the anthill) emerge from the millions of interactive applications of a few simple rules. Even the biblical book Proverbs from around 950 BCE, containing maxims attributed mainly to Solomon, refers to the ants. They have no king, no central authority, Proverbs says. Yet they are industrious and build great things.

What emerges has been called "organized complexity." It is complex because there are a large number of components, and, as a result, a dense throng of mini-programs running and interacting, cross-influencing each other. Such complex interaction and constant cross-adaptation is hard to model and understand using traditional deterministic, or Newtonian logic, because so much of it is non-linear and non-deterministic. But what it produces is not disorganized. Rather, it is organized: it creates higher-order behavior (like wandering in pursuit of food, or a huge ant stack that heats and cools and stores and incubates) as an amazing emergent product of the complex interactions between a multitude of simpler entities. Of course, the word "organized" is perhaps a bit problematic in this

context, as it suggests some kind of entity that does the organizing or controls it. Which is not the case. The organization is purely emergent, purely bottom-up, the result of local interactions and the cumulative effect of simple rules.

Behavior of the brain has also been seen as emergent. In fact, it is both ironic and inspiring that we no longer think of the mammalian brain, the central nervous system, as a centralized smart program either. Rather, brains are made up of a huge mass of much simpler components with simpler programs (neurons or nerve cells) that, in their action-reaction or even ON–OFF behavior cannot begin to show the sort of complex phenomena that a collective brain is capable of exhibiting. Indeed, one example of emergence that is often used is consciousness. The cells that make up the brain are not conscious (at least not as far as we know. And, for that matter, how could we ever know?). But their collective interaction produces something that we experience as consciousness. Of course, much of human factors, thanks to its reliance on information processing psychology, still treats the brain as a central executive function: as an entity that produces and directs intelligent behavior. But the components that make up the brain are hardly intelligent. Its intelligence cannot be reduced to its constituent components. Intelligence itself is an emergent property.

ACCIDENTS AS EMERGENT PROPERTIES

Since ideas about systemic accident models were first published and popularized, system safety has been characterized as an emergent property, something that cannot be predicted on the basis of the components that make up the system.[44] Accidents have also been characterized as emergent properties of complex systems.[45] They cannot be predicted on the basis of the constituent parts; rather, they are one emergent result of the constituent components doing their normal work.

But, you may object, isn't there a relationship between some components (people, or technical parts) not doing their work, and having an accident? Indeed, if you believe this, it would affirm the idea that accidents are resultant phenomena, not emergent. After all, the fact that an accident happens can be traced back, or reduced, to a component not doing its job. The accidents-as-resultant-phenomena idea is alive and seems intuitive, consistent with common sense. In Alaska 261, for example, the component that didn't do its job was the jackscrew in the tail of the MD-80 airplane. It broke and that's why there was an accident. In Tenerife, one of the components that didn't do his job was the co-pilot, who didn't speak up against a stubborn captain.[46] That human component didn't work, and that's why there was an accident. Such characterizations are quite popular and, in many circles, still hard to call into question.

Those who are involved with occupational health and safety issues may want to believe it too. For example, isn't there a relationship between the number of occupational accidents (people burning themselves, falling off stairs, not securing loads, and so on) and having an organizational accident? Isn't it true that having a lot of occupational accidents points to something like a weak safety culture, which

ultimately could help produce larger system accidents as well? Not necessarily, because it depends on how you describe the occupational accidents. If accidents are emergent properties, then the accident-proneness of the organization cannot be reduced to the accident-proneness of the people who make up the organization (again, if that is the model you want to use for explaining workplace accidents). In other words, you don't need a large number of accident-prone people in order to suffer an organizational accident. The accident-proneness of individual employees fails to predict or explain system-level accidents. You can suffer an organizational accident in an organization where people themselves have no little accidents or incidents, in which everything looks normal, and everybody is abiding by their rules.

The Alaska Airlines 261 accident, the topic of Chapter 2, was, as far as is known, not preceded by or correlated with a noticeable number of relevant occupational accidents. In fact, what is remarkable about this accident is that everybody was pretty much following the rules. Their local rules. The airline and its maintenance arm was abiding by recommendations from the manufacturer and the rules set by the regulator, the regulator was approving successive maintenance intervals on the basis of the evidence that was presented and deemed appropriate and sufficient at the time. The people doing the maintenance work followed the rules and procedures (however underspecified they might have been, but that was normal too) given to them for the execution of the lubrication task. Whether more or fewer maintenance workers would have suffered occupational incidents doesn't matter. It has no predictive value, it can carry no explanatory load relative to the eventual accident that emerged precisely because everybody was following their local rules. Just like the ants building their hill.

In Alaska 261, whether fewer or more technicians' hands got stuck and injured in the lubrication of the jack screw may bear no relationship to the eventual organizational accident. That is emergence. The behavior of the whole cannot be explained by, and is not mirrored in, the behavior of its constituent components. Instead, the behavior of the whole is the result – the emergent, cumulative result – of all these local components following their local rules and interact with each other in innumerous ways as they do so. Of course, unlike ants and slime mold, people across various organizations are not all the same. They are, in fact, quite different. But, like ants, all of them still respond to, and help shape, their local environment and follow or help shape the rules (written or otherwise, formal or informal) that apply there. So the properties of emergence would seem to hold. Behavior that is locally rational, that responds to local conditions and makes sense given the various rules that govern it locally, can add up to profoundly irrational behavior at the system level. Like an accident that takes scores, or hundreds, of lives – in a system where everybody locally works precisely to prevent that from happening.

A recent example from a major energy company operating in Louisiana near the Gulf of Mexico illustrates the problematic connection between the safety of individual components and system safety.[47] An employee of this company

injured himself when his swivel chair collapsed. The injury was severe enough for him to have to take a day off work. People from Occupational Health and Safety (OSHA) came in, inspected the situation, and issued a citation to the company for "not properly instructing employees how to sit in their chairs."

What was the company to do? Under pressure to show managerial action in response to the OSHA finding and citation, their health and safety department sent out a company-wide PowerPoint presentation, which demonstrated how to properly and safely sit in a swivel chair. In summary, it told employees to inspect their chair at least once every month, and to remove defective chairs from the workplace. Also, employees were to use the chair only for the purpose for which it was designed, and never stand in a chair to retrieve an object out of reach. The company president himself got personally involved with enforcing the OSHA sitting recommendations, citing employees who violated the new policy. His rationale was that "Permitting the smallest exceptions to our health, safety and environmental program is unacceptable and results in catastrophes such as the BP disaster."

Chairs tipping over because people don't sit in them in ways that they were designed for may have a relationship with a disaster the size of Deepwater Horizon, but it may not. Complexity doesn't allow us to draw such straight lines from the behavior of individual components to large events at system level. In fact, complexity understandings tipping to mean something different.

PHASE SHIFTS AND TIPPING POINTS

The emergent behavior of a complex system can be very different, from one situation or moment to the next, than the collection of its parts would suggest. One dynamic that is responsible for this is called the tipping point, the phase shift, or phase transition. A phase shift means that a bit more (or less) of the same leads to something *very* different. Of course, if the relationship between parts and system were as straightforward as Newton proposed, then this could never happen. A bit more behavior by the parts would lead to the same little bit more behavior by the system as a whole. But in complexity, that is not the case. A tiny little bit more (or a tiny little bit less) of the same in terms of the parts, can lead to something *completely* different for the whole system. It can lead to something qualitatively different.

The original idea for phase transitions comes from physics (or, more specifically: thermodynamics). It describes the shift of a thermodynamic system from one phase to another. As solid material is heated, it will transition to a liquid at some temperature, shifting phases, and changing system level properties. Heat it more, and it will change to a gas. Heat it eve more, and it may, in rare cases, become plasma. The behavior of the parts (molecules moving among each other) is not at all that different on either side of a phase transition point, yet the system properties undergo an abrupt change. For example, the volume taken up by steam

is vastly different from the volume that boiling water needs (yet the difference in temperature between these two only needs to be infinitesimal).

The idea gained currency in sociology in the 1960s. The term tipping point, or angle of repose, was introduced to describe how a previously rare social phenomenon could become rapidly and dramatically more common. Morton Grodzins was studying racially diverse neighborhoods in the U.S.A. in the 1960s, when he discovered that most of the white families remained in the neighborhood as long as the comparative number of black families remained small. Very small. At a certain point, however, when "one too many" black families arrived, the remaining white families would move out en masse, in a process that became known as "white flight." The phrase tipping point of course was borrowed from physics itself, analogous to the adding of a small weight to a balanced object that could cause it to suddenly and completely (and irreversibly) topple over.

On September 17, 2007, in Nisour Square, Baghdad, 17 people were killed and 24 injured when security teams from the private firm Blackwater U.S.A. unleashed a barrage of machine-gun fire. A bomb had gone off nearby just before. The first victim was Ahmed Haithem Ahmed, who was driving his mother, Mohassin, to pick up his father from the hospital where he worked as a pathologist. A Blackwater bullet tore through his head, but the car kept rolling toward the Blackwater convoy, and not much later 17 people were dead in a hail of gun fire, including Iraqis trying to escape to safety. No shot that could have provoked the Blackwater response had been heard.[48]

The event blew the extent of the privatization of warfare in Iraq into the open, with some in the U.S. congress arguing that if war was a profit motive, then peace could be hard to achieve. Blackwater lost its lucrative State Department contract to provide diplomatic security for the U.S. embassy in Baghdad after the incident.

Yet it would take two more years for the subtle drift towards the deadly incident to become apparent.

Blackwater U.S.A., a private security firm, was originally contracted to provide security for State Department and CIA employees after the September 11 attacks in 2001. In the spring of 2002, Erik Prince, the founder of Blackwater, offered to help the CIA guard its makeshift Afghan station in the Ariana Hotel in Kabul. Not long after signing that contract, dozens of Blackwater personnel, many of them former Navy Seals and Army Delta Force ex-soldiers, were sent out into the surrounding streets to provide perimeter security for the CIA station. From there, Blackwater operatives began accompanying CIA case officers on missions beyond the perimeter.

A similar progression happened in Iraq. Blackwater was first hired to provide static security for the CIA Baghdad station. Also, Blackwater employees were hired to provide personal security for CIA officers whenever they traveled in

either Iraq or Afghanistan. This meant that Blackwater personnel began to accompany CIA officers even on offensive operations, sometimes launched in conjunction with Delta Force or Navy Seals teams (that is, former colleagues). It will never be possible to find out who fired the first offensive, not security-defensive, shot from a Blackwater gun in these operations, but lines soon began blurring.

Blackwater employees started to play central roles in so-called "snatch-and-grab" operations, intended to capture or kill militants in Iraq and Afghanistan, particularly during the 2004–2006 height of the Iraqi insurgency. Blackwater exercised a strong influence on such clandestine CIA operations, under the banner of being able to decide what the safest ways were to conduct such missions. They filled all roles, from "drivers to gunslingers."[49]

The House Oversight and Government Reform committee found that Blackwater had been involved in at least 195 shootings over the previous two years, many of which involved cover-ups of fatal shootings by its staff.

The incident on the 16th of September, 2007, was perhaps not a large departure from where things had been drifting. The Blackwater convoy in question was in the square to control traffic for a second convoy that was approaching from the south. The second convoy was bringing diplomats who had been evacuated from a meeting after a bomb went off near the compound where the meeting was taking place. That convoy had not arrived at the square by the time the shooting started.

As the gunfire continued, at least one of the Blackwater guards began screaming, "No! No! No!" and gestured to his colleagues to stop shooting, according to an Iraqi lawyer who was stuck in traffic and was shot in the back as he tried to flee.[50]

In the blur between CIA, military and contractor roles that grew during the worst part of the 2004–2006 Iraqi insurgency, not much would have distinguished one "snatch-and-grab" raid from the next. Blackwater guards would have fired shots in defense of the CIA and military, consistent with their assignment. Until a shot was fired that was not in defense, or not entirely in defense, or not at all in defense, but rather a contribution to the offensive raid. It is not that hard to imagine. You know the guys by name, you remember them from the time in the Seals or in Delta, you go on raids side-by-side for weeks, you see them lift a gun and fire in the same direction that everybody else is shooting in, you even get a shout perhaps, or a taunt, or a question. You shoot too. What's one offensive shot between friends? The next raid, you may even be expected to help out that way. One more shot, a bit more of the same, and the system began to display vastly different properties. Blackwater started to go in ahead of the others, started to help plan operations and take the lead, started to play an offensive role in the

missions. Did the killing, even when unprovoked. A first offensive shot may have been a tipping point.

Despite original design requirements that the External Tank not shed debris, and the corresponding design requirement that the Shuttle not receive debris hits exceeding a trivial amount of force, debris impacted the Shuttle on each flight. Debris strikes were normal, in other words. Just like a lot of other technical problems – NASA engineers were, and always had been, working in an environment where technical problems proliferated. Flying with flaws was the norm. Over the course of 113 missions, foam-shedding and other debris impacts came to be regarded less as a hazard to the vehicle and crew. With each successful landing, it appears that NASA engineers and managers increasingly regarded the foam-shedding as inevitable, and as either unlikely to jeopardize safety or simply an acceptable risk.

The distinction between foam loss and debris events also appears to have become blurred. NASA and contractor personnel came to view foam strikes not as a safety of flight issue, but rather a maintenance, or "turnaround" issue. In Flight Readiness Review documentation, Mission Management Team minutes, In-Flight Anomaly disposition reports, and elsewhere, what was originally considered a serious threat to the *Orbiter* came to be treated as "in-family," a reportable problem that was within the known experience base, was believed to be understood, and was not regarded as a safety-of-flight issue. The reason why this problem was in the known experience base was that its result, heat tile damage, had occurred on previous occasions (in fact, was very normal) and occurred because of a variety of reasons. Here was just one more.

The foam-loss issue was considered insignificant enough that Flight Readiness Review documents included no discussion about it. There was no paper trail of concerns about foam debris tile damage that preceded the accident. This even fit the rules. According to Program rules, this discussion was not a requirement because the STS-112 incident was only identified as an "action," not an In-Flight Anomaly. Official definitions were assigned to each in-flight anomaly and ratified by the Flight Readiness Reviews. It limited the actions taken and the resources available for dealing with these problems.[51]

In the evaluation of damage caused by debris falling off the external tank prior to the 2003 Space Shuttle *Columbia* flight, you can see a similar phase shift. Under pressure to accommodate tight launch schedules and budget cuts (in part because of a diversion of funds to the international space station), it became more sensible to see certain problems as maintenance issues rather than flight safety risks. What was known as "debris events" now became "foam loss," a more innocuous label. Maintenance issues like foam loss could be cleared through a nominally simpler bureaucratic process, which allowed quicker turnarounds. In the enormous mass of assessments to be made between flights, foam debris

strikes became one more issue. Gradually converting this issue from safety to maintenance was not different from a lot of other risk assessments and decisions that NASA had to do as one Shuttle landed and the next was prepared for flight. It was quite normal. It may, however, have represented the kind of phase shift, or phase transition – one more decision, just like tens of thousands of other decisions, that produced fundamentally different system behavior in the end.

With the benefit of hindsight, of course, it is easy to point to the flaws in these logics and priorities, for example those that converted a flight safety problem into a maintenance problem. But what we really need to understand is how these conversions of language made sense to decision-makers at the time. After all, their objective cannot have been to burn up a Space Shuttle on re-entry. And the important question to ask ourselves is how organizations can be made aware early on that such shifts in language can have far-reaching consequences, even if those are hard to foresee. In complex systems, after all, it is very hard to foresee or predict the consequences of presumed causes. So it is not the consequences that we should be afraid of (we might not even foresee them or believe them if we could). Rather, we should be weary of renaming things that negotiate their perceived risk down from what it was before.

OPTIMIZED AT THE EDGE OF CHAOS

A common notion is that the functioning of complex systems is optimized at the edge of chaos. Originally a mathematical concept, the edge of chaos in socio-technical settings denotes a region of maximal complexity where systems are perched between order and disorder; between optimally organized complexity and chaos. Maximal functionality can be extracted from the system at this edge. This is where the system is tweaked to achieve its extreme diversity and variety, where complexity reaches its optimum. Here, the system can maximally, exhaustively and swiftly adapt and respond to changes in the environment (4WDs to speed boats to airplanes, day to night, Caribbean to Guinea). Indeed, this optimum, or maximum is determined very much in the relationship between complex system and environment.

In a sense, the edge of chaos is where the ecological arms race plays out in its most pure form – where competitors, or predator and prey are constantly trying to stay one step ahead of each other. The use of aerial surveillance will make trafficking over open ground less attractive, perhaps deflecting smugglers' routes into forests. The use of infrared or other dark-penetrating optics makes trafficking at night less protected, thus putting a greater premium on quiet daytime travel, when the sun is hot and high and policing might be less effective. The various actions and counteractions constantly affect each other to undermine what was previously a good strategy while at the same time forcing the creation of new strategies. Complex systems tend to settle at the edge where their responses are just good enough to stay ahead of the others, but where there is not such a huge cost to generating those responses that they will run out of business because of it. That is the edge of chaos.

There is a fundamental trade-off that interdependent agents in a complex system must make, a trade-off that lies behind much of the adaptation that the system can display, and also why such adaptation sometimes becomes less successful. This is the trade-off between exploitation versus exploration. Exploration is a necessary activity for survival inside a complex, adaptive, living system. It means searching for a solution that is optimal relative to the landscape of opportunities and constraints you now know. Flying cocaine via West Africa may be a great opportunity that can be the result of exploration if the Caribbean has gone solid with narcotics countermeasures, policing, interdiction and arrests. Exploration is that which generates new smuggling routes (for example, through Africa). But making those routes work takes investment, of time and money. Local politicians may need to be bought, local strongmen need to be found and patronized, the local geography and ecology of inlets, ports, airports, roads, forests, hiding places, mules and so forth, need to be mapped. So if there are still plenty of holes to wiggle cocaine through in the Caribbean, then West Africa may not be optimal. The Caribbean can still be exploited – until further.

Exploitation means taking advantage of what you already know, it means reaping the benefits of past searches.[52] If you stop exploring, however, and just keep exploiting, you might miss out on something much better, and your system may become brittle, or less adaptive. While you are exploiting Caribbean routes, competitors may have already set up shop in Guinea. By the time the Caribbean holes close, you are left with no alternatives other than declaring war on rival smugglers. At the edge of chaos, then, the system has reached an optimum point – not just in either exploration or exploitation, but in their complement. It gets enough return from exploiting that which was explored earlier, yet retains adaptive capacity by continued exploration of better options. In a complex system, however, it is difficult to say in advance what the returns of exploration are going to be. A new route may get discovered. But a key smuggler on a scouting mission may get caught, and give up his or her collaborators. Exploration can thus produce big events. This is because the optimum balance between exploration and exploitation puts a system near the edge of chaos.

The edge of chaos is what is called a critical state, or, in an explicit reference to thermodynamics, a dissipative state. While poised in a critical or dissipative state, a system is not in equilibrium and can be easily perturbed. In this state, the many components that make up the complex system never quite lock into place, yet hardly ever dissolve into an entire chaos. For example, by purchasing a faster Coast Guard aircraft to go after jet-bound smugglers across the Caribbean, there will be a response. The smugglers' jet may fly at night, or change routings more often. Perturb the system again, and there might be a completely different response, and one at a different scale – such as the opening of routes through West Africa. This is one very characteristic property at the edge of chaos: Small responses are common, big responses are rare. But big responses are possible, and what they are can be almost impossible to foresee.

Arriving at the edge of chaos is a logical endpoint for drift. At the edge of chaos, systems have tuned themselves to the point of maximum capability. Mixing

the base chemicals in a cup and smearing them across scratches and cracks and gouges in the foam covering of the Space Shuttle's external fuel tank was one such response to production pressures. Those actions got the tanks returned to service quickly by making evidence of any maintenance and possible flight safety problems go away. Production was served, and production had been called for (incentivized in various ways, even). The system had tweaked itself to maximum capacity in a constant, dancing interaction with the political, medial and technical environment in which it operated. It was in a critical state, which suddenly allowed a big response to the small, little actions that maintenance workers had applied hundreds of times. A large event happened on 16 January 2003: the separation of a chunk of iced-up foam so large that it managed to penetrate the heat shield on the wing's leading edge, which would later lead to the loss of *Columbia* while on its re-entry. Drift into failure, in these terms, is about optimizing the system until it is perched on that edge of chaos. There, in that critical state, big, devastating responses to small perturbations become possible. Large events are within the space of possibilities. Drift doesn't necessarily lead to failure. At least not until it does.

REFERENCES

1 Laszlo, E. (1996). *The systems view of the world: A holistic vision for our time.* Cresskill, NJ: Hampton Press.
2 Perrow, C. (1984). *Normal accidents: Living with high-risk technologies.* New York: Basic Books.
3 Perrow, C. (1984). Ibid.
4 Perrow, C. (1984). Ibid.
5 Seaways. (2010, July). London: The Nautical Institute, pp. 25–7.
6 See, for example, Billings, C.E. (1996). *Aviation automation: The search for a human-centered approach.* Hillsdale, NJ: Lawrence Erlbaum Associates.
7 Leveson, N.G. (2006). *System safety engineering: Back to the future.* Cambridge, MA: Aeronautics and Astronautics, Massachusetts Institute of Technology.
8 Mackall, D.A. (1988). *Development and Flight Test Experiences with a Flight-Critical Digital Control System* (NASA Technical Paper 2857). Lancaster, CA: National Aeronautics and Space Administration Dryden Flight Research Facility.
9 Leveson, N.G. (2006). Ibid.
10 Leveson, N.G. and Turner, C.S. (1993). An Investigation of the Therac-25 Accidents. *IEEE Computer*, 26(7), 18–41.
11 Rose, B.W. (1994). *Fatal dose: Radiation deaths linked to AECL computer errors.* Montreal, QC: Canadian Coalition for Nuclear Responsibility.
12 Leveson, N.G. (2006). Ibid.
13 United Nations Security Council (1). *Security Council, 4317th and 4318th meeting, condemns illegal exploitation of Democratic Republic of Congo's natural resources* (UN Press Release SC/7057). New York: United Nations.

14 Peat, F.D. (2002). *From certainty to uncertainty. The story of science and ideas in the twentieth century.* Washington, DC: Joseph Henry Press.
15 Capra, F. (1982). *The turning point.* New York: Simon & Schuster.
16 Capra, F. (1982). Ibid., p. 69.
17 Stapp, H.P. (1971). S-matrix interpretation of quantum theory. *Physical Review D*, 3.
18 Griffiths, P. (2009, September 8). *Climate talks must agree aviation CO_2 cap – UK adviser.* London: Thompson Reuters Foundation Alertnet.
19 Intergovernmental Panel on Climate Change (2007). *Fourth Assessment Report (AR4) of the IPCC.* More than 450 lead authors, 800 contributing authors, and an additional 2,500 reviewing experts from more than 130 countries contributed to AR4.
20 Wang, G. and Eltahir, E.A.B. (2002). Impact of CO_2 concentration changes on the biosphere-atmosphere system of West Africa. *Global Change Biology*, 8, 1169–82.
21 Intergovernmental Panel on Climate Change (2007). Ibid.
22 Burke, M.B., Miguel, E., Satyanath, S., Dykema, J.A., and Lobell, D.A. (2009). Warming increases the risk of civil war in Africa. *Proceedings of the National Academy of Sciences*, 106(49), 20670–74.
23 Anderson, C.A., Bushman, B.J., and Groom, R.W. (1997). Hot years and serious and deadly assault: Empirical tests of the heat hypothesis. *Journal of Personality and Social Psychology*, 73, 1213–23.
24 Hancock, P.A., Ross, J.M., and Szalma, J.L. (2007). A meta-analysis of performance response under thermal stressors. *Human Factors*, 49, 851–77.
25 Burke, M.B., Miguel, E., Satyanath, S., Dykema, J.A., and Lobell, D.A. (2009). Ibid., p. 20673.
26 Bertalanffy, L. von (1969). *General system theory: Foundations, development, applications.* New York: G. Brazilier.
27 Cilliers, P. (1998). *Complexity and postmodernism: Understanding complex systems.* London: Routledge.
28 Cilliers, P. (1998). Ibid.
29 Heylighen, F., Cilliers, P., and Gershenson, C. (2005). *Complexity and philosophy.* Brussels, Belgium: Vrije Universiteit, Evolution, Complexity and Cognition, p. 8.
30 Canadell, J.G. et al. (2007). Contributions to accelerating atmospheric CO_2 growth from economic activity, carbon intensity, and efficiency of natural sinks. *Proceedings of National Academy of Sciences*, 104(47), 18866–70.
31 National Science Foundation (2010, March 4). *Methane releases from Arctic shelf may be much larger and faster than anticipated* (Press Release 10–036). Washington, DC: NSF.
32 Kune, G. (2003). Anthony Eden's bile duct: Portrait of an ailing leader. *ANZ Journal of Surgery*, 73, 341–5. See also Pearson, J. (2003). *Sir Anthony Eden and the Suez Crisis: Reluctant gamble.* Basingstoke: Palgrave Macmillan.
33 Dekker, S.W.A. and Hugh, T.B. (2008). Laparoscopic bile duct injury: Understanding the psychology and heuristics of the error. *ANZ Journal of Surgery*, 78, 1109–114.
34 The West Africa connection: How drug cartels found new routes. *The Times* (of London), February 28, 2009.
35 Kauffman, S. (1993). *The origins of order.* Oxford, UK: Oxford University Press.
36 Nyce, J.M. and Dekker, S.W.A. (2010). IED casualties mask the real problem: it's us. *Small Wars and Insurgencies*, 21(2), 409–413.

37 Byrne, G. (2002). *Flight 427: Anatomy of an air disaster.* New York: Copernicus books.
38 Hutchins, E., Holder, B.E., and Alejandro Pérez, R. (2002). *Culture and flight deck operations* (Prepared for the Boeing Company, Sponsored Research Agreement 22–5003). La Jolla, CA: University of California San Diego.
39 Onderzoeksraad voor de veiligheid (2010). Aircraft Accident Report: *Neergestort tijdens de nadering, Boeing 737–800, nabij Amsterdam Schiphol Airport, 25 februari 2009.* The Hague, NL: OVV.
40 Capra, F. (1975). *The Tao of physics.* London: Wildwood House, p. 139.
41 Capra, F. (1996). *The web of life: A new scientific understanding of living systems.* New York: Anchor Books.
42 Hollnagel, E. (2004). *Barriers and accident prevention.* Aldershot, UK: Ashgate Publishing Co.
43 Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities and software.* New York: Scribner.
44 Leveson, N. (2006). Ibid.
45 Hollnagel, E. (2004). Ibid.
46 Weick, K.E. (1990). The vulnerable system: An analysis of the Tenerife air disaster. *Journal of Management*, 16, 571–593.
47 Daulerio, A.J. (2010). How one energy company will prevent catastrophic oil spills: Swivel-chair safety. [Online]. Available at: www.deadspin.com [accessed: 17 June 2010].
48 Glanz, J. and Rubin, A.J. (2007). From errand to fatal shot to hail of fire to 17 deaths. *New York Times*, October 3.
49 Risen, J. and Mazzetti, M. (2009). Blackwater guards teamed with C.I.A. on raids. *International Herald Tribune*, December 12–13, pp. 1 and 4.
50 Glanz, J. and Rubin, A.J. (2007, October 3). Ibid.
51 Columbia Accident Investigation (2003, August). *Report Volume 1.* Washington, DC: CAIB, pp. 122–3, 126 and 130.
52 Page, S.E. (2010). *Diversity and complexity.* Princeton, NJ: Princeton University Press.

7
MANAGING THE COMPLEXITY OF DRIFT

❖

It started out as pure, clear, legitimate deals. And each deal gets a little messier and messier. We started out just taking one hit of cocaine. Next thing you know, we're importing the stuff from Colombia.

Former Enron executive

COMPLEXITY, CONTROL AND INFLUENCE

The traditional view is that organizations are Newtonian–Cartesian machines with components and linkages between them. Accidents get modeled as a sequence of events (actions-reactions) between a trigger and an outcome. But such theories can say nothing about the build-up of latent failures, about a gradual, incremental loosening or loss of control that seems to characterize system accidents and drift into failure.

Remember the basic message of this book. The growth of complexity in society has got ahead of our understanding of how complex systems work and fail. Our technologies have got ahead of our theories. Our theories are still fundamentally reductionist, componential and linear. Our technologies, however, are increasingly complex, emergent and non-linear. Or they get released into environments that make them complex, emergent and non-linear. If we keep seeing complex systems as simple systems – because of the dominant logic and inherited scientific-engineering language of Newton and Descartes – we will keep missing opportunities for better understanding of failure. We will keep missing opportunities to develop fairer responses to failure, and more effective interventions before failure.

One of the patterns through which complex systems spawn accidents is the drift into failure: a slow but steady adaptation of unruly technology to environmental pressures of scarcity and competition. Traditional safety and decision theories are incapable of dealing with drift. They only take snapshots of a system (often an