Human Factors

VY Safety Letter Spezial III

DFS Deutsche Flugsicherung

member of FABEC

Vorwort

Dies ist die dritte Ausgabe des Safety Letter Spezial. Sie zeigt, dass das Thema Human Factors in der DFS zunehmend an Bedeutung und Wichtigkeit gewonnen und mittlerweile einen guten Stand erreicht hat. Dass die DFS im internationalen Vergleich sehr gut dasteht, wird unter anderem dadurch belegt, dass jede Ausgabe des Safety Letter Spezial durch die Beiträge führender und namhafter Wissenschaftler auf diesem Gebiet bereichert wird.

Ralph Riedle

Robert Schickling

Die Aufbauarbeit, die wir im Unternehmen in den vergangenen Jahren im Bereich Human Factors (HF) geleistet haben, trägt nun Früchte. HF ist etabliert und umfasst ein breites Themenspektrum von individueller Unterstützung nach kritischen Ereignissen (CISM) über die Human-Factor-Analysen bei Vorfällen (HERA), die Stärkung der Verhaltenssicherheit in Sondersituationen (Erstsprecher) bis hin zu Fragen der ergonomischen Gestaltung von Flugsicherungssystemen (Design Process Guide, DPG). Ich habe diese Entwicklung immer aktiv unterstützt und gefördert. Oftmals mussten wir uns mit Themen auseinandersetzen, die am Anfang unbequem waren, sind wir doch ein sehr technisch orientiertes Unternehmen. Doch der Weg hat sich gelohnt: So verstehen wir heute die Menschen als Garanten für den Erfolg des Unternehmens – sowohl für Sicherheit wie auch Wirtschaftlichkeit – und wir schätzen die menschliche Flexibilität, sich auf Anforderungen immer wieder neu einstellen zu können, als wesentlichen Erfolgsfaktor.

Ende des Jahres reiche ich den Stab weiter an meinen Nachfolger Robert Schickling, verbunden mit dem Wissen, dass wir uns im Bereich Human Factors nicht auf den Erfolgen ausruhen dürfen.

Ich übernehme den Stab sehr gerne und sehe den Bereich Human Factors als genauso wichtig an wie mein Vorgänger Ralph Riedle.

Die Welt hat sich verändert. Die Art und Weise, wie wir kommunizieren, die Technik, die wir einsetzen, die Schnelligkeit, mit der Informationen übermittelt und verarbeitet werden können, unterscheidet sich deutlich von der Welt vor 15 Jahren. Insofern sehe ich als große Herausforderung der nächsten Jahre, diese Veränderungen zu beleuchten, insbesondere mit Blick auf unsere Kernaufgabe Flugverkehrskontrolle. Passen unsere Denkmodelle noch? Benötigen wir neue Ansätze und Sichtweisen, um unser System besser zu verstehen und dadurch genauer kontrollieren und steuern zu können?

Die Aufbauarbeit, die wir im Unternehmen in den vergangenen Jahren im Bereich Human Factors (HF) geleistet haben, trägt nun Früchte.

Hierzu gibt es Ansätze und viele davon sind in den Safety-Letter-Spezial-Ausgaben beschrieben. Es sind innovative Ideen, gekennzeichnet von Querdenken und neuen Theorien. Und wie das so ist bei neuen Ideen, sind diese nicht immer leicht zugänglich und einfach verständlich. Sie wirken sperrig, da sie Bisheriges in Frage stellen, aber sie sind nötig, um Veränderung und Entwicklung voranzutreiben. Vor dieser Aufgabe sehe ich das gesamte Unternehmen DFS und im Besonderen den Bereich Human Factors als innovative Kraft.

Wir werden alle – Geschäftsführung, Führungskräfte und Mitarbeiter – dazu beitragen, dass wir in ein paar Jahren auch hier auf eine erfolgreiche Entwicklung zurückblicken können.

Ralph Riedle

Robert Schickling

Inhalt

Einleitung

Sie halten bereits die dritte Ausgabe der Reihe Safety Letter „Human Factors Spezial“ in Ihren Händen. Während in der ersten Ausgabe ein theoretisches Fundament für unterschiedliche Human-Factors-Themen gelegt wurde, hatte die zweite Ausgabe ihren Fokus auf die Vorfalluntersuchung in der DFS gelegt. In beiden Ausgaben findet sich eine Veränderung des Verständnisses unseres Arbeitsumfeldes von einer linearen zu einer komplexen Sichtweise.

„Wir leben in einem komplexen sozio-technischen System“ ist eine Aussage, die sich an mehreren Stellen wiederfindet. Es scheint fast schon so zu sein, dass der Komplexitätsbegriff zu einem Modewort geworden ist. Kaum ein Interview eines Managers, sei er aus der Finanzwelt, der Luftfahrt, der Autoindustrie oder eines anderen Industriezweiges, kommt ohne diesen Begriff aus. Finanzprodukte werden dann genauso wie das Cockpit eines neuen Automobils als komplex bezeichnet. Auch die Konstruktion eines neuen Flugzeuges bekommt schnell diese Bezeichnung (man denke nur an die Einführung der neuen Airbus A380). Was meinen wir eigentlich damit, wenn etwas komplex ist und welche Implikationen ergeben sich dadurch? Im Alltag bezeichnen wir Systeme häufig als komplex, die wir nicht oder nur mit großem Aufwand verstehen. Man denke nur daran, wie kompliziert es war, die Uhrzeit eines Videorecorders von Sommer- auf Winterzeit umzustellen.

Die Wissenschaft, die sich in den letzten Jahrzehnten vermehrt um die Erforschung von komplexen technischen und sozialen Systemen gekümmert hat, hat uns ein heterogenes Bild hinterlassen. Wir haben, wie es häufig üblich ist, mehr Fragen als Antworten generiert. Handliche Tools, die uns helfen, Komplexität zu managen, sind rar gesät. Einigkeit besteht aber darin, dass wir mit unserem jetzigen Verständnis und unseren derzeitigen Methoden nicht in der Lage sind, die immer größer werdende (gefühlte) Komplexität unserer Systeme zu verstehen und Erfolg versprechende Maßnahmen abzuleiten.

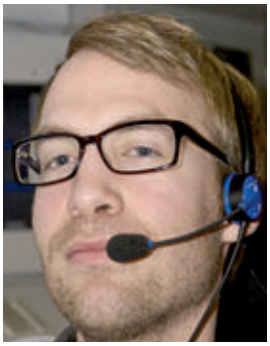
„Our technologies have gotten ahead of our theories.” So fasst es Prof. Sidney Dekker zusammen. In seinem Artikel werden neben den grundsätzlichen Merkmalen der Komplexitätstheorie auch konkrete Implikationen für die Vorfalluntersuchung aufgezeigt. Wenn wir die Perspektive der Komplexität einnehmen, heißt das auch, dass wir einfache Antworten zurückweisen. Vielmehr ist uns aber an unterschiedlichen Sichtweisen gelegen und wir versuchen, aus der daraus resultierenden Diversität den für uns am besten geeigneten Ansatz nutzbar zu machen. Die DFS begegnet der Komplexität ihrer Produkte, Prozesse und Systeme, indem sie neue Wege geht und wagt, Altbewährtes kritisch zu hinterfragen und gegebenenfalls auch aufzugeben. Mit dem Projekt „Weak Signals“ versuchen wir, frühe Anzeichen von Veränderungen besser zu verstehen. Dabei macht es keinen Unterschied, ob diese Veränderungen negativ oder positiv erscheinen. Die Vorfalluntersuchung in der DFS versuchen wir weiterzuentwickeln. Wir schulen Mitarbeiter und Führungskräfte und vermitteln Inhalte und Implikationen von unterschiedlichen Human-Factors-Ansätzen.

Die vorliegende dritte Spezialausgabe des Safety Letter „Human Factors Spezial“ befasst sich dementsprechend schwerpunktmäßig mit dem Thema Komplexität und dem Verständnis von Sicherheit. Uns ist bewusst, dass die Artikel sehr anspruchsvoll und nicht immer einfach zu lesen sind, aber wir sind überzeugt, dass nur so eine Weiterentwicklung auf einem bereits hohen Niveau machbar ist.

Wie immer freuen wir uns über Anregungen, Kritik und einen regen Austausch.
Viel Spaß beim Lesen!

Christoph Peters und Jörg Leonhardt

Was verstehen wir unter Safety?



Christoph Peters ist seit 2000 in der DFS und arbeitet zu 50 % als Lotse bei Düsseldorf APP in der Niederlassung Mitte und zu 50 % als Referent Human Factors im Unternehmenssicherheitsmanagement (VYIH). Peters hat 2009 ein Psychologiestudium an der Universität Mainz abgeschlossen und ist seit 2010 hälftig bei VYIH tätig.

Von Christoph Peters

„Aus Vorfällen lernen ist eine der zentralen Aufgaben." Diese Aussage stand in der letzten Ausgabe ganz am Anfang des Vorwortes und unterstreicht somit auch den Anspruch und das Ziel, dem wir im Safety-Management nachgehen. Doch wie können wir die richtigen Schlüsse aus Vorfällen ziehen? Sind Vorfälle überhaupt geeignet, um daraus zu lernen?

Um diese Fragen beantworten zu können, gilt es zuerst zu klären, welches Grundverständnis wir überhaupt darüber haben, wie sich Unfälle ereignen. Hier gibt es weder eine objektive Mustererklärung noch eine generelle Übereinstimmung, sondern unterschiedliche Annahmen, Überzeugungen und Vorstellungen. Ist man beispielsweise davon überzeugt, dass die Ursache der meisten Unfälle auf menschliches Versagen zurückzuführen ist, wird man viele Indizien und Informationsstücke im Licht dieser Bewertung sehen. Dementsprechend wird man in einer Vorfalluntersuchung eher nach technischen Aspekten suchen, sofern man der Überzeugung ist, dass unausgereifte Technik die Hauptursache von Unfällen darstellt. Diese Einstellungen bündeln sich in einem sogenannten Unfallmodell, das entweder explizit oder implizit jeder Vorfalluntersuchung unterliegt. Das Modell wird die Untersuchung in eine bestimmte Richtung lenken und dazu führen, manche Aspekte sehr genau zu betrachten und andere Informationen eher auszublenden. Eine Vorfalluntersuchung folgt somit dem Prinzip „What you look for is what you find" (Lundberg et al., 2009).

Die Bewertungen und Vorstellungen, wie sich Unfälle ereignen, haben sich dabei in den letzten Jahrzehnten signifikanten Veränderungen und Entwicklungen unterzogen (Hollnagel, 2004). In der Regel hat erst eine Beinahekatastrophe oder ein Großschadensereignis ein Umdenken bzw. Ablösen alter Vorstellungen und Überzeugungen vorangetrieben. In der historischen Betrachtung lassen sich drei große Faktoren identifizieren, die in Unfallmodellen und Safetytheorien eine wesentliche Rolle spielen: Technik, Mensch und Organisation. Je nach Modell wird möglicherweise ein Faktor stärker als der andere berücksichtigt oder es blendet einen oder gar mehrere Faktoren aus. Die Veränderungen des jeweiligen gesellschaftlichen und technologischen Kontextes machten eine Anpassung des Unfallmodells notwendig, um z. B. neue oder weitere Faktoren bzw. dynamische oder komplexe Zusammenhänge in der Analyse zu berücksichtigen.

Das Voran- und Weiterschreiten in der Entwicklung von Unfallmodellen kann bis zum Jahrtausendwechsel in drei große „Zeitalter" eingeteilt werden (Hale & Hovden, 1998). Jeder Abschnitt bediente sich eines bestimmten Modells und kam somit auf einen für diese Zeit charakteristischen Erklärungsansatz. Das Verständnis von Safety und die Erklärungsansätze für Unfälle haben sich stets verändert. Auch wir in der DFS müssen stets kritisch hinterfragen, ob unsere Ansätze noch geeignet sind, die gegenwärtigen Gegebenheiten zutreffend abzubilden und ob Erkenntnisse für die Verbesserung unseres Systems daraus abzuleiten wären.

Doch fangen wir der Reihe nach an und blicken auf das erste Zeitalter, das die frühen Jahre und ersten Entwicklungen im Luftverkehr abbildet.

„Accidents happen because technology fails."

Die Notwendigkeit, sich systematisch mit Unfällen und deren Verhütung auseinanderzusetzen, begann vor allem mit der Industriellen Revolution, die am Anfang des 19. Jahrhunderts in England

startete und einige Jahrzehnte später auch in Deutschland entstand. Die Fabriken, Werkstätten und Produktionsstandorte hatten zum ersten Mal eine derartige Größe und Personenstärke erreicht, dass eine Katastrophe nicht nur einige wenige, sondern eine enorme Anzahl von Menschen getroffen hätte. Dies erhöhte die Notwendigkeit, sich systematisch mit dem Thema Sicherheit auseinanderzuset-

Nach dieser Logik kann der Unfall verhindert werden, indem ein oder mehrere Dominosteine weggenommen werden oder das Fallen eines Steines verhindert wird.

zen. Es dauerte aber noch bis zum Anfang des 20. Jahrhunderts, ehe Herbert William Heinrich 1931 sein Buch „Industrial Accident Prevention: a scientific approach“ veröffentlichte, das ein Meilenstein in der Sicherheitsforschung darstellt.

In diesem Buch wurde auch ein Accident-Modell spezifiziert, das sich auch heutzutage noch größter Beliebtheit erfreut: das sogenannte Domino-Modell. Nach dieser Annahme treten Ereignisse vor einem Unfall in einer linearen und festgelegten Reihenfolge auf, wobei am Ende dieser Ereigniskette ein Unfall steht. Dies führte auch zu der Vorstellung einer finalen Hauptursache („Root Cause“), die am Anfang der Dominoreihe stand und sie anstieß. Modelle, die in dieser Tradition stehen, werden auch als „Sequence of events“-Modelle bezeichnet. Nach dieser Logik kann der Unfall verhindert werden, indem ein oder mehrere Dominosteine weggenommen werden oder das Fallen eines Steines verhindert wird. Im Dominomodell erhöht man die Sicherheit auch, indem man die Abstände zwischen den Dominosteinen vergrößert und dadurch verhindert, dass ein fallender Stein den anderen umstößt. Safety Briefings und Awareness-Kampagnen, die darauf abzielen, dass der Dominostein stabilisiert wird oder nicht fällt, sind Beispiele dieses Ansatzes. Im Kontext Flugsicherung könnte das z. B. auch bedeuten, nicht zu eng an die Sicherheitsmargen heranzugehen (z. B. im Endanflug an verkehrsreichen Flughäfen). Weitere Faktoren, die auf das System einwirken (z. B. Produktionsdruck) werden mit diesem Unfallmodell nicht betrachtet. Auch die täglichen Anpassungen, die in komplexen Systemen nötig sind (Performance-Variabilität) werden mit diesem Modell nicht betrachtet, was zu dieser Zeit jedoch auch nicht zwingend nötig war. Viele Handlungsabläufe waren aufgrund der weit verbreiteten

Fließbandarbeit sehr linear und einfach. Das Domino-Modell ist heutzutage immer noch eine beliebte Metapher. Verändert hat sich, dass darin nicht mehr nur die Technik, sondern auch der Mensch und die Organisation einfließen. Die Grundidee des Modells ist aber gleich geblieben, nämlich die rein lineare Struktur des Modells und die Überzeugung, dass es um das Scheitern von (isolierbaren) Komponenten geht.

In diesem ersten Zeitalter, das vor allem den technischen Erfindungsreichtum und auch die Anfänge kommerzieller Luftfahrt kennzeichnete, lag der Fokus auf dem Faktor Technik. Dies ergab für die damaligen Zeitumstände auch durchaus Sinn. Man denke nur an die Unfallserie des ersten in Serie gebauten Düsenverkehrsflugzeugs der Welt, die de Havilland „Comet“. Anfang der 1950er Jahre musste die BOAC (British Overseas Airways Corporation) eine Serie von Totalverlusten beklagen. Es stellte sich später heraus, dass der entscheidende Faktor die Materialermüdung war. Sowohl Probleme mit der Druckkabine als auch das Design von rechteckigen Fenstern übten einen derart hohen Druck auf die Struktur des Flugzeuges aus, dass das Flugzeug in der Luft regelrecht zerbrach. Der Mensch rückte in dieser Zeit noch nicht in den Fokus der Aufmerksamkeit. Dies sollte sich jedoch bald ändern.

Man denke nur an die Unfallserie des ersten in Serie gebauten Düsenverkehrsflugzeugs der Welt, die de Havilland „Comet“.

„Accidents happen because the human (factor) fails.“

Nachdem ab den 1950er Jahren Technik stetig zuverlässiger wurde, war es eine logische Konsequenz, dass nun der Mensch in den Vordergrund der Aufmerksamkeit geriet. Schnell bildete sich die Kategorie „Human Error“ neben dem technischen Faktor heraus. Die Ursache von Unfällen wurde nun entweder dem menschlichen Versagen oder der Technik zugeschrieben. Dass diese beiden Faktoren die Vorfalluntersuchung maßgeblich geprägt haben, sieht man beispielsweise an dem Fall der DC 10 von Air New Zealand, die im November 1979 über der Arktis in den „Mount Erebus“ flog (für eine ausführliche Schilderung dieses Falles siehe Ausgabe 2 des Safety Letter Spezial).

Im Rahmen der ersten Untersuchung wurde zuerst versucht, technische Mängel als Absturzursache zu finden. Als sich dieser Verdacht jedoch nicht erhärten ließ, war die logische Konsequenz, menschliches Versagen als Ursache anzunehmen. Als Folge dieser Überzeugung wurden viele (neutrale) Informationsstücke als Indiz für Fehlhandlungen der Pilotencrew ausgelegt. Erst eine zweite Untersuchung machte jedoch deutlich, dass der Ansatz „Human Error“ deutlich zu kurz griff und keine befriedigende Erklärung für das Zustandekommens dieser Tragödie war. Es war dann auch dieser

Absturz und einige Monate früher die partielle Kernschmelze im Kernkraftwerk „Three Mile Island“ in Harrisburg (USA), die aufzeigten, dass die vorhandenen Unfallmodelle nicht in der Lage waren, zufriedenstellende Antworten auf die erhöhte Systemkomplexität zu geben. Somit wurde, neben den bereits etablierten Faktoren „Mensch“ und „Technik“, die „Organisation“ als weiterer Baustein in die Vorfallanalyse eingeführt, was zu einer Überarbeitung des vorherrschenden Paradigmas geführt hat.

Prominentester Vertreter dieser sogenannten epidemiologischen Unfallmodelle ist zweifelsfrei das latente Fehlermodell, besser bekannt als „Swiss Cheese“-Metapher (Reason, 1990; 1997). Diese Modelle vergleichen das Zustandekommen von Unfällen mit der Verbreitung von Krankheiten (einer Epidemie). Hierzu benötigt es ein Zusammenwirken von manifesten und latenten Faktoren in ähnlicher zeitlicher und räumlicher Nähe. Statt eines großen „Single Failure“, wie es noch im Domino-Modell angenommen wurde, ist nun das Auftreten von mehreren beitragenden Faktoren und das Versagen bestimmter Merkmale entscheidend. Ganz in der Analogie zu einem Schweizer Käse braucht es eine spezifische Kombination von Faktoren, die die Löcher in allen Käsescheiben passieren können.

Mit der Verbreitung der „Swiss Cheese“-Metapher wurden gleichzeitig zwei Charakteristika eingeführt, die im Domino-Modell noch nicht zu finden waren: Zum einen war dies die Idee einer Barriere (also einer Käsescheibe). Eine Barriere kann dazu beitragen, die Katastrophe oder ihre Entwicklung zu verhindern. Es handelt sich zumeist um ein Merkmal, das eine spezielle Abwehraufgabe erfüllt, also z. B. das STCA-System, TCAS oder das Vier-Augen-Prinzip. Eine Organisation verfügt nun über eine ganze Reihe von Barrieren, die die Absicht verfolgen, eine potenziell gefährliche Entwicklung zu stoppen. Zum anderen finden sich in dieser Modellfamilie latente Bedingungen. Hierbei handelt es sich um Systembedingungen, die schon längere Zeit im System „schlummern“ bzw. unentdeckt sind und erst bei einem unerwünschten Ereignis sichtbar bzw. gefunden werden.

Diese Unfallmodelle stellen einen enormen Fortschritt dar, weil sie den Unfall als komplexes Zusammenspiel mehrerer Faktoren ansehen. Für das Zustandekommen einer Katastrophe ist jetzt nicht mehr eine große singuläre Ursache verantwortlich, sondern das Auftreten mehrerer Umstände, die alle Verteidigungsbarrieren eines Systems überwinden und durch alle Löcher (latente Systembedingungen) laufen. Des Weiteren wurde damit die Linse einer Vorfalluntersuchung verbreitert und der Schwerpunkt von den sogenannten Front-Line-Operateuren („Sharp End“) auf die in einer Organisation gerichteten Bedingungen verlagert („Blunt End“).

Die Nachwehen zum Reaktorunglück in Harrisburg bedeuteten nicht nur einen wichtigen Impuls zur Weiterentwicklung von Unfallmodellen, sie stießen auch die Entwicklung von zwei wichtigen Theorien an, die die weitere Entwicklung auf dem Gebiet der Sicherheitsforschung entscheidend prägten.

In der sogenannten „Normal Accident Theory“ NAT (Perrow, 1984) wird die Grundidee vertreten, dass Unfälle in komplexen und eng gekoppelten Systemen letztendlich unvermeidlich sind. Paradoxerweise, so Perrow, erhöhen die Barrieren und Redundanzen, die eigentlich zum Schutz eines Systems existieren, dessen Komplexität. Die Notwendigkeit, ein System zuverlässiger zu machen, macht es aber gleichzeitig auch komplexer. Dabei sind zwei wesentliche Dimensionen zu betrachten. Dies ist zum einen der Komplexitätsgrad, den Perrow in „linear“ und „komplex“ unterscheidet. Komplexe Systeme sind z. B. dadurch gekennzeichnet, dass nicht alle Prozesse in einer Organisation komplett verstanden werden können, viele Systemparameter eine Vielzahl an Interaktionen

aufweisen oder indirekte Informationsquellen existieren. Systeme können nicht nur komplex oder linear sein, sondern sich auch in einer zweiten Dimension unterscheiden, dem Kopplungsgrad. Eine enge Kopplung bedeutet beispielsweise, dass Verzögerungen oder zeitliche Stopps im Produktionsablauf nicht möglich sind und diese Abläufe speziell ausgebildetes Personal erfordern. In Systemen, die sowohl einen hohen Grad an interaktiver Komplexität und eine enge Kopplung aufweisen, ist eine Katastrophe nach der NAT ein systeminhärentes Ereignis. Trotz der besten Vorstellungskraft wird es immer unvorhersehbare Entwicklungen geben. Diese können dann durch die enge Verschachtelung im System schnell durch das gesamte System strömen und in einem Desaster eskalieren.

Da es bei vielen Unfällen ähnliche Muster zu identifizieren gibt, müssen wir besser verstehen, wie unser System überhaupt funktioniert.

„Things that never happened before happen all the time."

(Sagan, 1993)

Ein zweiter, wesentlich optimistischerer Ansatz wurde von einer Forschergruppe aus Berkeley, die Todd LaPorte gründete, formuliert (Rijpka, 1997): die sogenannte „High Reliability Theory" HRT. Im Rahmen dieses Ansatzes wurden Organisationen erforscht, die einen fehlerfreien oder extrem fehlerarmen Betriebsablauf aufweisen, obwohl sie in einem Hochrisikoumfeld operieren. Dazu gehören beispielsweise amerikanische Atom-U-Boote oder Flugzeugträger (Rochlin, LaPorte & Robert, 1987). Die Vertreter dieser Schule beanspruchen für sich, Strategien gefunden zu haben, mit denen komplexe und eng gekoppelte Organisationen eine sehr gute Safetybilanz erzielen können. Zu diesen Prinzipien, die eine „Mindful Organisation" ausmachen, gehören z. B. die Abneigung zur Vereinfachung oder die Entwicklung eines guten Gespürs für betriebliche Abläufe. Wissen bzw. Expertise wird ein großer Stellenwert zugeschrieben, sodass Entscheidungen von denjenigen getroffen werden sollten, die in einer Situation das optimale Wissen haben, unabhängig von ihrem Dienstgrad (Weick & Sutcliffe, 2007).

Beide Denkschulen bilden letztendlich zwei verschiedene Ansätze, die zu unterschiedlichen Schlussfolgerungen kommen, aber auf das gleiche Phänomen schauen: Safety. Unabhängig davon, ob man eher der NAT oder der HRT zuneigt, beide Theorien lehnen einfache Erklärungsmuster ab. Beide Ansätze berücksichtigen die gestiegene Komplexität in Organisationen. Beide Denkschulen waren mit den vorhandenen Denkmodellen unzufrieden und entwickelten alternative Ansätze. Da nun nicht nur die Technik und der Mensch, sondern jetzt vor allem auch die Organisation mit berücksichtigt wurde, lautet das Motto des dritten Zeitalters, das sich Anfang der 1990er Jahre bildete:

„Accidents happen because the organisation fails."

Die epidemiologischen Unfallmodelle, allen voran „Swiss Cheese", haben zwar die Organisation als einen weiteren wichtigen Aspekt aufgenommen, erklären aber nicht andere wesentliche Aspekte, z. B. wo genau denn die Löcher im Käse auftreten, wie sie sich verändern und wieder geschlossen werden können. Sie haben die Simplifizierung im Dominomodell zwar korrigiert, indem sie von einer Kombination von Faktoren sprechen. Sie erklären aber nicht, weshalb sich latente Bedingungen (Löcher im Käse) über die Zeit verändern, sowohl in ihrer Größe als auch an ihrem Ort. Des Weiteren bleibt die Grundstruktur dieser Modelle sequenziell und linear. Sie suchen weiterhin nach fehlerhaften Komponenten, die nun entweder der Kategorie Technik, Mensch oder Organisation zugeschrieben werden.

Die wissenschaftliche Betrachtung von Unfällen der jüngeren Zeit (z. B. Dekker, 2004; Snook, 2000; Vaughan, 1996) hat jedoch gezeigt, dass es notwendig ist, die Betrachtungsebene zu ändern und ein neues Verständnis von Safety zu entwickeln. Dabei steht nicht mehr die Jagd nach fehlerhaften Komponenten bzw. Barrieren im Vordergrund, sondern das System in seiner Gesamtheit. So ist beispielsweise das langsame und häufig unentdeckte Driften von einem ursprünglich sicheren Zustand zu den Grenzen

und Abgründen eines Systems ein normales, alltägliches Nebenprodukt komplexer Organisationen, die unter Ressourcenknappheit leiden (Dekker, 2011). Diese „Drift into Failure“ geschieht Schritt für Schritt, wobei die einzelnen Veränderungen meistens klein sind und daher unbemerkt bleiben. Wir benötigen daher Modelle, die analysieren, wie die tägliche Arbeit überhaupt funktioniert. Verstehen wir überhaupt, wieso in über 99,9 Prozent aller Situationen alles gut geht? Erst eine systemische Betrachtung liefert uns ein Verständnis darüber, wo wir überhaupt im Gesamtsystem stehen und wohin wir treiben. Die Suche nach fehlerhaften Komponenten wird unseren Erkenntnisstand in dieser Sache nicht erhöhen. Nur wenige Unfallmodelle beschäftigen sich intensiver mit dieser Thematik, doch die Anfänge sind mit einem systemischen Modell wie z. B. STAMP gemacht (Leveson, 2004). Auch die Analyse alltäglichen Arbeitens wird durch die sogenannte „Functional Resonance Analysis Method“ näher beleuchtet (Hollnagel, 2012). Mit dieser Methode wird z. B. der Fragestellung nachgegangen, welche Systemfunktionen für einen bestimmten und wichtigen Ausschnitt aus der Flugsicherungswelt existieren, wie sie aufgebaut und vernetzt sind und welche Auswirkungen sie auf das Gesamtsystem haben.

Die Welt verändert sich stetig. Dies macht sich in vielen unterschiedlichen Lebensbereichen und veränderten Abläufen in Organisationen bemerkbar. Die Art und Weise, wie wir arbeiten und kommunizieren, hat sich enorm verändert. Die Schnelligkeit und Dynamik, in der Informationen ausgetauscht werden, ist beträchtlich gestiegen. Veränderungen, die sich um uns herum ereignen, müssen auch wir als Safety-Management antizipieren, damit geeignete Maßnahmen ergriffen werden können. Wir brauchen Methoden und Tools, die eine praktische Relevanz haben und mehr als einen theoretischen Wert besitzen. Dafür ist es aber auch wichtig zu verstehen, welche Denkmodelle welches Lernpotenzial besitzen und welche Ergebnisse liefern. Incidents können dabei einen Beitrag zum Lernen liefern, da sie uns aufzeigen, wie das System mit einer unerwünschten Situation umgegangen ist und inwieweit es gut oder schlecht reagiert hat. Allerdings dürfen wir uns aber nicht ausschließlich um negative Ereignisse kümmern, sondern sollten den Blick auch auf das positive und alltägliche Funktionieren unserer Systeme richten.

Der Bereich VY/H hat diesbezüglich mehrere Aktivitäten gestartet, um die gestiegene Komplexität in unserem Gesamtsystem besser zu verstehen. Dazu gehören unter anderem die Auseinandersetzung mit dem Thema „Resilience Engineering“ und das von EUROCONTROL ausgeschriebene Projekt „Weak Signals“, das die DFS für sich gewinnen konnte. Die neuen Erkenntnisse werden in die Schulung der Vorfalluntersucher einfließen. Diese Anstrengungen sind notwendig, um die Dynamik und Komplexität in dem System Flugsicherung weiter besser zu verstehen und es dadurch weiter sicher zu halten.

Literaturhinweise:

Dekker, S.W.A. (2004). Why we need new accident models. *Human Factors and Aerospace Safety 4*(1), 1-18.

Dekker, S.W.A. (2011). *Drift into Failure: From Hunting Broken Components to Understanding Complex Systems*. Aldershot: Ashgate.

Hale, A.R. & Hovden, J. (1998). Management and culture: the third age of safety. A review of approaches to organizational aspects of safety, health and environment. In: Feyer, A.M & Williamson, A. (Eds.). *Occupational Injury. Risk Prevention and Intervention*. London: Taylor & Francis.

Heinrich, W.A. (1931). *Industrial accident prevention*. New York: McGraw-Hill.

Hollnagel, E. (2004). *Barriers and Accident Prevention*. Aldershot: Ashgate.

Hollnagel, E. (2012). *FRAM-The Functional Resonance Analysis Method*. Farnham, UK: Ashgate.

Leveson, N. G. (2004). A New Accident Model for Engineering Safer Systems. *Safety Science, 42*(4), 237-270.

Lundberg, J., Rollenhagen, C. & Hollnagel, E. (2009). What-You-Look-For-Is-What-You-Find- The consequences of underlying accident models in eight accident investigation manuals. *Safety Science*, 1297-1311.

Perrow, C. (1984). *Normal Accidents. Living with high-risk technologies*. New York: Basic Books.

Reason, J.T. (1990). *Human Error.* Cambridge, U.K: Cambridge University Press.

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

Rijpma, J.A. (1997). Complexity, Tight-Coupling and Reliability: Connecting Normal Accidents Theory and High Reliability Theory. *Journal of Contingencies and Crisis Management*, 5(1), 15-23.

Rochlin, G., LaPorte, T.R. & Robert, K.H. (1987). The self-designing high reliability organization: Aircraft carrier flight operations at sea. *Naval War College Review*, (Autumn), 76-90.

Sagan, S.D. (1993). *The Limits of Safety: Organizations, Accidents and Nuclear Weapons*. Princeton, NJ: Princeton University Press.

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

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

Weick, K.E. & Sutcliffe, K.M. (2007). *Managing the unexpected: resilient performance in an age of uncertainty* (2nd edition). San Francisco: Jossey-Bass.

Charles B. Perrow. Professor Emeritus of Sociology, Yale University. Author of the award winning book "Normal Accidents: Living with high risk industries".

Für diese Ausgabe des Safety Letter Spezial sind wir sehr stolz darauf, dass wir Prof. Charles Perrow für einen Gastbeitrag gewinnen konnten. Charles Perrow ist emeritierter Professor für Soziologie an der Yale-Universität, hat sechs Bücher und mehr als 50 Papers veröffentlicht. Einem größeren Publikum wurde Prof. Perrow durch die Publikation seines Buches „Normal accidents: Living with high risk industries" im Jahre 1984 bekannt. In diesem Werk formulierte Prof. Perrow die sogenannte „Normal Accident Theory" (NAT). Die grundsätzliche Aussage dieses Ansatzes ist, dass in komplexen und eng gekoppelten Organisationen ein Unfall zwar selten, aber unvermeidbar ist. Zusammen mit einem alternativen Ansatz, der sogenannten „High Reliability Theory", regte der sich anschließende wissenschaftliche Diskurs eine Vielzahl von Forschungsarbeiten an. Prof. Perrow hat uns freundlicherweise erlaubt, eine wissenschaftliche Arbeit über die Nuklearkatastrophe von Fukushima abzudrucken. In diesem Artikel werden auch einige Aspekte der NAT aufgegriffen. Des Weiteren hat Prof. Perrow extra für diese Publikation seinen Text aufgrund neuerer Entwicklungen in Japan angepasst.

Fukushima and the inevitability of accidents

By Charles B. Perrow

The March 11, 2011 disaster at the Fukushima Daiichi Nuclear Power Station in Japan replicates the bullet points of most recent industrial disasters. It is outstanding in its magnitude, and eventually might even surpass Chernobyl in its effects. But in most other respects, it simply indicates the risks that we run when we allow high concentrations of energy, economic power, and political power to form. Just how commonplace – prosaic, even- this disaster was illustrates just how risky the industrial and financial world really is.

Nothing is perfect, no matter how hard people try to make things work, and in the industrial arena there will always be failures of design, components, or procedures. There will always be operator errors and unexpected environmental conditions. Because of the inevitability of these failures, and because there are often economic incentives for business not to try very hard to play it safe, government regulates risky systems in an attempt to make them less so. Formal and informal warning systems constitute another method of dealing with the inherently risky systems of industrial society. And society can always be better prepared to respond when accidents and disasters occur.
But for many reasons, even quality regulation, close attention to warnings, and careful plans for responding to disaster cannot eliminate the possibility of catastrophic industrial accidents. Because that possibility is always there, it is important to ask whether some industrial systems have such huge catastrophic potential that they should not be allowed to exist.

Regulations

Nuclear safety is problematic when nuclear plants are in private hands because private firms have the incentive and, often, the political and economic power to resist effective regulation. That resistance often results in regulators being captured in some way by the industry. In Japan and India, for example, the regulatory function concerned with safety is subservient to the ministry concerned with promoting nuclear power and, therefore, is not independent. The United States had a similar problem that was partially corrected in 1975 by putting nuclear safety into the hands of an independent agency, the Nuclear Regulatory Commission (NRC), and leaving the promotion of nuclear power in the hands of the Energy Department. Japan is now considering such a separation. It should make one. Since the accident at Fukushima, many observers have charged that there is a revolving door between industry and the nuclear regulatory agency in Japan – what the New York Times called a "nuclear power village" – compromising the regulatory function.

Regulatory capture is widespread in many risky US industrial systems and often subtle – but not always.

Of course, even in Europe, where forprofit firms have less power, there are safety problems that have needed more effective oversight. But by and large, European nuclear plants, which are generally part of a state-run industry, appear to be safer than the privately owned, poorly regulated nuclear plants in the United States, Japan, and other countries.

Systemic regulatory failure – as opposed to simple error – is tricky to identify accurately. After an accident in a risky industry, it is always possible to find some failure of a regulatory agency. Everything, after all, is subject to error, in regulatory agencies as well as chemical or power plants. To say that regulation failed on a system-wide basis, one must have strong evidence of agency incompetence or collusion.

The Union Carbide chemical plant in Institute, West Virginia, is my favorite example of regulatory incompetence; in this case, it was a matter of regulators seeing what they were apparently predisposed to see. Shortly after a Union Carbide pesticide plant in Bhopal, India, leaked methyl isocyanate gas in December 1984, killing thousands, the Occupational Safety and Health Administration (OSHA) inspected the company´s West Virginia plant and gave it a clean bill of health. What happened in Third World India could not happen in the United States, it was said.

Nine months later, an accident quite similar to Bhopal occurred at the plant, though the gas released was not as toxic and the wind was in a favorable direction, so only some 135 people were hospitalized (Perrow, 2011: 179-180). OSHA looked again and, predictably, found “an accident waiting to happen” and cited the plant for numerous violations, despite its clean bill of health nine months before. There was a trivial fine and a Union Carbide promise to store only the small amounts of the toxic gas actually needed for production.

Before 2011, these experts were largely ignored. Japan has 53 nuclear power plants drawing cooling water from the ocean.

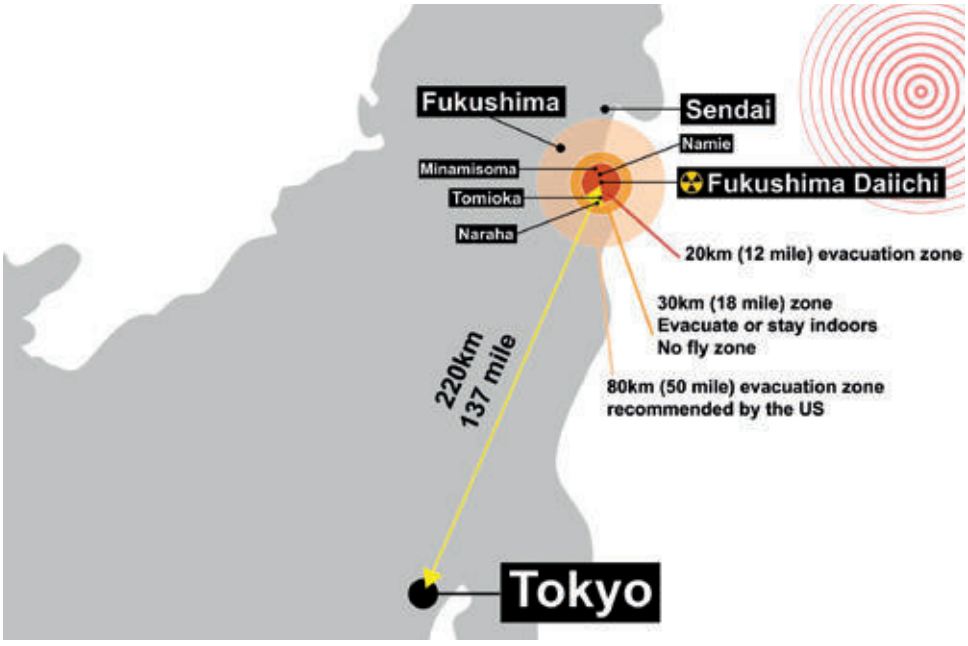
Union Carbide soon resumed massive storage of methyl iscocyanate. Bayer subsequently took over the plant and, in 2008, an explosion killed two workers and threatened to release 4,000 gallons of the deadly gas. Subsequent investigation by the US Chemical Safety Board again found an accident waiting to happen. OSHA appears not to have noticed that its strictures on the amount of storage were violated.

Regulations will always be imperfect. They cannot cover every exigency, and, unfortunately, almost anything can be declared the cause of an accident. One can also make the case that too much regulation interferes with safe practices, as nuclear plant operators have always claimed in the United States. But the overregulation complaint is undermined by the following anecdote: A few years ago, the NRC sharply increased the number of inspections of nuclear power plants following some embarrassing near-misses. A then-powerful US senator, Pete Domenici of New Mexico, a recipient of large campaign donations from the industry, called in top NRC officials and threatened to cut the agency´s budget in half if it did not reduce the number of inspections (Mangels, 2003). The NRC reduced its inspections. I doubt that anything similar could take place in Europe.

Regulatory capture is widespread in many risky US industrial systems and often subtle –but not always. In the Interior Department´s Materials Management Service, for example, representatives of the oil industry and regulators who were supposed to be overseeing oil exploration exchanged sexual favors and drugs. This intramural partying was disclosed just before the BP-leased Deepwater Horizon oil rig blew up in the Gulf of Mexico, resulting in the largest oil spill of its kind, making the regulatory failure especially dramatic.

Charges of regulatory failure were also levied in the 2010 Massey Energy coal mine disaster in West Virginia, which killed 29; the explosion at BP´s Texas City, Texas, refinery in 2005, which killed 15 and injured at least 170; and BP´s massive oil pipeline break in 2006 in Prudhoe Bay, Alaska.

There are many forms of regulatory failure. Regulations on the books can lie dormant by the common consent of regulators and industry. A worker at the Millstone nuclear power plant in Connecticut kept warning management that the spent fuel rods were being put too quickly into the spent storage pool



and that the number of rods in the pool exceeded specifications. Management ignored him, so he went directly to the NRC, which eventually admitted that it knew of both of the forbidden practices, which happened at many plants, but chose to ignore them. The whistleblower was fired and blacklisted.

Rather than completely ignore regulations, a captured regulatory agency may just lower the standards it uses. The NRC has consistently lowered standards for emergency electric power supplies in US nuclear plants. And in the wake of the Fukushima disaster, the government of Japan is lowering standards for allowable doses of radiation.

Regulations are only as good as their enforcement, and here the evidence is fairly uniform: Enforcement is generally lax and often all but nonexistent. Workers at Fukushima reported that they had advance warnings of inspections, and inspectors regularly winked at violations. The record of the NRC is similar in the United States; for example, when utilities complained about the standards for fire prevention at nuclear plants in recent years, the regulators lowered the standards.

Not surprisingly, we also do not find that warnings of looming industrial and financial disasters have much impact.

Even when safety inspections find violations, there is no guarantee that the regulated firm will be moved to change its practices. In many cases, the fines levied are too small to be a deterrent. After BP´s huge spill in Prudhoe Bay, the company was fined less than its profits for one day of operation. After the Texas City refinery explosion, the New York Times reported that OSHA had levied a record fine of $87 million against the firm. According to BP, it made a profit of about $14 billion in 2009, meaning the fine amounted to about six-tenths of a percent of its profit for the year. An official of OSHA subsequently testified to the agency´s weakness and the power of the petrochemical industry by noting that the size of the fines levied did not deter; firms repeated the same glaring mistakes despite their costs, ignored warnings, and harassed workers who warned of wrongdoing.

Warnings

Even if a risky system is only loosely regulated, a point will come when warnings are loud enough to attract attention. Catastrophes are expensive; no one wants them. The overall experience with warnings about global warming, however, should caution against expecting warnings to be too effective. A well-funded but factually challenged campaign to deny that climate change is the result of human activity has managed to ice the climate-warming warnings of a consensus of thousands of the world´s top climate scientists.

Not surprisingly, we also do not find that warnings of looming industrial and financial disasters have much impact. At Fukushima, the regulatory authorities required a seawall that was a bit taller than the largest tsunami that locale had experienced in the last 1,000 years. So the danger was, indeed, recognized. But the seawall design was based on probabilistic thinking, not thinking about what is possible, and the seawall was horribly inadequate to the 2011 tsunami.

Some Japanese experts had done possibility analysis. They pointed to historical records of a huge tsunami in the year 869; three huge tsunamis on the Pacific Ring of Fire, along which Japan lies, in the last 100 years; and a geological record of relentless collision between two tectonic plates underneath Japan. Before 2011, these experts were largely ignored.

Japan has 53 nuclear power plants drawing cooling water from the ocean. Before Fukushima, 14 lawsuits charging that risks had been ignored or hidden were filed in Japan, revealing a disturbing pattern in which operators underestimated or hid seismic dangers to avoid costly upgrades and keep operating. But all the lawsuits were unsuccessful.

A representative in the Japanese parliament in 2003 warned that the nuclear plants were not sufficiently protected; a seismology professor at Kobe University resigned in protest from a nuclear safety board in 2006 due to a lack of attention to earthquake and tsunami risks. Even though there had been a 30-foot tsunami in 1993 on Japan´s west coast from a much smaller 7.8 earthquake, the former head of the Toyko Electric Power Company (Tepco) said that the risk of a tsunami never crossed his mind when he was president of that firm. He obviously did not have a possibilistic mind set.

But warnings can be slippery and hard to use effectively, regardless of the attitudes of the people being warned. Too often warnings are imprecise, which was why Condoleezza Rice, as national security adviser at the time of 9/11, dismissed the warnings of a terrorist attack using airplanes. The warnings did not specify the time and place. Many warnings are, in fact, so general as to be useless, e.g., “nuclear power is dangerous” or “radical Moslems will strike the United States.”

There is the problem that warnings are often seen as mere obstructionism. This was the view of a representative for a Japanese utility who brushed away the possibility that two backup electrical generators would fail simultaneously. He said that worrying about such possibilities would “make it impossible to ever build anything.” Warnings may also be false, especially if based upon information that has little credibility, as with the weapons of mass destruction that Iraq was supposed to have. Many seemingly credible warnings never materialized, e.g., that President Barack Obama would not live through his first year in office. Florida coast residents are said to have stopped paying much attention to hurricane warnings after there were two evacuations for storms that didn´t make landfall in the state.

And to be sure, there are major accidents that occur without warning, including the Three Mile Island nuclear incident and some chemical plant accidents, such as the toxic releases from Union Carbide´s West Virginia plant. But these no-warning events are few. Credible warnings before major accidents are much more common.

The most credible ones are specific and in-house: The night before the launch of the space shuttle Challenger, engineers wanted to delay it because of the cold, saying, “We have never launched

at this temperature, and cold affects the O rings." Before the re-entry that burned up Columbia, a technician on the shuttle´s launch team tried to get pictures of the extent of the damage caused by chunks of insulation that had fallen off a fuel tank during the liftoff. Before the Deepwater Horizon was destroyed in a fiery blowout, Halliburton managers warned BP officials that there were not enough stabilizing rings installed on the drill pipe to continue drilling safely. Before BPÕs Prudhoe pipeline leaked, workmen placed hand-painted signs in the parts of the pipeline that did the most shaking, warning people to stand back because it might rupture (it did, but in an isolated area, and the break was not discovered for some days). Testimony has revealed that Massey managers regularly told supervisors to ignore warnings of dangerous concentrations of methane.
The warning at Texas City a few days before the plant blew was less specific than these others, but more ominous. At a company safety meeting, a slide was shown that simply said, "This is not a safe plant to work in." It was the view of management at the plant; they were unable to maintain the plant in safe conditions because of budget cuts and production pressures by top management.

Other warnings are more general and long range, but significant anyway. Scientists had regularly warned that the erosion of the wetlands protecting New Orleans was making it more vulnerable to hurricanes. They were more specific about the negative consequences of building a new ship canal that did, as predicted, channel Katrina´s storm surge directly into the city.

Scientists had regularly warned that the erosion of the wetlands protecting New Orleans was making it more vulnerable to hurricanes.

There were multiple warnings in the United States before the 2008 economic meltdown. They came from some directors of impacted firms and from many risk managers and department heads, worried about the risks of their highly profitable mortgage business. They also came from government regulatory agencies, including the Securities and Exchange Commission, and watchdog agencies such as the Government Accountability Office; from bills proposed in Congress; from chief executives of financial firms not at direct risk; from financial gurus; and from journalists at leading magazines such as the Economist. There were also some 7,000 news stories containing the phrase "the housing bubble" from 2000 to 2006, meaning there were seven years of warnings before that bubble burst late in 2007. Indeed, according to a book by a respected journalist (Sorkin, 2011), the newly appointed secretary of the treasury in the Bush administration, Henry Paulson, delivered warnings about a dangerous mortgage bubble at his first cabinet meeting in 2006. He proposed that investment banks be regulated much as commercial banks were, but Goldman Sachs, where Paulson had just served as president, and other major investment banks that dominated Wall Street would not hear of it. Credible warnings were dense, but the profits the firms were making drowned them out.

Coping

So how do organizations cope with disasters once one occurs? The record here is just as dismal as with regulation and warnings.

There are vastly more cases of creative coping from citizens than from organizations. The true first responders to disaster – co-workers, neighbors, passersby – have almost always performed splendidly, as with the completely self-organized flotilla that evacuated thousands from lower Manhattan in the 9/11 crisis. And in a few cases, governmental agencies and private firms do successfully cope with disaster.

Though failed space flights do not have the catastrophic potential of the other systems I have mentioned (they affect only a handful of people), they are complex and risky. The rescue of the crippled Apollo 13 capsule is a prime case of skill and innovation in dealing with and overcoming a failure. There are many outstanding examples of coping by airplane pilots, though, again, the potential loss

of life in these cases is relatively small. The response of President Lyndon Johnson´s administration to the 1964 Alaskan earthquake – at 9.2 magnitude, the biggest ever in North America – is a model of what can be done by government to help victims and rebuild a city. But examples of creative coping by organizations are rare.

The poster child for official failure to cope with disaster is the response to Hurricane Katrina, the subject of so many books and articles that I will not dwell upon it (although it should be noted that the US Coast Guard performed extremely well and the unofficial response by citizens and private firms was often innovative and effective). In the case of Fukushima, there was official denial and secrecy, refusal to accept outside help, the failure to evacuate citizens at risk, and an attempt by the prime minister to halt the cooling of the damaged reactors by seawater shortly after the process had been started (fortunately, the plant manager lied, saying he had stopped using seawater even as he continued to use the ocean to cool the crippled plants, thus preventing a far worse catastrophe).

At Bhopal, plant officials initially denied any chemical release and then said it was not dangerous, even as they themselves were fleeing upwind of the toxic fumes. The Soviet Union refused to admit there had been an accident at Chernobyl, even after the Swedish nuclear agency had concluded that the radioactive materials they were detecting had to come from Chernobyl. Worse yet, the USSR waited two days before evacuating the town next to the plant. BP officials and American officials consistently minimized the damage of the oil spill in the Gulf of Mexico and kept reporters and scientists away from the scene. Little of the equipment oil companies are required to have on hand in case of a big spill was present when the Exxon Valdez ran aground. Similarly, Massey Energy didn´t have equipment available to handle a mine explosion.

Crises may bring out the best in citizens. But, in some cases, they often raise the rate of routine errors made by distressed, tired managers and workers. At Fukushima, workers desperately trying to assemble a huge tank that would remove radioactive substances from the salt water being poured over the damaged reactor and its spent fuel pool saw the system fail on its first trial – because a valve had been installed backward. Workers at the Fort Calhoun Nuclear Power Plant near Omaha were surrounded by a flooding Missouri River with dikes close to being topped. They assembled an emergency berm to protect the vital electrical system. It was a 15-foot-wide, eight-foot-high plastic doughnut filled with water, a literal example of defense in depth. But someone backed a truck into it, and all the water poured out onto the soggy plant grounds.

This litany of regulatory failures, failures to heed warnings, and commonplace failures is independent of normal accident theory. That theory says that even if we had excellent regulation and everyone played it safe, there would still be accidents in systems that are highly “interactively complex,” and if the systems are tightly coupled, even small failures will cascade through them. The theory is useful for its emphasis on system complexity and tight coupling; these concepts play a huge role in analyzing the failures of any source in risky systems. In the financial meltdown, for example, the mounting complexity of the overall system allowed fraud and self-dealing to go undetected, and the tight coupling of many systems allowed the failures to cascade.

In my work on “normal accidents”, I have argued that some complex organizations -such as chemical plants, nuclear power plants, nuclear weapons systems, and, to a more limited extent, air transport networks- have so many nonlinear system properties that eventually the unanticipated interaction of multiple failures may create an accident that no designer could have anticipated and no operator can understand.

Everything is subject to failure -designs, procedures, supplies and equipment, operators, and the environment. The government and businesses know this and design safety devices with multiple redundancies and all kinds of bells and whistles. But nonlinear, unexpected interactions of even small failures can defeat these safety systems. If the system is also tightly coupled, no intervention can prevent a cascade of failures that brings it down.

I use the term "normal" because these characteristics are built into the systems; there is nothing one can do about them other than to initiate massive system redesigns to reduce interactive complexity and to loosen coupling. Companies and governments can modularize integrated designs and deconcentrate hazardous material. Actually, though, compared with the prosaic cases previously mentioned, normal accidents are rare. (Three Mile Island is the only accident in my list that qualifies.) It is much more common for systems with catastrophic potential to fail because of poor regulation, ignored warnings, production pressures, cost cutting, poor training, and so on.

All of the organizational faults I have noted have their counterpart in daily life. Like organizations and their leaders, people seek wealth and prestige and a reputation for integrity. In the process, they occasionally find it necessary to be deceitful, engaging in denials and cover-ups, cheating and fabrication. Everyone has violated regulations, failed to plan ahead, and bungled in crises. But people are not, as individuals, repositories of radioactive materials, toxic substances, and explosives, nor do they sit astride critical infrastructures. Organizations do. The consequences of an individual's failures can only be catastrophic if they are magnified by organizations. The larger the organizations, the greater the concentration of destructive power. The larger the organizations, the greater the potential for political power that can influence regulations and ignore warnings.

Modern society is not likely to deconcentrate big organizations and toxic substances, so what can be done? High-reliability theory is correct, of course, to say that government and business can do much more than they do to prevent serious accidents through constant training and mindfulness. More important is system design: Modular systems are less vulnerable than integrated ones, and the toxic and explosive potential is more dispersed in modular systems than in tightly coupled ones. Even more can be done through regulation; highly regulated nuclear power plants in Europe do much better than poorly regulated ones in the United States and Japan. Technical improvements can make systems safer, of course, and we do learn from past disasters. Emergency power facilities are being upgraded at nuclear plants in the United States because of Fukushima.

High-reliability theory is correct, of course, to say that government and business can do much more than they do to prevent serious accidents through constant training and mindfulness.

And the learning is continuous. Before the Three Mile Island incident, some held that during a loss-of-coolant accident in a nuclear plant, there would be a possibility of a zirconium-water reaction that consumes oxygen and frees hydrogen, which is explosive. One nuclear scientist scoffed at such a possibility in a publication, which was released shortly before he was, unfortunately, designated to be the key scientific adviser to Pennsylvania Governor Richard Thornburg during the Three Mile Island accident. (Later, the scientist was appointed chairman of the NRC.) The scientist was of course wrong. The appearance of hydrogen meant there was hydrogen "burn," as it is called, at Three Mile Island. Fortunately, the hydrogen accumulation was small, and the damage was minimal. With this march of knowledge, engineers learned to install vents to prevent the explosive accumulation of hydrogen in the reactor buildings of nuclear plants in case of a loss-of-coolant accident. But the vents failed at Fukushima. There was no power to open them, and the radiation levels were so high workers could not get to them to open them manually. The hydrogen explosions sent radioactive materials and gasses into the environment.

Don't despair, though. Learning from disaster still goes on: US plants have been asked to make sure their vents can be opened even manually in order to prevent a hydrogen explosion!

The US, however, has not learned about the dangers of storing spent fuel rods in a building adjacent to the containment building as we do in 34 of our nuclear power plants.

Presently, in Fukushima, fuel rods with 10 times the cesium 137 released at Chernobyl sit in a cooling pool 100 feet into the air in a damaged building that has lost its roof because of an explosion and is further damaged by the earthquake. Japanese authorities estimate it will take at least four years to remove the rods. But the building could collapse on its own or in a storm before then, and more earthquakes in the area this year is almost a certainty. One could bring down the pool.

Collapse of the pool would require the evacuation of Tokyo. The resulting meltdown of the uncooled rods could make access to a much larger storage site about 50 yards away impossible, releasing a total of 80 times as much of the deadly cesium as released at Chernobyl. This would endanger the whole planet. It would be too late for nuclear regulators to learn this lesson, and it might not matter. Cooling of the 405 plant cores and spent fuel rods worldwide, even if they could be quickly shutdown, might not be possible because of the radiological incapacitation or death of qualified personnel.

Prosaic organizational failures will always be with us, and knowledge is always incomplete or in dispute. Even highly reliable systems are subject to everyday failures, and even if we avoid these, there is always the possibility of normal accidents -rare but inevitable in interactively complex, tightly coupled systems. Some complex systems with catastrophic potential are just too dangerous to exist, not because we do not want to make them safe, but because, as so much experience has shown, we simply cannot.

References:

Mangels, J. (2003). NRC cracks down: Industry strikes back. Cleveland Plain Dealer, June 25.

Perrow, C. (2011): The Next Catastrophe: Reducing *Our Vulnerabilities to Natural, Industrial, and Terrorist Disasters*. Princeton, NJ: Princeton University Press.

Sorkin, A. (2011): *Too Big to Fail.* New York: Penguin.

Das Thema Komplexität ist in dieser Ausgabe ein Schlüsselthema. Der immer höhere Komplexitätsgrad führt zu der Notwendigkeit, neue Modelle und Theorien zu formulieren und alte Pfade zu verlassen. Ein möglicher Zugang zur Komplexität bietet die sogenannte Komplexitätstheorie. In dem vorliegenden Artikel, der uns freundlicherweise von Prof. Dekker zu Verfügung gestellt wurde, geht es ganz konkret um die Implikationen der Komplexitätsforschung auf die Vorfalluntersuchung. Es handelt sich hier um einen Artikel, den Prof. Dekker zusammen mit dem renommierten Komplexitätsforscher Paul Cilliers (Stellenbosch-Universität, Südafrika) und Jan-Hendrik Hofmeyr verfasst hat. Eine umfassende Behandlung der Thematik Komplexität findet sich im 2011 erschienenen Buch „Drift into Failure", das ebenfalls von Prof. Dekker geschrieben wurde und dem geneigten Leser ans Herz gelegt wird.

Sidney Dekker. Professor in the School of Humanities at Griffith University in Brisbane, Australia. Previously Professor at Lund University, Sweden and Director of the Leonardo Da Vinci Center for Complexity and Systems Thinking. Author of several books on system failure and human error, e.g. "Behind Human Error" (2010) and "Drift into Failure" (2011).

The complexity of failure: Implications of complexity theory for safety investigations

By Sidney Dekker

1. Introduction

Complexity is a defining characteristic of today's high-technology, high-consequence systems (Perrow 1984), and recent submissions to this journal highlight the importance of taking a systems- or complexity view of failure in such systems (Goh, Brown et al. 2010). Despite this, single-factor explanations that condense accounts of failure to individual human action or inaction often prevail. An analysis by Holden (Holden 2009) showed that between 1999 and 2006, 96% of investigated US aviation accidents were attributed in large part to the flight crew. In 81%, people were the sole reported cause. The language used in these analyses also points to failures or shortcomings of components. "Crew failure" or a similar term appeared in 74% of probable causes; 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 (Holden 2009).

Single-factor, judgmental explanations for complex system failures are not unique to aviation – they are prevalent in fields from medicine (Wachter and Pronovost 2009), to military operations (e.g. Snook, 2000), to road traffic (Tingvall and Lie 2010). Much discourse around accidents 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?" (Douglas 1992; Catino 2008; Cook and Nemeth 2010). The problem of reverting to condensed, single-factor explanations rather than diffuse and system-level ones (Galison 2000) has of course been a central preoccupation of system safety (Reason 1997; Maurino, Reason et al. 1999; Dekker 2006), and the difficulty in achieving systemic stories of failure has been considered from a variety of angles (Fischhoff 1975; Woods, Dekker et al. 2010).

This paper adds to the literature by critiquing the traditional philosophical-historical and ideological bases for sustained linear thinking about failure in complex systems. We mean by linear thinking a process that follows a chain of causal reasoning from a premise to a single outcome. In contrast, systems thinking regards an outcome as emerging from a complex network of causal interactions, and, therefore, not the result of a single factor (Leveson 2002). We lay out how a Newtonian analysis of failure makes particular assumptions about the relationship between cause and effect, foreseeabil-

ity of harm, time-reversibility and the ability to come up with the "true story" of an accident. Whereas such thinking has long been equated with science and rationality in the West, an acknowledgement of the complex, systemic nature of many accidents necessitates a different approach. We take steps toward the development of such an approach, of what we like to call a post-Newtonian analysis, in the second half.

2. The Cartesian-Newtonian worldview and its implications for system safety

The logic behind Newtonian science is easy to formulate, although its implications for how we think about accidents are subtle and pervasive. Classical mechanics, as formulated by Newton and further developed by Laplace and others encourages a reductionist, mechanistic methodology and worldview. Many still equate "scientific thinking" with "Newtonian thinking." The mechanistic paradigm is compelling in its simplicity, coherence and apparent completeness and largely consistent with intuition and common sense.

2.1. Reductionism and the eureka part

The best known principle of Newtonian science, formulated well before Newton by the philosopher-scientist Descartes, is that of analysis or reductionism. 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 – is very common (Galison 2000). The investigators of the Trans World Airlines 200 crash off New York called it their search for the "eureka part": the part that would have everybody in the investigation declare that the broken component, the trigger, the original culprit, had been located and could carry the explanatory load of the loss of the entire Boeing 747. But for this crash, the so-called "eureka part" was never found (Langewiesche 1998).

The defenses-in-depth metaphor (Reason 1990; Hollnagel 2004) relies on the linear parsing-up of a system to locate broken layers or parts. This was recently used by BP in its own analysis of the Deepwater Horizon accident and subsequent Gulf of Mexico oil spill (BP 2010), seen by others as an effort to divert liability onto multiple participants in the construction of the well and the platform (Levin 2010).

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 analyze phenomena into their basic components. This is applied in the search for psychological sources of failure, for example. Methods that subdivide "human error" into further component categories, such as perceptual failure, attention failure, memory failure or inaction are in use in for example air traffic control (Hollnagel and Amalberti 2001). It is also applied in legal reasoning in the wake of accidents, by separating out one or a few actions (or inactions) on the part of individual people (Douglas 1992; Thomas 2007; Catino 2008).

2.2. Causes for effects can be found

In the Newtonian vision of the world, everything that happens has a definitive, identifiable cause and a definitive effect. There is 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, but assumes that physical effects can be traced back to physical causes (or a chain of causes-effects) (Leveson 2002). The assumption that effects cannot occur without specific causes influences legal reasoning in the wake of accidents too. For example, to raise a question of negligence in an accident, harm must be caused by the negligent action (GAIN 2004). Assumptions

about cause-effect symmetry can be seen in what is known as the outcome bias (Fischhoff 1975). The worse the consequences, the more any preceding acts are seen as blameworthy (Hugh and Dekker 2009).

Newtonian ontology is materialistic: all phenomena, whether physical, psychological or social, can be reduced to (or understood in terms of) matter, that is, 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. The Newtonian model may have become so pervasive and coincident with "scientific" thinking that, if analytic reduction to determinate cause-effect relationships cannot be achieved, then the accident analysis method or agency isn't entirely worthy. The Chairman of the NTSB at the time, Jim Hall, raised the specter of his agency not being able to find the Eureka part in TWA 200, which would challenge its entire reputation (p. 119): "What you're dealing with here is much more than an aviation accident… What you have at stake here is the credibility of this agency and the credibility of the government to run an investigation." (Dekker 2011)

If such knowledge is in principle attainable, then harmful outcomes are foreseeable too.

2.3. 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 momentum of the components constituting a system, as well as the forces acting on those components (which are not only external forces but also those 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.

If such knowledge is in principle attainable, then harmful outcomes are foreseeable too. Where people have a duty of care 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 (i.e. were they incompetent, unknowledgeable)? Did they fail to plot out the possible effects of their actions? Indeed, legal rationality in the determination of negligence follows this feature of the Newtonian model (p. 6): "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" (GAIN 2004).

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 based on this Newtonian commitment to foreseeability too. For example, a road traffic law in a typical Western country might specify how a motorist should adjust speed so as to be able stop the vehicle before colliding with some obstacle, at the same time remaining aware of the circumstances that could influence such selection of speed. Both the foreseeability of all possible hindrances and the awareness of circumstances (initial conditions) critical for determining speed are steeped in Newtonian epistemology. Both are also heavily subject to outcome bias: if an accident suggests that an

obstacle or particular circumstance was not foreseen, then speed was surely too high. The system's user, as a consequence, is always wrong (Tingvall & Lie, 2010).

2.4. 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. Such assumptions give accident 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 reconstruction reaffirms and instantiates Newtonian physics: knowledge about past events is not original, but merely the result of uncovering a pre-existing order. The only thing between an investigator and a good reconstruction are the limits on the accuracy of the representation of what happened. It follows that accuracy can be improved by "better" methods of investigation (Shappell and Wiegmann 2001).

Knowledge about past events is not original, but merely the result of uncovering a pre-existing order.

2.5. Completeness of knowledge

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) and it was the task of the investigator to discover this hidden order underneath the apparent disorder (Feyerabend 1993). It follows that the more facts an analyst or investigator collects, the more it leads, inevitably, to a better investigation: a better representation of "what happened." In the limit, this can lead to a perfect, objective representation of the world outside (Heylighen 1999), or one final (true) story, the one in which there is no gap between external events and their internal representation.

Those equipped with better methods, and particularly those who enjoy greater "objectivity," (i.e. those who have no bias which distorts their perception of the world, and who will consider all the facts) are better positioned to construct such a true story. Formal, government-sponsored accident investigations can sometimes enjoy this idea of objectivity and truth – if not in the substance of the story they produce, then at least in the institutional arrangements surrounding its production. Putative objectivity can be deliberately engineered into the investigation as a sum of subjectivities: all interested parties (e.g. vendors, the industry, operator, unions, professional associations) can officially contribute (though dissenting voices can be silenced or sidelined (Perrow 1984; Byrne 2002)). Other parties often wait until a formal report is produced before publicly taking either position or action, legitimizing the accident investigation as original arbitrator between fact and fiction.

2.6. Newtonian bastardization of the response to failure in complex systems

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

- To understand a failure of a system, investigators 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 (e.g. the more egregious) the cause must have been.
- If they put in more effort, people can more reliably foresee outcomes. After all, they would have a better understanding of the starting conditions and they 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, harmful states be foreseen and 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.
- One official account of what happened is possible and desirable. 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.

These assumptions can remain largely transparent and closed to critique in safety work precisely because they are so self-evident and commonsensical. People involved in system safety may be 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. 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 reasoning and the identification of malfunctioning components (Shappell and Wiegmann 2001). Also, organizations and those held accountable for internal failures need something to "fix," which further valorizes condensed accounts. If these processes fail to satisfy societal accountability requirements, then courts can decide to pursue individuals criminally, which also could be said to represent a hunt for a broken component (Thomas 2007; Dekker 2009). A socio-technical Newtonian physics is thus often read into events that could yield much more complexly patterned interpretations.

3. The view from complexity and its implications for system failure

Analytic reduction cannot tell how a number of different things and processes act together when exposed to a number of different influences at the same time. This is complexity, a characteristic of a system. Complex behaviour arises because of the interaction between the components of a system. It asks us to focus not on individual components but on their relationships. The properties of the system emerge as a result of these interactions; they are not contained within individual components. Complex systems generate new structures internally, they are not reliant on an external designer. In reaction to changing conditions in the environment, the system has to adjust some of its internal structure. Complexity is a feature of the system, not of components inside of 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. The behavior of the system cannot be reduced to the behavior of the constituent components. If we wish to study such systems, we have to investigate the system as such. It is at this point that reductionist methods fail.

3.1. Complicated versus complex

There is an important distinction between "complex" and "complicated" systems. Certain systems may be quite intricate and consist of a huge number of parts, e.g. a jet airliner. Nevertheless, it can be taken apart and put together again. Even if such a system cannot practically be understood completely by a single person, it is understandable and describable in principle. This makes them complicated. Complex systems, on the other hand, come to be in the interaction of the components. Jet airliners become complex systems when they are deployed in a nominally regulated world with cultural diversity, receiver-oriented versus transmitter-oriented communication expectations, different hierarchical gradients in a cockpit and multiple levels of politeness differentiation (Orasanu and Martin 1998), effects of fatigue, procedural drift (Snook 2000), varied training and language standards

(Hutchins, Holder et al. 2002), as well as cross-cultural differences in risk perceptions, attitudes and behavior (Lund and Rundmo 2009). This is where complicated systems become complex because they are opened up to influences that lie way beyond engineering specifications and reliability predictions.

Complex systems are held together by local relationships only. Each component is ignorant of the behavior of the system as a whole, and cannot know the full influences of its actions. Components respond locally to information presented to them, and complexity arises from the huge, multiplied webs of relationships and interactions that result from these local actions. The boundaries of what constitutes the system become fuzzy; interdependencies and interactions multiply and mushroom.

Complex systems have a history, a path-dependence, which also spills over those fuzzy boundaries. Their past, and the past of events around them, is co-responsible for their present behavior, and descriptions of complexity should take history into account. Take as an example the take-off accident of SQ006 at Taipei where deficient runway and taxiway lighting/signage was implicated in the crew's selection of a runway that was actually closed for traffic (ASW 2002). A critical taxiway light had burnt out, and the story could stop there. But as only one of two countries in the world (the Vatican being the other), Taiwan is not bound by airport design rules from the International Civil Aviation Organization, a United Nations body. Cross-strait relationships between Taiwan (formerly Formosa) and China, and their tortured post-WWII history (Mao's communists versus Chang-Kai-Shek's Kuomintang Nationalists) and their place in a larger global context are responsible for keeping Taiwan out of the UN and thus out of formal reach for international safety regulations that apply to even the smallest markings and lights and their maintenance on an airport. A broken light and exceptionally bad weather on the night of the SQ006 take-off constituted the small changes in starting conditions that led to a large event in a system that had been living far from equilibrium for a long time – all consistent with complexity.

A critical taxiway light had burnt out, and the story could stop there.

In describing the macro-behaviour (or emergent behaviour) of the system, not all its micro-features can be taken into account. The description on the macro-level is thus a reduction or compression of complexity, and cannot be an exact description of what the system actually does. Moreover, the emergent properties on the macro-level can influence the micro-activities, a phenomenon sometimes referred to as "top-down causation." Nevertheless, macro-behaviour is not the result of anything else but the micro-activities of the system, keeping in mind that these are not only influenced by their mutual interaction and by top-down effects, but also by the interaction of the system with its environment. These insights have important implications for the knowledge-claims made when analyzing accidents in complex systems. Since we do not have direct access to the complexity itself, our knowledge of such systems is in principle limited. What follows is an outline of the implications of these insights for (the ethics of) system failure, revisiting to the topics used in the section on Newtonian science.

3.2. Synthesis and holism

Perrow pointed out how Newtonian focus on parts as the cause of accidents may sponsored a false belief that redundancy is the best way to protect against hazard (Perrow 1984; Sagan 1993). The downside is that barriers, as well as professional specialization, policies, procedures, protocols, redundant mechanisms and structures, all add to a system's complexity. They entail an explosion of new relationships (between parts and layers and components) that spread through the system. System accidents result from the relationships between components, not from the workings or malfunctioning of any component part. This insight had already grown out of systems engineering

and system safety, pioneered in part by 1950's aerospace engineers who were confronted with increasing complexity in aircraft and ballistic missile systems at that time. The interconnectivity and interactivity between system components made that greater complexity led to vastly more possible interactions than could be planned, understood, anticipated or guarded against (Leveson 2002).

Newtonian analytic reduction of complex systems not only fails to reveal any culprit part (as broken parts are not necessary to produce system failures) but would also eliminate the phenomenon of interest – the interactive complexity of the system itself that gives rise to conditions that help produce an accident. Sequence-of-events and barrier models can say nothing about the build-up of latent failures, about a gradual, incremental loosening or loss of control that characterizes system accidents (Rasmussen 1997). The processes of erosion of constraints, of attrition of safety, of drift towards margins, cannot be captured because reductive approaches are static metaphors for resulting forms, not dynamic models oriented at processes of the formation, the evolution of relationships.

3.3. Emergence

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 (Leveson 2002). Accidents have similarly been characterized as emergent properties of complex systems (Hollnagel 2004). They cannot be predicted on the basis of the constituent parts. Rather, they are one emergent feature of constituent components doing their (normal) work. A systems accident is possible in an organization where people themselves suffer no noteworthy incidents, in which everything looks normal, and everybody is abiding by their local rules, common solutions, or habits (Vaughan 2005). This means that the behavior of the whole cannot be explained by, and is not mirrored in, the behavior of constituent components. Snook (Snook 2000) expressed the realization that bad effects can happen with no causes in his study of the shoot-down of two U.S. Black Hawk helicopters by two U.S. fighter jets in the no-fly zone over Northern Iraq in 1993:

“This journey played with my emotions. When I first examined the data, I went in puzzled, angry, and disappointed – puzzled how two highly trained Air Force pilots could make such a deadly mistake; angry at how an entire crew of AWACS controllers could sit by and watch a tragedy develop without taking action; and disappointed at how dysfunctional Task Force OPC must have been to have not better integrated helicopters into its air operations. Each time I went in hot and suspicious. Each time I came out sympathetic and unnerved... If no one did anything wrong; if there were no unexplainable surprises at any level of analysis; if nothing was abnormal from a behavioral and organizational perspective; then what…?”

Snook’s impulse to hunt down the broken components (deadly pilot error, controllers sitting by, a dysfunctional Task Force) led to nothing. There was no “eureka part.” Again, an accident defied Newtonian logic.

Asymmetry or non-linearity means that an infinitesimal change in starting conditions can lead to huge differences later on. This sensitive dependence on initial conditions removes proportionality from the relationships between system inputs and outputs. The evaluation of damage caused by debris falling off the external tank prior to the fatal 2003 Space Shuttle Columbia flight can serve as an example (CAIB 2003; Starbuck and Farjoun 2005). Always under pressure to accommodate tight launch schedules and budget cuts (in part because of a diversion of funds to the international space station), certain problems became seen as maintenance issues rather than flight safety risks. Maintenance issues could be cleared through a nominally simpler bureaucratic process, which allowed quicker Shuttle vehicle turnarounds. In the mass of assessments to be made between flights, the effect of foam debris strikes was one. 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 – one more decision, just like tens of thousands of other

decisions. While any such decision can be quite rational given the local circumstances and the goals, knowledge and attention of the decision makers, interactive complexity of the system can take it onto unpredictable pathways to hard-to-foresee system outcomes.

This complexity has implications for the ethical load distribution in the aftermath of complex system failure. Consequences cannot form the basis for an assessment of the gravity of the cause (or the quality of the decision leading up to it), something that has been argued in the safety and human factors literature (Orasanu and Martin 1998). It suggests that everyday organizational decisions, embedded in masses of similar decisions and only subject to special consideration with the wisdom of hindsight, cannot be fairly singled out for purposes of exacting accountability (e.g. through criminalization) because their relationship to the eventual outcome is complex, non-linear, and was probably impossible to foresee (Jensen 1996).

3.4. Foreseeability of probabilities, not certainties

Decision makers in complex systems are capable of assessing the probabilities, but not the certainties of particular outcomes (Orasanu and Martin 1998). With an outcome in hand, its foreseeability becomes quite obvious, and it may appear as if a decision in fact determined an outcome, that it inevitably led up to it (Fischhoff and Beyth 1975). But knowledge of initial conditions and total knowledge of the laws governing a system (the two Newtonian conditions for assessing foreseeability of harm) is unobtainable in complex systems.

That does not mean that such decisions are not singled out in retrospective analyses. That they do is but one consequence of Newtonian thinking: accidents have typically been modeled as a chain of events. While a particular historical decision can be cast as an "event," it becomes very difficult to locate the immediately preceding "event" that was its cause. So the decision (the human error, or "violation") is cast as the aboriginal cause, the root cause (Leveson 2002).

3.5. Time-irreversibility

The conditions of a complex system are irreversible. The precise set of conditions that gave rise to the emergence of a particular outcome (e.g. an accident) is something that can never be exhaustively reconstructed. Complex systems continually experience change as relationships and connections evolve internally and adapt to their changing environment. Given the open, adaptive nature of complex systems, the system after the accident is not the same as the system before the accident – many things will have changed, not only as a result of the outcome, but as a result of the passage of time.

This also means that the any predictive power of retrospective analysis of failure is limited (Leveson 2002). Decisions in organizations, for example, to the extent that they can be described separately from context at all, are not the single beads strung along some linear cause-effect sequence that they may seem afterward. Complexity argues that they are spawned and suspended in the messy interior of organizational life that influences and buffets and shapes them in a multitude of ways. Many of these ways are hard to trace retrospectively as they do not follow documented organizational protocol but rather depend on unwritten routines, implicit expectations, professional judgments and subtle oral influences on what people deem rational or doable in any given situation (Vaughan 1999). Reconstructing events in a complex system, then, is impossible, primarily as a result of the characteristics of complexity. The system that is subjected to scrutiny after the fact is never the same system that produced the outcome. It will already have changed, partly because of the outcome, and partly because of passing time and the nature of complexity. But psychological characteristics

of retrospective investigation make it so too. As soon as an outcome has happened, whatever past events can be said to have led up to it, undergo a whole range of transformations (Fischhoff and Beyth 1975; Hugh and Dekker 2009). Take the idea that it is a sequence of events that precedes an accident. Who makes the selection of the "events" and on the basis of what? The very act of separating important or contributory events from unimportant ones is an act of construction, of the creation of a story, not the reconstruction of a story that was already there, ready to be uncovered. Any sequence of events or list of contributory or causal factors already smuggles a whole array of selection mechanisms and criteria into the supposed "re"-construction. There is no objective way of doing this – all these choices are affected, more or less tacitly, by the analyst's background, preferences, experiences, biases, beliefs and purposes. "Events" are themselves defined and delimited by the stories with which the analyst configures them, and are impossible to imagine outside this selective, exclusionary, narrative fore-structure (Cronon 1992).

3.6. Perpetual incompleteness and uncertainty of knowledge

The Newtonian belief that is both instantiated and reproduced in official accident investigations is that there is a world that is objectively available and apprehensible. An independent world exists to which investigators, with proper methods, can gain objective access. Observer and the observed are separable, knowledge is a mapping of facts from one to the other. Investigation, in this belief, is not a constitutive or creative process: it is merely an "uncovering" of distinctions that were already there and that are simply waiting to be observed (Heylighen, Cilliers et al. 2006).

Complexity, in contrast, suggests that the observer is not just the contributor to, but in many cases the creator of, the observed (Wallerstein 1996). Cybernetics acknowledged the intrinsically subjective nature of knowledge, as an imperfect tool used by an intelligent agent to help it achieve its goals. Not only does the agent not need an objective mapping of reality, it can actually 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 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 (Heylighen, Cilliers et al. 2006).

Different descriptions of a complex system, then (from the point of view of different agents), decompose the system in different ways. It follows that the knowledge gained by any description is always relative to the perspective from which the description was made. This does not imply that any description is as good as any other. It is merely the result of the fact that only a limited number of characteristics of the system can be taken into account by any specific description. It is not that some complex readings are "truer" in the sense of corresponding more closely to some objective state of affairs (as that would be a Newtonian commitment). Rather, the acknowledgement of complexity in safety work can lead to a richer understanding and thus it holds the potential to improve safety and help to expand the ethical response in the aftermath of failure. Examples of alternative accounts of high-visibility accidents that all challenged the "official" readings include the DC-10 Mount Erebus crash (Vette 1984), the Dryden Fokker F-28 icing accident (Maurino, Reason et al. 1999), the Space Shuttle Challenger launch decision (Vaughan 1996), and the fratricide of two Black Hawk helicopters over Northern Iraq (Snook 2000). All such revisionist accounts, generated by a diversity of academics, system insiders, and even a judge, have offered various audiences the opportunity to embrace a greater richness of voices and interpretations. Together, they better acknowledge the complexity of the events, and the systems in which they happened, than any single account could. In complex-

ity there is no procedure for deciding which narrative is correct, even though some descriptions will deliver more interesting results than others – depending of course on the goals of the audience. The selection of "causes" (or "events" or "contributory factors") is always an act of construction by the investigation or a revisionist narrative. There is no objective way of doing this – all analytical choices are affected, more or less explicitly, by the author's own position in a complex system. The most interesting "truth" about an accident, then, may lie in the diversity of possible accounts, not in the coerciveness of a single one (Hoven 2001; Cilliers 2010).

3.7. A post-Newtonian analysis of failure in complex systems

Complexity and systems thinking denies the existence of one objective reality that is accessible with accurate methods. This has implications for what can be considered useful and ethical in the aftermath of failure (Cilliers 2005):

- An investigation must gather as much information on the event as possible, notwithstanding the fact that it is impossible to gather "all" the information.
- An investigation can never uncover one true story of what happened. That people have different accounts of what happened in the aftermath of failure should not be seen as somebody being right and somebody being wrong. It may be more ethical to aim for diversity, and respect otherness and difference in accounts about what happened as a value in itself. Diversity of narrative can be seen as an enormous source of resilience in complex systems, not as a weakness. The more angles, the more there can be to learn.
- An investigation must consider as many of the possible consequences of any finding, conclusion or recommendation in the aftermath of failure, notwithstanding the fact that it is impossible to consider all the consequences.
- An investigation should make sure that it is possible to revise any conclusion as soon as it becomes clear that it has flaws, notwithstanding the fact that the conditions of a complex system are irreversible. Even when a conclusion is reversed, some of its consequences (psychological, practical) may remain irreversible.

4. Conclusion

When accidents are seen as complex phenomena, there is no longer an obvious relationship between the behavior of parts in the system (or their malfunctioning, e.g. "human errors") and system-level outcomes. Instead, system-level behaviors emerge from the multitude of relationship and interconnections deeper inside the system, and cannot be reduced to those relationships or interconnections. Investigations that embrace complexity, then, might stop looking for the "causes" of failure or success. Instead, they gather multiple narratives from different perspectives inside of the complex system, which give partially overlapping and partially contradictory accounts of how emergent outcomes come about. The complexity perspective dispenses with the notion that there are easy answers to a complex systems event – supposedly within reach of the one with the best method or most objective investigative viewpoint. It allows us to invite more voices into the conversation, and to celebrate their diversity and contributions.

-> Aufgrund der besseren Übersicht haben wir uns entschieden, die enorme Anzahl an Literaturangaben in diesem Text nicht abzudrucken. Der interessierte Leser sei auf die von Prof. Dekker bereits erschienenen Bücher verwiesen

Nancy Leveson. Professor of Aeronautics and Astronautics and Professor of Engineering Systems at Massachusetts Institute of Technology (MIT). She is an elected member of the National Academy of Engineering (NAE). Prof. Leveson conducts research on the topics of system safety, software safety, software and system engineering, and human-computer interaction. She has published over 200 research papers and is author of two books, "Safeware: System Safety and Computers" published in 1995 by Addison-Wesley and "Engineering a Safer World" published in 2012 by MIT Press.

Nancy Leveson ist eine weitere herausragende Wissenschaftlerin, die sich intensiv mit dem Thema Komplexität beschäftigt hat. Ihr Renommee auf dem Gebiet der Sicherheitsforschung wird unter anderem durch ihr Engagement als Beraterin im Untersuchungsgremium zum Absturz der Raumfähre Columbia und der Ölkatastrophe „Deepwater Horizon" unterstrichen. Sie formulierte ein neues Unfallmodell, das eine Abkehr traditioneller Annahmen bedeutet. In dem von ihr formulierten STAMP (System-Theoretic Accident Model and Processes) werden Unfälle als dynamisches Kontrollproblem angesehen. Es betrachtet das gesamte sozio-technische System mit all seinen Interaktionen statt den Blick isoliert auf fehlerhafte Komponenten zu richten. Prof. Leveson hat für diese Ausgabe in Zusammenarbeit mit Cody Fleming, John Tomas (beides Mitarbeiter am MIT) und Chris Wilkinson (Honeywell) einen wissenschaftlichen Beitrag zusammengestellt, der anhand eines konkreten Beispiels der traditionellen Herangehensweise zur Sicherheitsbetrachtung ihre neue gegenüberstellt.

Safety assurance of changes to the air transport system

By Nancy Leveson

1. Introduction

Both the U.S. and Europe are developing complex changes to their respective air transportation systems, respectively called NextGen and SESAR. Because the research described in this paper was supported by NASA's Aviation Safety Program, which is focused on NextGen, a NextGen example is used. The approaches described, however, apply equally to SESAR.

The NextGen plan is to change the system through an evolution from a ground-based air traffic control system to a satellite-based system of air traffic management [1]. The overarching goals of NextGen (which are shared with SESAR) are to 1) reduce flight delays by improving airport operations; 2) improve aviation's impact on the environment through reduced CO2 emissions and fuel use; and 3) make the airspace safer via more precise tracking, improved information-sharing, and implementing a Safety Management System [2].

As changes are designed and implemented to realize these goals, assurance is necessary that the current high level of safety will not be degraded. The complexity of the current system and the changes envisioned makes this process challenging. Powerful tools will be required to assure aircraft and airspace safety.

Traditional approaches to safety analysis assume that accidents are caused by component failures [3, 4]. They therefore focus on reliability analysis techniques, particularly fault tree or event tree analysis. The goal is to determine scenarios of component failures that together will lead to an accident or loss event. Failures may be single or multiple and are usually assumed to be random. After the component failure scenarios are identified, engineers use fault tolerance or fail-safe techniques to protect against hazards caused by the identified failures and to increase individual component integrity. A fly-fix-fly approach augments the design techniques with investigation of accidents and potentially serious incidents in great depth and recommendations made from the results to prevent reoccurrences.

This approach has been very effective in the past because there have been relatively few changes in the basic aircraft or air traffic control design; the systems are relatively simple; technology has changed slowly; engineers have been able to use very conservative design approaches; and the sys-

tem components can be effectively decoupled so that interactions can be anticipated, simplified, and guarded against. This approach, by itself, is becoming less effective, however, as these assumptions start to be violated in new next generation ATM systems.

Software is increasingly an important part of systems and allows enormously more complex and tightly coupled systems to be constructed. The potential for accidents arising from unsafe interactions among non-failed components, i.e., unplanned system and software behavior, is increasing. NextGen components, for example, may involve more than just one aircraft and one onboard system and include multiple aircraft, ground controllers, space-based systems, and communication links between aircraft. The traditional hardware-oriented safety engineering techniques focusing on failures do not adequately handle these types of new accident causes.

In addition, human roles are changing from direct control to supervision of automation, which requires more cognitively complex human decision-making. Like software, the changing roles of pilots and ground controllers introduces the potential for new causes of accidents that are not well handled by today's failure-oriented and hardware-oriented approaches.

To deal with these new accident causes, more powerful tools are needed. This paper describes and demonstrates a new approach to safety analysis based on systems and control theory rather than reliability theory. Safety is treated as a control problem rather than a "prevent failures" problem, allowing not only consideration of the causes of the component failure accidents that were predominant in the past but also the new causality factors that are increasingly important today. This approach can be applied to NextGen and to other upgrades to complex transportation systems.

To demonstrate and evaluate the approach, we use a new ATC procedure, called ATSA-ITP (Airborne Traffic Situational Awareness In-Trail Procedure). ATSA-ITP, or simply ITP, was chosen because the safety analysis and underlying methodology has already been documented in DO-312 [5]. ITP provides a real-world case study with which to compare our safety assurance philosophy and analytical techniques being proposed to those being used by the FAA, EUROCONTROL, and their associated organizations.

2. Using STAMP and STPA for Safety Assurance

The significant technical changes envisioned for NextGen create a necessity for a new, more powerful model of accident causality that better represents today's complex, socio-technical systems. The new model used in our analysis, called STAMP (Systems-Theoretic Accident Model and Processes) [6], extends the types of accidents and causes that can be considered by including non-linear, indirect, and feedback relationships among events. In this way, the traditional causality model is extended to consider new types of accident causes arising from component interactions (rather than just component failures), cognitively complex human mistakes, management and organizational errors, software errors (particularly requirements errors), etc. Accidents or unacceptable losses can result not only from system component failures but also from interactions among system components – both physical and social – that violate system safety constraints.

In systems theory, emergent properties (like safety) associated with a set of components are related to constraints upon the degree of freedom of those components' behavior [7]. System safety, then, can be reformulated as a system control problem rather than a component reliability problem: accidents or losses occur when component failures, external disturbances, and/or dysfunctional interactions among system components are not handled adequately or controlled – where controls may be managerial, organizational, physical, or operational.

In a systems-theoretic view of safety, the emergent safety properties are controlled or enforced by a set of safety constraints related to the behavior of the system components. Safety constraints specify those relationships among system variables or components that constitute the non-hazardous or safe system states: for example, the power must never be on when the access door to the high-power source is open; two aircraft must never violate minimum separation requirements; pilots in a combat zone must be able to identify targets as hostile or friendly; and the public health system must prevent the exposure of the public to contaminated water and food products. Accidents result from interactions among system components that violate these constraints – in other words, from a lack of appropriate constraints on component and system behavior.

STPA (System Theoretic Process Analysis) is a hazard analysis technique built on STAMP. As described above, accidents are viewed in STAMP as resulting from inadequate enforcement of constraints on system behavior. Figure 1 shows a generic (example) safety control structure to enforce safety constraints. Each hierarchical level of the control structure represents a control process and control loop with actions and feedback. Two control structures are shown in Figure 1 – system development (on the left) and system operations (on the right), both of which have different responsibilities with respect to enforcing safe system behavior. The reason behind the inadequate enforcement may involve classic component failures, but it may also result from unsafe interactions among components operating as designed or from erroneous control actions by software or humans.

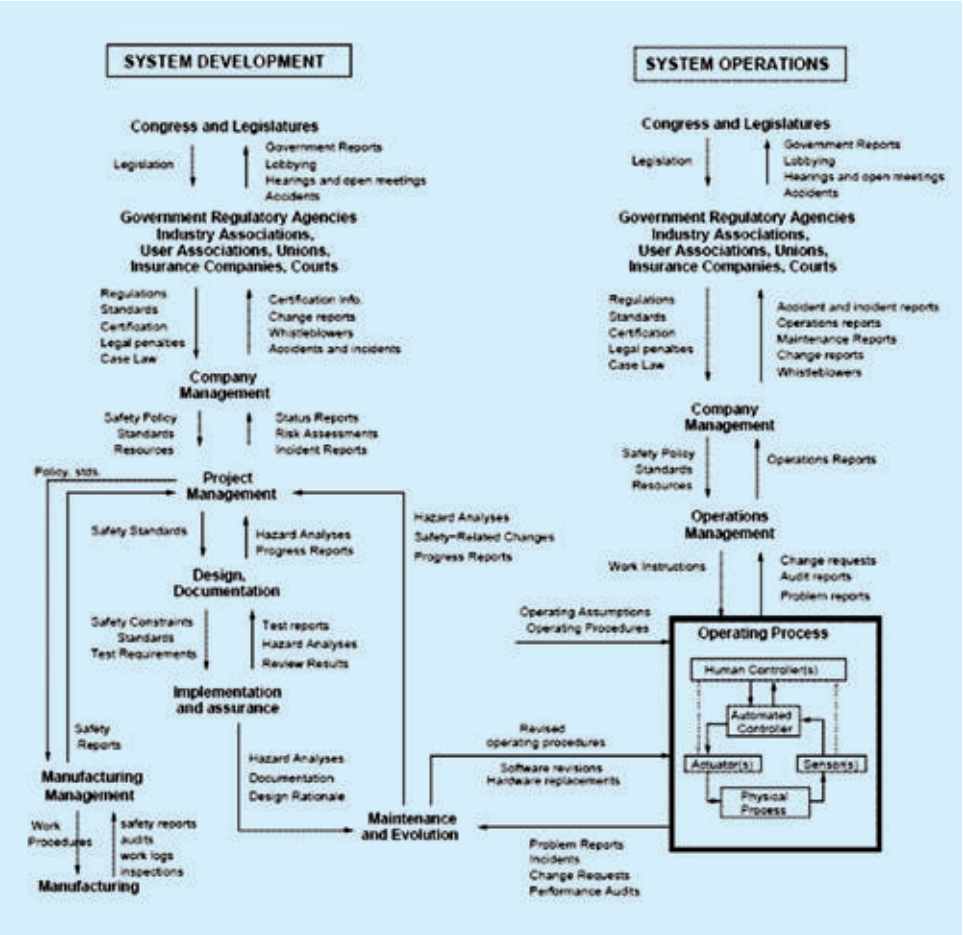


Figure 1: Example Model of a Socio-Technical Safety Control Structure

Human and automated controllers use a process model, usually called a mental model for humans, which they use to determine what control actions are needed (Figure 2). The process model contains the controller's understanding of (1) the current state of the controlled process, (2) the desired state of the controlled process, and (3) the ways the process can change state. Software and human errors often result from incorrect process models, e.g., the controller thinks the aircraft are in a different position than they really are and provides an unsafe control action (advisory). Accidents can therefore occur when an incorrect or incomplete process model causes a controller to provide control actions that are hazardous. While process model flaws are not the only causes of accidents involving software and human errors, they are a major contributor.

There are four types of hazardous control actions that need to be eliminated or controlled to prevent accidents:

1) A control action required for safety is not provided or is not executed.
2) An unsafe control action is provided that leads to a hazard.
3) A potentially safe control action is provided too late, too early, or out of sequence.
4) A safe control action is stopped too soon or applied too long.

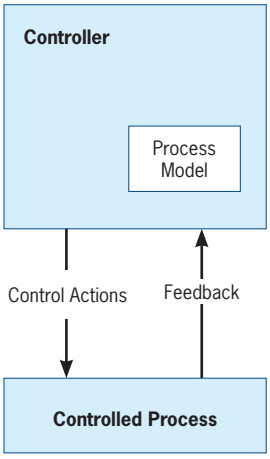


Figure 2: A Simple Control Loop and Process Model

Identifying the potentially unsafe control actions for the specific system being considered is the first step in STPA. These unsafe control actions are used to create safety requirements and constraints on the behavior of both the system and its components. Additional analysis can then be performed to identify the detailed scenarios leading to the violation of the safety constraints and used to generate more detailed safety requirements. As in any hazard analysis, these scenarios are the basis for designing controls and mitigation measures for the hazards. Any hazards that cannot be adequately controlled at the system level must be allocated in the form of behavioral requirements on the lower-level system components.

To demonstrate STPA, it is used ITP and the results compared to the hazard analysis method currently being used for NextGen/SESAR.

3. ATSA-ITP

Because of the limited radar coverage in oceanic and other remote airspace, air traffic controllers have historically relied on procedural separation rules to ensure safe traffic flow and minimum separation between aircraft. Aircraft in these sectors generally fly long, pre-defined flight paths (tracks). Air traffic controllers have limited options for knowing the positions of all aircraft in a sector at the same time or the ability to directly communicate with aircraft. To compensate for the limitations, these sections of the airspace often have much larger separation requirements than those applied to airspace with greater surveillance and communication coverage. The large separation minima, however, place severe limits on the capacity of a given track in a remote airspace.

The new Airborne Traffic Situational Awareness In-Trail Procedure (ATSA-ITP, referred to here as just ITP) will allow many of these previously blocked flight level changes to occur. ITP enables either leading or following Same Track aircraft to perform a climb or descent to a requested flight level through intervening flight levels. Note that standard separation minima between aircraft may not hold at times during the ITP maneuver. The ITP equipment must ensure that the reduced separation minima, as defined for ITP, are observed.

In the ITP procedure, the flight crew determines if the criteria for an ITP request are met. If the criteria are satisfied, the flight crew may request an ITP, identifying the Reference Aircraft in the request.

ATC verifies that the ITP and Reference Aircraft are Same Track and that the maximum Closing Mach Differential will not be exceeded. If the controller determines that separation minima will be met with all Other Aircraft, the climb or descent request may be granted. The controller does not determine or verify the separation distance from the Reference Aircraft." [5]

3.1 Current Process Used to Assure Safety in ITP

The safety analysis process adopted for NextGen and used for the ITP procedure uses traditional modeling and hazard analysis techniques. A Collision Risk Model was created to identify the aircraft state vectors needed to avoid a collision. The state vectors provide the information to determine the parameters that must be satisfied before the ITP is performed. We assume here that this model is correct and that the associated Operational Performance Assessment appropriately correlates with the risk model.

Figure 3: Summary of the official process to ensure safety in ITP

This paper concentrates on the hazard analysis, which for ITP is called the Operational Hazard Analysis and is documented in DO-312 [5]. Figure 3 shows the overall process being used in NextGen/SESAR, which relies heavily on probabilistic risk analysis using event trees and fault trees. This approach is contrasted in Section 4 with the alternative approach being proposed, which does not rely on probabilities that cannot possibly be known before extensive use of the system.

3.2 Applying STPA to ITP

The new hazard analysis technique based on STAMP is called STPA (System-Theoretic Process Analysis). The process involved and results are very different than the more traditional approach used on ITP. As in the traditional safety analysis, the process starts by identifying hazards, although hazards are not equated to failures, as is currently being done for ITP. Instead, as in system safety engineering for defense and space systems, a hazard is defined as a system state that under worst-case environmental conditions will lead to a loss or accident. This definition encompasses more than simply the states following component failures (the definition used in DO-312) but undesired states that can result from any cause. The general hazards for aircraft include:

- H-1: A pair of controlled aircraft violate minimum separation standards
- H-2: Aircraft enters unsafe atmospheric region
- H-3: Aircraft enters uncontrolled state
- H-4: Aircraft enters unsafe attitude (excessive turbulence or pitch/roll/yaw that causes passenger injury but not necessarily aircraft loss)
- H-5: Aircraft enters a prohibited area

Because ITP is at first only going to be used on long oceanic tracks, HI is the most relevant at this time and was the focus of our analysis and leads to the high-level system safety requirement/constraint: "The IPT must not cause a pair of controlled aircraft to violate minimum separation standards." The STPA hazard analysis identifies ITP system and component requirements necessary to enforce this constraint. In the future, if ITP is to be allowed in other locations, the analysis must be augmented to include any of the other hazards that become relevant.

STPA is performed on a functional control diagram of the system, which is shown in Figure 4 for the ITP-related parts of the system. Levels of control above the ATC manager are not shown.

STPA is broken into two steps in order to better organize the process. The first step identifies hazardous control actions for each component that could produce a system-level hazard by violating the system safety constraint. These identified unsafe control actions are used to refine the high-level safety constraint/requirement into more detailed safety requirements.
The second part of STPA analyzes the system to determine the potential scenarios that could lead to providing one of the unsafe control actions. These scenarios can then be used to derive detailed safety requirements/constraints for the system components to ensure that all the components operating together cannot create a system hazard, in this case H-1 (violation of minimum separation standards). The system and component-level requirements are used to design controls and to certify the safety of the system and of its components.

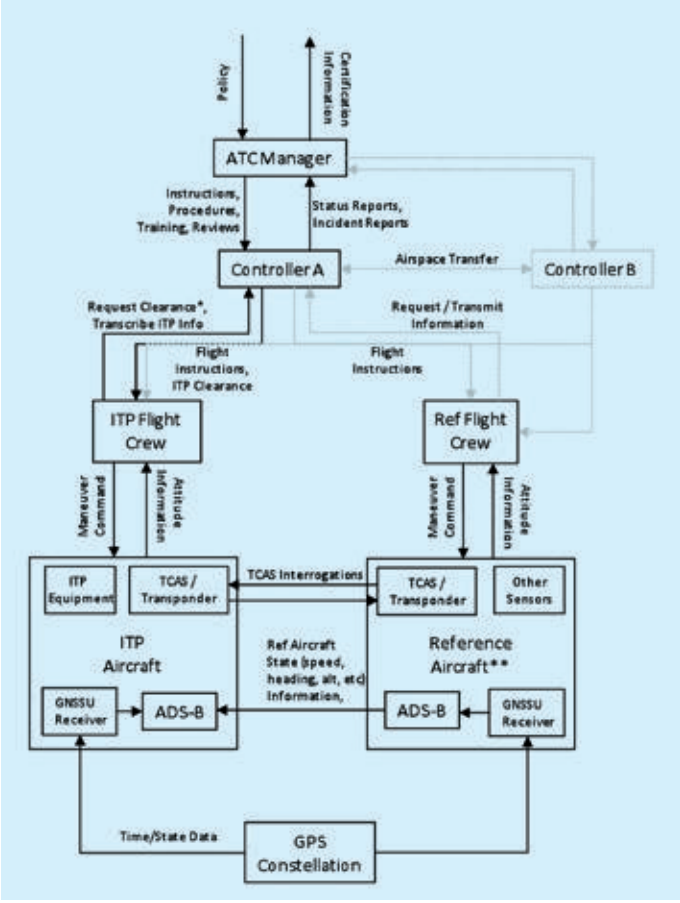


Figure 4: Safety control structure for ATSA-ITP

STPA Step One: The first step of STPA identifies control actions for each component that can lead to one or more of the defined system hazards. The four general types of unsafe control actions were shown above, i.e., providing a control action that leads to a hazard, not providing a control action that is needed to prevent a hazard, incorrect timing or sequencing, and applying a control action too long or stopping it too soon. A table can be used to document the hazardous control actions identified, as in Table 1 and Table 2. The hazardous control actions can then be translated into high-level system and component safety requirements and constraints.

Control Action	Not Providing Causes Hazard	Providing Causes Hazard	Wrong Timing/Order Causes Hazard	Stopped Too Soon/ Applied Too Long
Execute ITP		ITP executed when not approved ITP executed when ITP criteria are not satisfied ITP executed with incorrect climb rate, final altitude, etc	ITP executed too soon before approval ITP executed too late after reassessment	ITP aircraft levels off above requested FL 1 ITP aircraft levels off below requested FL2
Abnormal Termination of ITP	FC continues with maneuver in dangerous situation	FC aborts unnecessarily FC does not follow regional contingency procedures while aborting		

Table 1: Potentially hazardous control actions for the flight crew

[1] The unsafe control actions "ITP aircraft levels off above [and below] requested flight level" are not included in the following as separate unsafe control actions because they are equivalent to the unsafe control action "ITP executed incorrectly"

To produce the table, each potential entry is evaluated to determine whether that control action can lead to the system hazard (violation of minimum separation assurance). Consider the potential control action "Flight Crew Executes ITP" In Table 1. If the flight crew does not provide that control action, hazard H-1 does not result and the table entry is empty. On the other hand, there are several conditions under which providing the control action (execute ITP) could lead to the hazard, namely: executing the ITP procedure when it is not approved; executing it when the ITP criteria are not satisfied; and executing it with incorrect parameters (e.g., an incorrect climb rate or final altitude).

Control Action	Not Providing Causes Hazard	Providing Causes Hazard	Wrong Timing/Order Causes Hazard	Stopped Too Soon or Applied Too Long
Approve ITP request		Approval given when criteria are not met Approval given to incorrect aircraft	Approval given too early Approval given too late	
Deny ITP request				
Abnormal Termination Instruction	Aircraft should abort but instruction not given	Abort instruction given when abort is not necessary	Abort instruction given too late	

Table 2: Potentially hazardous control actions for ATC

Once the tables are created, the identified unsafe control actions are rewritten as high-level system safety constraints. The constraints are refined further, in a top-down system engineering process, during STPA Step Two. Although some (or most) of these may seem obvious, they are used not only in the refinement to more detailed requirements but also to provide a chain back to the unsafe control actions from the more detailed requirements and constraints in order to document where these requirements came from and therefore why they are needed.

The constraints on ATC identified in Table 2 are:

SC-ATC.1 Approval of an ITP request must be given only when the ITP criteria are met.

SC-ATC.2 Approval must be given only to the requesting aircraft.

SC-ATC.3 Approval must not be given too early or too late.

SC-ATC.4 An abnormal termination instruction must be given when continuing the ITP would be unsafe.

SC-ATC.5 An abnormal termination instruction must not be given when it is not required to maintain safety and would result in a loss of separation.

SC-ATC.6 An abnormal termination instruction must be given immediately if an abort is required.

The constraints on the ITP flight crew are:

SC-FC.1 The flight crew must not execute the ITP when it has not been approved by ATC.

SC-FC.2 The flight crew must not execute an ITP when the ITP criteria are not satisfied.

SC-FC.3 The flight crew must execute the ITP with correct climb rate, flight levels, Mach number, and other associated performance criteria.

SC-FC.4 The flight crew must not continue the ITP maneuver when it would be dangerous to do so.

SC-FC.5 The flight crew must not abort the ITP unnecessarily. (Rationale: An abort may violate separation minimums)

SC-FC.6 When performing an abort, the flight crew must follow regional contingency procedures.

SC-FC.7 The flight crew must not execute the ITP before approval by ATC.

SC-FC.8 The flight crew must execute the ITP immediately when approved unless it would be dangerous to do so.

SC-FC.9 The crew shall be given positive notification of arrival at the requested FL.

STPA Step Two: The second step of STPA examines each control loop in the safety control structure to identify potential causal factors for each hazardous control action, i.e., the scenarios for causing a hazard. This process will basically refine the high-level safety requirements identified in Step One. Figure 5 shows a generic control loop that can be used to guide this step. While STPA Step One focuses on the provided control actions (the upper left corner of Figure 9), STPA Step Two expands the analysis to consider causal factors along the rest of the control loop.

One reason a safety constraint might be violated is that the process model of the controller is incorrect, for example, the flight crew thinks it is safe to execute the ITP when it is not. The incorrect process model, in turn, may be the result of inadequate feedback provided by a failed or erroneous sensor or the feedback may be delayed or corrupted. Alternatively, the designers may have omitted a feedback signal or the flight crew may have received incorrect input from ATC or from other input devices (such as ADS-B).

Once the second step of STPA has been applied to determine potential causes for each hazardous control action identified in STPA Step One, the causes should be eliminated or controlled in the design at the system level or detailed requirements must be levied on the system components.

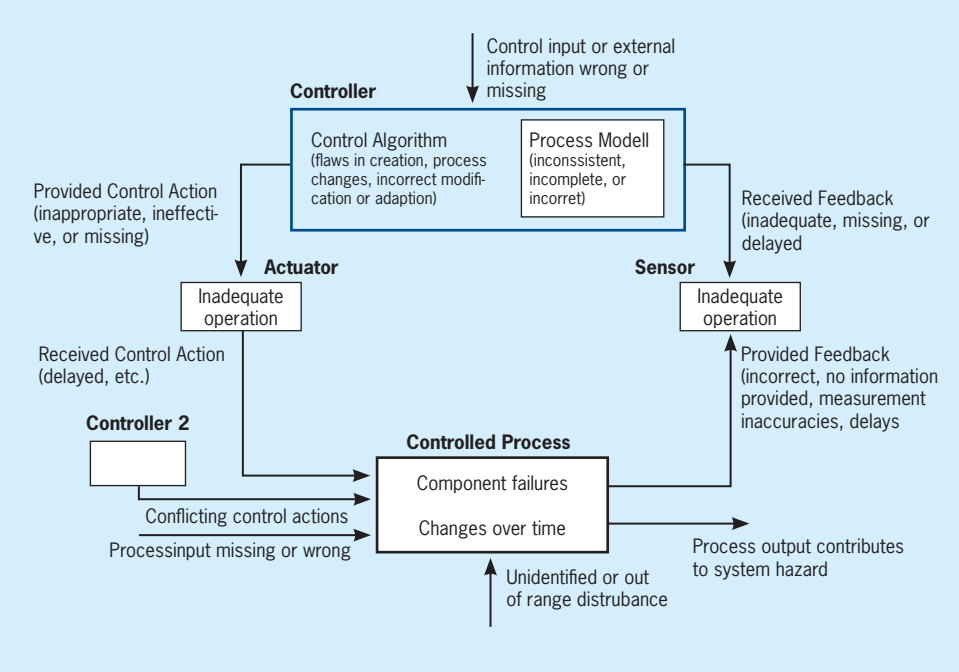


Figure 5: General control loop with causal factors

Figure 6 shows part of the Step Two analysis for unsafe flight crew behavior in executing the ITP. A more complete analysis can be found in [8]. The required process model and responsibilities of the flight crew are shown in the Flight Crew Controller Box. The unsafe control actions are written on the top left connection from the Controller to the actuator while some of the possible causes of unsafe flight crew behavior are shown on the other connections in the control loop.

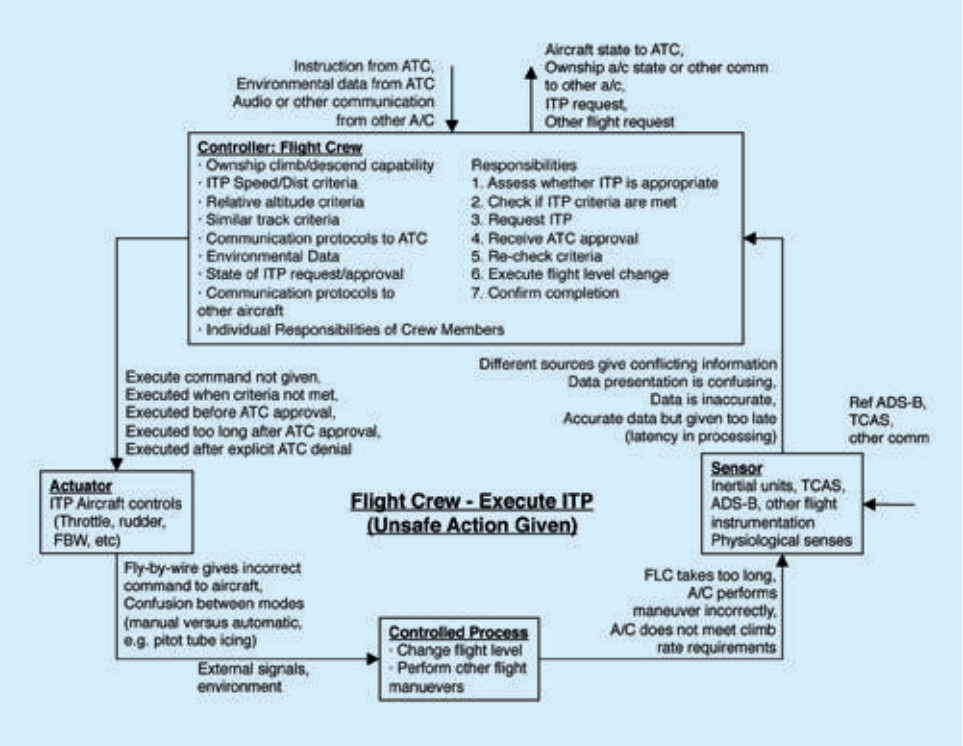


Figure 6: Partial causal analysis for the flight crew executing the ITP unsafely

The identified causal factors can be restated as requirements and then used to improve the system design to try to eliminate or mitigate such human errors, such as ensuring that the air traffic controllers and pilots have access to the information they need to make safe decisions and create safe control commands and that this information is available when they need it. The information produced by the Step 2 analysis may also be used in developing ATC and flight crew operational procedures and training.

Compare the results in Figure 6 with the fault tree used in the official hazard analysis of ITP as shown in Figure 7. The top level event in the fault tree figure is that the procedure is performed when one of the criteria for safe ITP is not satisfied and neither the flight crew nor ATC detects this non-compliance. According to the analysis, failure to comply can occur either because the flight crew does not understand the minimum distance required or the ATC does not receive the required data or fails to detect non-compliance due to a communication error involving corruption of data during transport. There are, of course, many other reasons for communication errors but these are ignored in the analysis. The leaf nodes in the tree are assigned an overall probability of occurrence for the hazardous event at the top of the tree.

Note that the fault tree assumes independent behavior, however the interaction and behavior of the flight crew and ATC may be coupled, with the parties exerting influence on each other or both being influenced by high-level system conditions.

In contrast, the STPA causal factors include the basic communication errors included in the fault tree, but also include additional reasons for communication errors as well as guidance for understanding human error within the context of the system. Communication errors may result, for example, because there is confusion about multiple sources of information (for either the flight crew or ATC), confusion about heritage or newly implemented communication protocols, or simple transcription or speaking errors. There is no way to quantify or verify the probabilities of any of these sources of error for many reasons, particularly because the errors are dependent on context and the operator environments are highly dynamic and, in fact, not necessarily designed yet. Perhaps that is why they were omitted from the fault trees.

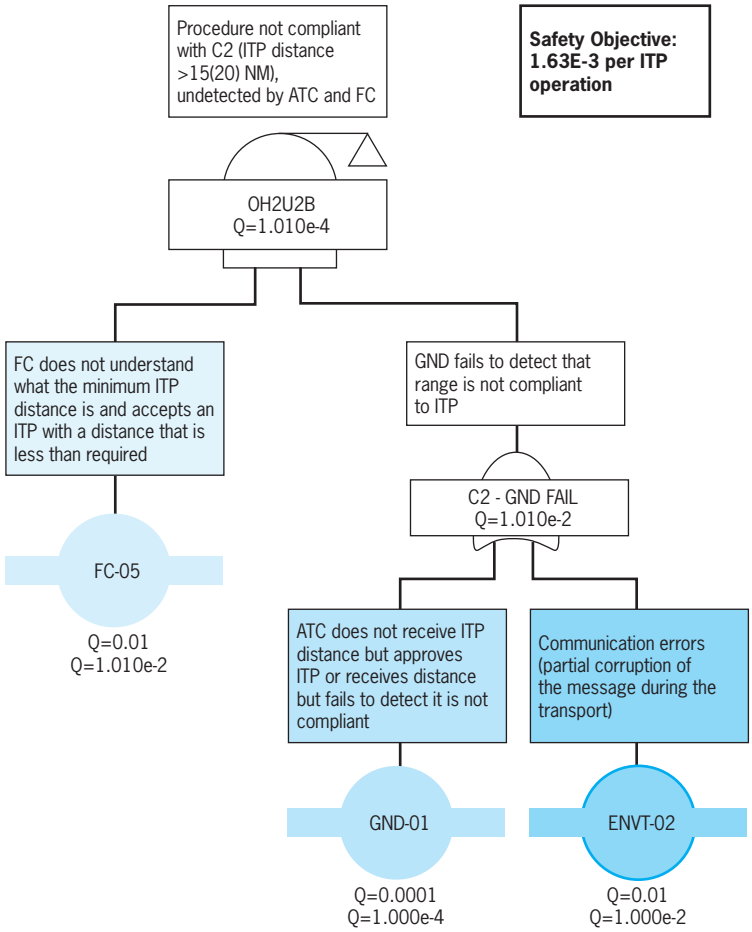


Figure 7: Example Fault Tree from DO-312

4. Comparing the results of the two analyses

The current approach to safety analysis for NextGen can be compared to our new approach both in terms of the actual safety requirements generated and in the underlying philosophical differences. In our comparison of the safety requirements generated by the two hazard analysis approaches, we included with the DO-312 analysis a set of additional requirements provided by the FAA for ITP [9]. Our analysis identified nineteen safety requirements that were not in either of the two official NextGen documents. One example of an important omitted requirement involves the reference aircraft. DO-312 assumes that the reference aircraft will not deviate from its flight plan during ITP execution. There needs to be a contingency or protocol in the event that the reference aircraft does not maintain its expected speed and trajectory, for example, because of an emergency requiring immediate action.

The more important comparison is not with the results of the specific techniques provided in DO-312 – they could be changed or improved – but rather the basic philosophical differences between the traditional approach to hazard analysis represented by the methods used in DO-312 and that of the new systems-theoretic approach described in this paper. The results of any hazard analysis are inextricably tied to the overall philosophy and viewpoint of the analysis approach, the causal factors identified, and the certification method implied by the safety analysis approach. Table 3 summarizes the comparison.

	DO-312	STAMP/STPA
Analysis Philosophy	Success oriented, i.e. it assumes nominal case then tries to predict probability of deviation Provides set of contingencies for off-nominal behavior	Assumes worst-case scenario, i.e. it starts with accident, then hazards, then causal factors and assumes that any of the causal factors can happen
	Emphasis on preventing or reducing failures	Emphasis on enforcing constraints on system (and thus component) behavior
	Assumes most failure modes are independent	Accounts for sub-system interactions and how these influence safety-related behavior
Causal Factors	Considers only hardware failures, or treats operators and software as if they are hardware (e.g. leaves on a fault tree with assigned probabilities of failure)	Assumes accidents are caused by lack of enforcement of safety constraints on the behavior of the system and its components. Assumes that software does not "fail" but can still be hazardous due to flawed requirements or unsafe interactions with rest of system Human operators perform within the context of a larger system design and, like software, do not necessarily "fail" but can make unsafe decisions
Certification Method	Assign performance goals or necessary probabilities of failure, then manufacturer attempts to assure compliance	Specify safety constraints derived from STPA, based on safety-related control actions and required component behavior, which manufacturer implements.

Table 3: General comparison of approaches

An important difference is that DO-312 starts with failures whereas STPA starts from hazard. All failures usually do not lead to hazards, and hazards may result from unsafe interactions among components that have not "failed." The equating of a hazard and a failure leads to omitting some unsafe states (see Figure 8). While some failure scenarios are unsafe, some unsafe scenarios lie outside the realm of a failure and are not included in the safety analysis when only failures are considered (as is the case for the safety analysis of ITP in DO-312).

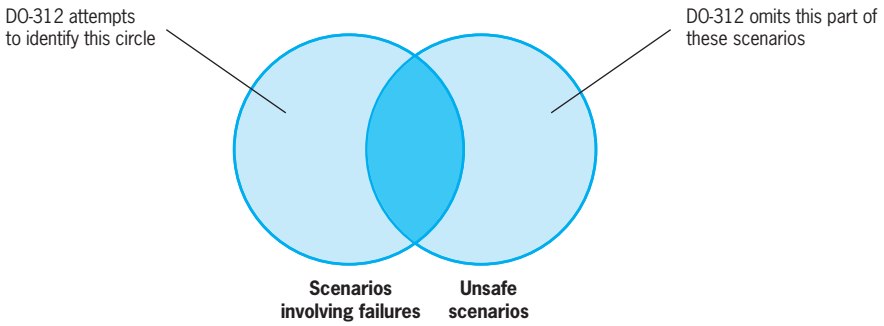


Figure 8: The Consequences of Equating Safety and Reliability

In addition, DO-312 assumes or prescribes independent probabilities for off-nominal behavior (which is common in analyses using fault trees or event trees). STPA, in contrast, accounts for component interaction accidents (where no components may have failed) and, in fact, assumes that not only can causal factors be dependent but also that the behavior of one component might be highly influential on other aspects of the system.

STPA also recognizes that software does not "fail" but merely performs the way it was designed: it can therefore be hazardous due to flawed requirements (or implementation) or unsafe interactions with the rest of the system. Most software-related accidents arise due to flaws in the software requirements [3] so getting a complete and correct set of safety-related requirements/constraints on software behavior is key to preventing software-related accidents.

Human operators also do not fail in the sense that hardware does and most of their failures are not random. Instead, humans are influenced by the design and operation of the overall system and the operational context and can thus make unsafe decisions due to the factors in Figure 5, such as incorrect mental models of the process they are controlling, possibly due to missing or incorrect feedback.

Human operators also do not fail in the sense that hardware does and most of their failures are not random.

The human error identification process used in DO312 for ITP is very incomplete, but the problems involve more than just incompleteness. Human error is treated in exactly the same way as a physical failure, that is, as a deviation from a predefined behavior or procedure. Unfortunately, this treatment of human error oversimplifies it as a binary decision between right and wrong. Many of the most important situations involved in accidents are overlooked because they are difficult or impossible to model in this way, including when:

- The correct behavior is not predefined or not clear
- The prescribed behavior is thought to be incorrect by the person responsible for following it
- Procedures conflict with each other, or it is not clear which procedure applies
- The person has multiple responsibilities or goals that may conflict
- The information necessary to carry out a procedure is not available or is incorrect
- Past experiences and current knowledge conflict with a procedure
- The procedure is misunderstood or the responsibility for the procedure is unclear
- The procedure is incorrect

The DO-312 analysis produced a short list of human errors such as "flight crew fails to detect inadequate climb/descent rate." Because basic causes are the "lowest level of failure" human errors are not analyzed in further detail. No attempt is made to understand why the human errors may arise or to prevent them. Instead, all errors are assumed to occur randomly at a given probability rate. If necessary, mitigation attempts are made to reduce the chance that human errors will lead to a hazard. Mitigation is done by adding barriers called "mitigation measures". Interestingly, every mitigation measure is simply a new procedure that is imposed on the humans.

Because human behavior was treated as random and no attempt was made to explain or understand the potential human errors, the possibility of eliminating or reducing errors was precluded. Human behavior is usually not random but is influenced by the current context, the information observed (rather than simply the information available [10]), and constructed beliefs about the operation of the system. Human behavior is also heavily dependent on interactions with other system components and past experiences. The treatment of human error as independent, random events precludes

the possibility of eliminating or reducing errors in the first place. Eliminating errors through system design is typically more effective than managing hazards through mitigation alone. There is also no guarantee that humans will perform better when additional (mitigation) procedures are added, and they may actually perform worse because of the added workload.

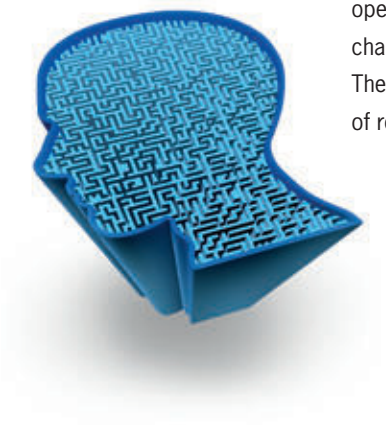
Instead, a safety analysis should not only identify what humans can do wrong, but also why and how it can be avoided. This is the goal of STPA. Assuming that the flight crew and ATCO are not intentionally malicious, this identification requires understanding the conditions under which each erroneous decision can make sense to them and modifying or adding system design requirements to help make the correct decisions obvious.

A final major difference between the two approaches is the certification method that results from them. The certification process for DO-312 compliance is performance based: The document specifies performance goals or necessary minimum probabilities of component and system failure in order to meet the assumptions used in the fault or event trees. For the performance goal approach, manufacturers must assure compliance with the performance metrics. Although many electro-mechanical components have sufficient heritage to yield an accurate probabilistic assessment, such individual physical component statistics may not hold in a complex system and are not useful for new components or old components operating in new environments, as is anticipated for NextGen (and SESAR). Even greater problems arise in defining probabilities for software and human errors. A good case can be made that such probabilities do not exist or cannot be determined even if they do. In practice, manufacturers either ignore the parts of their systems for which probabilistic data cannot be obtained, they estimate them, or they may even assign arbitrary numbers such as 10-4 for all software errors.

In contrast, the analysis produced by STPA results in a specification of behavioral constraints (requirements) that must be satisfied by the manufacturers to have their products certified for use in the system. This approach, in fact, was the one used for TCAS where behavioral requirements (using AND/OR tables as shown earlier in Figure 12) were provided and the manufacturers of TCAS boxes used standard assurance methods, such as DO-178B, to show that the requirements were satisfied.

5. Conclusions

The method being used in the hazard analysis for NextGen incorporates traditional hazard analysis methods (fault trees and event trees). The basic problem is that these methods were developed almost 50 years ago for systems composed primarily of hardware and of much less complexity than the transportation systems we are building today. They are not powerful enough to handle accidents involving interactions of components (system design errors) that do not involve failure of individual system components, software, and the cognitively complex tasks being assigned to human operators today. More powerful methods are needed to assure and certify safety in upgrades and changes to the NAS and other transportation systems. This paper describes a possible alternative. The alternative approach identifies a larger class of causes and provides a more comprehensive set of requirements.



References:

[1] FAA. What is NextGen, http://www.faa.gov/nextgen/why_nextgen_matters/what/.

[2] EUROCAE ED-78A / RTCA DO-264: Guidelines for Approval of the Provision and Use of Air Traffic Services Supported by Data Communications, March 2002.

[3] Leveson, N.G. (1995). Safeware. Amsterdam: Addison-Wesley Longman.

[4] Roland, H.E. & Moriarty, B. (1983). *System Safety Engineering and Management*, New York: John Wiley & Sons.

[5] RTCA. "Safety, Performance and Interoperability Requirements Document for the In-Trail Procedure in the Oceanic Airspace (ATSA-ITP) Application," DO-312, Washington DC, June 19, 2008.

[6] Leveson, N.G. (2012). *Engineering a Safer World*. Cambridge: MIT Press.

[7] Checkland, P. (1981). *Systems Thinking, Systems Practice*. Chichester, UK: Wiley.

[8] Leveson, N.G., Mats P.E. Heimdahl, H. Hildreth, H., Jon D. Reese, "Requirements *Specification for Process-Control Systems," IEEE Transactions on Software Engineering*, SE-20, No. 9, September, 1994.

[9] Leveson, N.G., "Intent Specifications: An Approach to Building Human-Centered Specifications," *IEEE Transactions on Software Engineering*, SE-26, No. 1, January 2000.

[10] Fleming, C.H., Spencer, M., Leveson, N.G. & Wilkinson, C. Safety Assurance in NextGen, NASA Technical Report NASA/CR-2012-217553.

[11] FAA. Interim Policy and Guidance for Automatic Dependent Surveillance Broadcast (ADS-B) Aircraft Surveillance Applications Systems Supporting Oceanic In-Trail Procedures (ITP), 2010.

[12] Dekker, S.W.A. (2004). *Ten Questions about Human Error*. Mahwah, NJ: Lawrence Erlbaum.

[13] Dekker, S.W.A. (2006). *A Field Guide To Understanding Human Error*. Aldershot, UK: Ashgate.

Erik Hollnagel, Professor at the University of Southern Denmark, Industrial Safety Chair at MINES Paris-Tech (France) and Professor Emeritus at University of Linköping (Sweden). He has worked at universities, research centres, and industries in several countries and in many domains, including nuclear power generation, aerospace and aviation, software engineering, land-based traffic, and healthcare. He has published widely and is the author/editor of 18 books including "FRAM: The Functional Resonance Analysis Method for Modelling Complex Socio-technical Systems" and "The ETTO Principle: Why things that go right, sometimes go wrong."

Resilience Engineering stellt einen neuartigen Weg dar, um sich dem Themenkomplex Safety zu nähern. Unfälle bedeuten in dieser Denkrichtung keinen Zusammenbruch von Systemkomponenten mehr, sondern eher das Nichtfunktionieren von Anpassungsprozessen, die einen reibungslosen Betriebsablauf garantieren. Diese Adaptationen sind jedoch notwendig, um mit der Systemkomplexität umzugehen. Safety wird nicht mehr als die Abwesenheit von Unfällen oder anderen unerwünschten Ereignissen angesehen, sondern als die Anwesenheit dieser Anpassungsprozesse. Denn diese Prozesse sind in komplexen Systemen mit begrenzten Ressourcen stets nötig. Resilience Engineering will dabei das Negative nicht weiter reduzieren, sondern eher das Positive ausbauen. Auch wir in der DFS sollten nicht nur ausschließlich darauf schauen, was nicht gut gegangen ist, sondern uns auch diejenigen Situationen zum Vorbild nehmen, in denen alles gut gelaufen ist, um daraus zu lernen. Erik Hollnagel war maßgeblich daran beteiligt, diesen Ansatz zu formulieren. Für diese Ausgabe hat er noch einmal die Kerngedanken von Resilience Engineering zusammengefasst und den Wandel von Safety I zu Safety II beschrieben.

From reactive to proactive safety management

By Erik Hollnagel

Abstract

The sustained existence of modern societies depends on the safe and efficient functioning of multiple systems, functions, and specialised services. Because these often are tightly coupled, safety cannot be managed simply by responding whenever something goes wrong. Both theory and practice demonstrate that safety management that follows developments rather than leads them runs a significant risk of lagging behind and of becoming reduced to uncoordinated and fragmentary firefighting. In order to prevent this from happening, safety management must look ahead, not only to avoid that things go wrong but also – and more importantly – to ensure that they go right. Proactive safety management must focus on how everyday performance usually goes well rather than on why it occasionally fails, and must actively try to improve the former rather than simply prevent the latter.

Safety as the freedom from unacceptable risk

Safety is traditionally defined as a condition where the number of things that go wrong is acceptably small. But this means that safety is defined by its opposite, by what happens when it is missing. It also means that safety is measured indirectly, not by its presence or as a quality in itself, but by its absence or the consequences of its absence.

From a utilitarian point of view it makes sense to focus on situations where things go wrong, both because such these by definition are unexpected and because they may lead to loss of life and property. Classical safety concerns, addressed risks related to passive technology such as buildings, bridges, ships, etc. The rapid mechanisation of work that followed the industrial revolution in the eighteenth century led to a growing number of hitherto unknown types of accidents, where the common factor was the breakdown, failure, or malfunctioning of active technology. The focus on technology as the main source of both problems and solutions in safety went largely unchallenged until 1979, when the accident at the Three Mile Island nuclear power plant demonstrated that safeguarding technology was not enough. Following that it became necessary to consider human failure and malfunctioning as a potential risk. Seven years later the accident in Chernobyl required yet another extension, this time by adding the influence of organisational failures and safety culture to the common lore.

Historically speaking, new types of accidents have been accounted for by introducing new types of causes (e.g., metal fatigue, 'human error,' organisational failure) rather than by challenging or changing the basic underlying assumption of causality. We have therefore through centuries become so accustomed to explaining accidents in terms of cause-effect relations – simple or compound – that we no longer notice it. And we cling tenaciously to this tradition, although it has becomes increasingly difficult to reconcile with reality.

Looking at what goes wrong rather than looking at what goes right

To illustrate the consequences of looking at what goes wrong rather than looking at what goes right, consider Figure 1. This represents the case where the (statistical) probability of a failure is 1 out of 10,000 – technically written as $p = 10^{-4}$. This means that for every time we expect that something will go wrong (the thin line), there are 9,999 times where we should expect that things will go right and lead to the outcome we want (the grey area).

As an illustration of this, consider the statistics for Frankfurt airport (EDDF). In 2011 there were a total of 490,007 movements, but only 10 infringements of separation and 11 runway incursions. This corresponds to a ratio of 2.04 10-5 and 2.25 10-5, respectively, or roughly 2 cases out of every 100.000.

The tendency to focus on what goes wrong is reinforced by requirements from regulators and authorities; it is supported by models and methods; it is documented in countless databases; it is described in literally thousands of papers, books, and conference proceedings; and there is an untold number of experts, consultants, and companies that constantly remind us of the need to avoid risks, failures, and accidents – and of how their services can help to do that. The net result is abundant information about how things go wrong. This focus also conforms to our stereotypical understanding of what safety is and that safety should be managed by following the simple 'find and fix' principle: look for failures and malfunctions, try to find their causes, and try to eliminate causes and/or improve barriers.

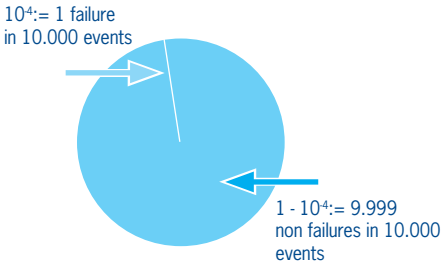


Figure 1: The imbalance between things that go right and things that go wrong

But there is little encouragement to consider that which goes right, i.e., the 9,999 events out of the 10,000 – or in the case of EDDF, the 99,998 events out of 100,000. There are no requirements authorities and regulators to look at what works well; there are few theories or models about how human and organisational performance succeeds, and few methods to help us study how it happens; examples are few and actual data difficult to locate; it is hard to find papers, books or other forms of scientific literature about it; and there are not many people who claim expertise in this area or even consider it worthwhile. It furthermore clashes with the traditional focus on failures, and even those who find it a reasonable endeavour are at a loss when it comes to the practicalities.

Safety-I: Avoiding That Things Go Wrong

The traditional definition of safety as a condition where the number of adverse outcomes is as low as possible can be called Safety-I. The purpose of managing Safety-I is therefore to achieve and maintain that state. The International Civil Aviation Organization defines safety as "the state in which harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and risk management."

The focus on failures creates a need to find the causes of what went wrong. When a cause has been found, the next logical step is either to eliminate it or to disable suspected cause-effect links. Following that, the outcome is then measured by counting how many fewer things go wrong after the intervention. Safety-I thus assumes that safety can be achieved by eliminating or weakening the causes of adverse events and that it can be measured by a reduced number of adverse outcomes. Safety-I also tacitly assumes that systems work because they are well designed and scrupulously maintained, because procedures are complete and correct, because designers can foresee and anticipate even minor contingencies, and because people behave as they are expected to – and more importantly as they have been taught or trained to do. This unavoidably leads to an emphasis on compliance in the way work is carried out.

Safety-I is reactive, because it based on responding to something that either has gone wrong or has been identified as a risk – as something that could go wrong. The response typically involves looking for ways to eliminate the cause – or causes – that have been found, or to control the risks, either by finding the causes and eliminating them, or by improving options for detection and recovery.

Safety-II: Ensuring That Things Go Right

As technical and socio-technical systems have continued to develop, work environments have gradually become more intractable. Since the models and methods of Safety-I assume that work is tractable in the sense that it is well-understood and well-behaved, they are less and less able to deliver the required and coveted 'state of safety.' As this inability cannot be overcome by 'stretching' the tools of Safety-I even further, it makes sense to change from Safety-I, where the goal is to avoid that something goes wrong, to Safety-II, where the goal is to ensure that everything goes right. Safety-II is thus the ability to succeed under varying conditions, so that the number of intended and acceptable outcomes is as high as possible – essentially paraphrasing the definition of resilience engineering. Safety-II explicitly assumes that systems work because people learn to identify and overcome design flaws and functional glitches, because they can recognise the actual demands and adjust their performance accordingly, and because they interpret and apply procedures to match the conditions. Systems also work because people can detect and correct when something goes wrong or when it is about to go wrong, hence intervene before the situation becomes seriously worsened. The consequence of this definition is that the basis for safety and safety management now becomes an understanding why things go right, which means an understanding of everyday activities.

In contrast to Safety-I, Safety-II acknowledges that systems are incompletely understood, that descriptions can be complicated, and that changes are frequent and irregular rather than infrequent and regular. Safety-II, in other words, acknowledges that systems are intractable rather than tractable. While the reliability of technology and equipment in such systems may be high, workers and managers frequently trade-off thoroughness for efficiency, the competence of staff varies and may be inconsistent or incompatible, and reliable operating procedures are scarce. Under these conditions humans are clearly an asset rather than a liability and their ability to adjust what they do to the conditions is a strength rather than a threat. Since safety cannot be achieved by constraining performance variability, efforts are needed to support the ubiquitous performance adjustments by clearly representing the resources and constraints of a situation and by making it easier to anticipate the consequences of actions.

For proactive safety management to work, it is necessary to foresee what could happen with acceptable certainty and to have the appropriate means (people and resources) to do something about it.

From a Safety-II perspective, the purpose of safety management is to ensure that as much as possible goes right, in the sense that everyday work achieves its stated purposes. This cannot be done by responding alone, since that will only correct what has happened. Safety management must instead be proactive, so that adjustments can take place before something happens and therefore affect how it happens or even prevent something from happening. A main advantage is that early responses, on the whole, require less effort because the consequences of the event will have had less time to develop and spread. And early responses obviously save valuable time.

For proactive safety management to work, it is necessary to foresee what could happen with acceptable certainty and to have the appropriate means (people and resources) to do something about it. That in turn requires an understanding of how the system works, how its environment develops and changes, and of how functions may depend on and affect each other. This understanding can be developed by looking for patterns and relations across events rather than for causes of individual events. To see and find those patterns, it is necessary to take time to understand what happens rather than spend all resources on fire-fighting.

Conclusion

While day-to-day safety management rarely is purely reactive, the pressure in most work situations is to be efficient rather than thorough, which reduces the incentive to be proactive. Proactive safety management does require that some effort is spent up front to think about what could possibly happen, to prepare appropriate responses, and to allocate resources and make contingency plans. Here are some practical suggestions for how to begin that process:

- Look at what goes right, as well as what goes wrong. Learn from what succeeds as well as from what fails. Indeed, do not wait for something bad to happen but try to understand what actually took place in situations where nothing out of the ordinary seemed to happen. Things do not go well because people simply follow the procedures. Things go well because people make sensible adjustments according to the demands of the situation. Find out what these adjustments are and try to learn from them!

- When something has gone wrong, look for everyday performance variability rather than for specific causes. Whenever something is done, it is a safe bet that it has been tried before. People are quick to find out which performance adjustments work and soon come to rely on them – precisely because they work. Blaming people for doing what they usually do is therefore counterproductive. Instead one should try to find the performance adjustments people usually make as well as the reasons for them. Things go wrong for the same reasons that they go right, but it is far easier and less incriminating to study how things go right.

Whenever something is done, it is a safe bet that it has been tried before.

- Look at what happens regularly and focus on events based on how often they happen (frequency) rather than how serious they are (severity). It is much easier to be proactive for that which happens frequently than for that which happens rarely. A small improvement of everyday performance may count more than a large improvement of exceptional performance.

- Allow time to reflect, to learn, and to communicate. If all the time is used trying to make ends meet, there will no time to consolidate experiences or replenish resources – including how the situation is understood. It must be legitimate within the organisational culture to allocate resources – especially time – to reflect, to share experiences, and to learn. If that is not the case, then how can anything ever improve?

- Remain sensible to the possibility of failure – and be mindful. Try to think of – or even make a list of – undesirable situations and imagine how they may occur. Then think of ways in which they can either be prevented from happening, or be recognised and responded to as they are happening. This is the essence of proactive safety management.

Jörg Leonhardt. Leiter des Bereichs Human Factors im Unternehmenssicherheits-management (VY/H). Er ist ein Human-Factors-Experte und hält ein Master Degree in Human Factors und System Safety der Universität Lund, Schweden.

“You can see how all the pieces fit, when you watch them fall apart.“ (Willie Nelson)

Von Jörg Leonhardt

In den vergangenen Ausgaben des Safety Letter Spezial wurden verschiedene Aspekte des Wandels von einem reaktiven zu einem pro-aktiven Safety-Management unter Human-Factors-Gesichtspunkten beleuchtet.

Der Mensch als derjenige, der durch seine Anpassungsfähigkeit an sich ständig ändernde Anforderungen die Sicherheit durch sein aktives Tun in einem komplexen und interaktiven System herstellt, steht dabei immer im Fokus der Aufmerksamkeit. Kommt es zu ungewollten Ereignissen und folgt eine Untersuchung des Vorfalles, sieht man oft Zusammenhänge, die man vorher nicht erkannt hat oder nicht erkennen wollte. Durch die Rückschau und die Kenntnis über das Ergebnis des Vorfalles ist es kaum vorstellbar, dass die jetzt klar auf der Hand liegenden Zusammenhänge vorher nicht erkannt wurden.

Der sogenannte „Hindsight Bias“ beeinträchtigt unser Urteilungsvermögen und führt oft zu Schlussfolgerungen, die aus der Offensichtlichkeit nach Eintritt des Ereignisses getroffen werden. „Der Lotse hätte nur…“ ist dann oftmals die schnell gefundene Erklärung und die Maßnahmen, die dann folgen, sind bekannt.

Da der Vorfall aber in einem Kontext passiert, indem es verschiedene Wirkungskräfte gibt, ist eine Reduktion auf eine singuläre Ursache geradezu fahrlässig. Es ist davon auszugehen, dass eine getroffene Entscheidung in diesem Kontext für den Entscheider Sinn gemacht hat. Daher gibt die Analyse des Kontextes mehr Aufschluss über die Ursachen. Organisationen, die in der heutigen Zeit noch an Human Error als Ursache und Erklärung von Vorfällen festhalten, haben aus den Katastrophen der Vergangenheit – sei es Überlingen, Deepwater Horizon oder Fukushima – nichts gelernt.

Im Nachhinein zu beurteilen und zu wissen, was falsch war, ist nicht hilfreich, da es den Kontext außer Acht lässt und sich nur unzureichend Verbesserungen ableiten lassen. Das betrifft im Übrigen nicht nur die operativen Mitarbeiter eines Unternehmens, auch Managementfehler oder Fehler in Projekten lassen sich mit der Kenntnis über das Ergebnis ebenso schnell finden und beschreiben. Das Entscheiden in einer hochkomplexen Situation, die gekennzeichnet ist von Produktions-, Zeit- und Kostendruck, ist nicht vergleichbar mit einer Analyse mit ausreichend Zeit und Einblick in Zusammenhänge, die erst im Nachhinein bekannt wurden.

Die Wirkfaktoren Kosten und Zeit spielen heutzutage auf allen Ebenen der Organisation eine enorme Rolle.

Wer konnte schon ahnen, dass der Ausbruch eines Vulkans in Island das gesamte System Luftfahrt binnen drei Tagen an den Rand seiner ökonomischen Überlebensfähigkeit bringt? Wer konnte ahnen, dass die Vergabe von Krediten an Kunden, die nicht über ausreichend Einkommen verfügen, um diesen je zurückzuzahlen, das Währungssystem in Europa nahe an einen Kollaps bringt? Im Nachhinein erscheint dies offensichtlich.

„The past seems incredible, the future implausible“ haben es einmal David D. Woods und Richard Cook treffend formuliert (Woods & Cook, 2002).

Doch was ist die Konsequenz aus dieser Erkenntnis?
Foresight anstelle von Hindsight könnte der richtige Weg sein. Die relativ neue Theorie des Resilience Engineering, die seit 2005 immer bedeutender wird, bietet für den Wechsel von einem retrospektiven zu einem prospektiven Safety-Management die Grundlage (Hollnagel, Woods & Leveson, 2005).

Im Kern geht es um die „Überlebensfähigkeit“ eines Systems, einer Organisation. Nicht nur im Sinne von Safety, sondern auch im Sinne ökonomischer Überlebensfähigkeit.

Werden die Zusammenhänge erst dann verstanden, wenn das Ganze auseinanderbricht, ist es zu spät.

„Foresight“ oder proaktives Safety-Management bedeutet nicht, die Zukunft vorhersagen zu können, sondern sich über den Systemstatus bewusst zu sein und abschätzen zu können, welche Wirkung durch Veränderungen oder äußere Einflussfaktoren erzeugt werden und wieviel „Puffer“ vorhanden ist, um mit diesen Anforderungen umgehen zu können.

“We mistake hazards under control for eliminating hazards, because we have effective mechanism we forget that we still have hazards.” (Woods, 2012)

Was macht eine Organisation resilient?
„Resilience is the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions.“ (Hollnagel et al., 2011)

Das setzt voraus, dass Bereiche und Abteilungen zusammen und nicht gegeneinander arbeiten und dass Feedback eine wichtige Informationsquelle darstellt.

Resilience Engineering baut auf die Eckpunkte:
Monitoring: Zu wissen, wo man steht
Anticipate: Zu wissen, was auf einen zukommen kann
Respond: Reagieren zu können, wenn es nötig ist
Learning: Aus der Erfahrung lernen und sich verändern

Monitoring bezieht sich auf die Fähigkeit, sein System zu kennen – den Überblick über Abläufe, Kürzungen und Doppelarbeiten genauso wie den Zustand des Gesamtsystems. Hierzu können Kennzahlen genutzt werden. Deren Aussagekraft über das Gesamtsystem und seiner Komponenten ist aber gering und sie liefern mehr einen Rahmen als ein Ergebnis.

Monitoring beinhaltet auch, dass Informationen zusammenfließen und somit der Gesamtüberblick erhalten bleibt. Das setzt voraus, dass Bereiche und Abteilungen zusammen und nicht gegeneinander arbeiten und dass Feedback eine wichtige Informationsquelle darstellt.

Die NASA hat einen schweren und harten Lernprozess hinter sich, als deutlich wurde, dass Organisationsstrukturen und ausgeprägtes Bereichsdenken („Structural Secrecy“) einen entscheidenden Beitrag zum Verlust der Raumfähre Challenger lieferten.

Die kritischen Denker, die Querdenker im Unternehmen, können einen Beitrag leisten, um die Fähigkeit der Antizipation zu erhöhen. Konformität im Denken hingegen schränkt sie ein.

Monitoring heißt auch, nicht ausschließlich anzunehmen, dass eine Organisation nach ihren Organigrammen, Regeln und Vorgaben läuft. *„Work as imagined or work as done?”* wie es Professor Hollnagel formuliert.

Anticipate ist die Fähigkeit zur Vorausschau. Die Kreativität, sich Dinge vorzustellen zu können, zu denen es noch keine Erfahrungen oder wenig Erfahrungen gibt. Das heißt aber nicht, in die Glaskugel zu schauen und die Zukunft vorherzusagen, aber eben auch nicht den Erfolg der Vergangenheit als Garantie für die Zukunft zu nehmen.

Auch hier hat die NASA eine schwere Lektion lernen müssen. Der Columbia-Katastrophe ist vorausgegangen, dass man sich intensiv mit den potenziellen Gefährdungen durch Einschläge von Debris im All, aber nicht beim Launch beschäftigt hat. Die Beschädigung der Hitzekacheln während der Startphase hat letztlich dazu geführt, dass der Shuttle beim Wiedereintritt in die Erdatmosphäre verglüht ist.

Die kritischen Denker, die Querdenker im Unternehmen, können einen Beitrag leisten, um die Fähigkeit der Antizipation zu erhöhen, Konformität im Denken hingegen schränkt sie ein.

Respond meint, ob und wie das Unternehmen in der Lage ist, auf unvorhergesehene Ereignisse zu reagieren. Wie viel Puffer bleibt, wie viele Ressourcen können wie schnell aktiviert werden und wie ist der Handlungsspielraum.

Dazu ein Beispiel aus dem Krankenhaus, viele Unfälle passieren im Krankenhaus, also jemand befindet sich z.B. wegen einer Herzerkrankung stationär in einem Krankenhaus. Beim Umbetten stürzt der Patient und bricht sich die Hüfte. Durchschnittliche Dauer vom Unfall (Hüfte gebrochen) bis zur Operation (Hüfte) in einem großen kanadischen Krankenhaus = 48 Stunden.

Die Operationssäle sind ausgelastet und durchgetaktet, die Personaldecke ist ausreichend für das Tagesgeschäft, aber nicht für unerwartete Ereignisse und die Konzentration liegt auf den – gewinnbringenden - Kernaufgaben (z.B. Herzkatether).

Das System ist optimiert - aus Sicht der Betriebswirtschaft. Ein Krankenhaus ist aber auch ein hochkomplexes System und in der Interaktivität der Systemkomponenten durchaus mit einer Flugsicherung vergleichbar. Es verfolgt mehrere Ziele - nicht nur betriebswirtschaftliche - und muss diese gleichwertig berücksichtigen, um resilient zu sein.

Learning steht für die Fähigkeit, zu lernen und sich weiterzuentwickeln. Nicht – nur – aus Fehlern und Vorfällen, sondern aus den Prozessen und Erkenntnissen des Monitorings, Antizipierens und Reagierens. Die Welt hat sich verändert und verändert sich weiter. Dies kann man schnell an den eigenen Erfahrungen mit der Nutzung von Mobiltelefonen, E-Mail, Internet oder den neuen sozialen Medien feststellen.

Wir können daher nicht die Probleme der Zukunft mit Modellen aus der Vergangenheit lösen, Lernen ist ein vorausschauender Prozess. Hilfreich ist auch, wenn die Erkenntnisse aus anderen Branchen und aus Großschadensereignissen im Unternehmen kommuniziert und diskutiert werden.

Wir können daher nicht die Probleme der Zukunft mit Modellen aus der Vergangenheit lösen, Lernen ist ein vorausschauender Prozess.

Ein „das trifft auf uns nicht zu. Das kann man nicht vergleichen“ limitiert die Möglichkeiten, hinzuzulernen.

Nimmt man diese vier Eckpfeiler des Resilience Engineering wäre die Richtung für ein Umdenken und eine Neuausrichtung aufgezeigt.

Nicht nur die DFS, auch andere ANSPs und Organisationen, die in einem komplexen und interaktiven Umfeld operieren, müssen umdenken.

Das Ausbalancieren von akutem Kostendruck gegenüber einer chronischen Anforderung an die Sicherheit benötigt Systemkenntnis.

Systemkenntnis kann erlangt werden durch beständiges Monitoren der aktuellen Prozesse und Abläufe, um die Manövrierfähigkeit und die zusätzlichen Ressourcen (Personal, Technik und Finanzen) realistisch einschätzen zu können und zu wissen, wie viel Spielraum vorhanden ist, um auf unerwartete Ereignisse reagieren zu können.

Entscheider benötigen verlässliche Informationen, die über das Zusammenfassen von Kennzahlen hinausgehen. Wissen wo die Unternehmung - das Gesamtsystem – steht, um nachher nicht zusehen zu müssen, wie alles zerfällt.

Die Anfänge sind gemacht. Das „neue Denken“ findet Einzug in die DFS, die Qualität der Vorfalluntersuchung verbessert sich stetig und damit das Lernen aus Vorfällen.

Der Bereich VY/H arbeitet im Auftrag von EUROCONTROL, zusammen mit den Niederlassungen und führenden Wissenschaftlern, an einem Konzept zur Erfassung von „Weak Signals“.

Die „Weak Signals“ sollen uns helfen, frühe Anzeichen von Veränderung und von Drift zu erkennen, um antizipierende Maßnahmen zu ergreifen.

Die DFS ist hier mal wieder innovativ und wegweisend für andere ANSPs. Wünschenswert wäre, wenn neue Ideen auch im Unternehmen stärker wertgeschätzt und gefördert würden.

„Das einzig Beständige ist die Veränderung“ war mal ein Leitgedanke der DFS.

Literaturhinweise:

Hollnagel, E. (2009). The *ETTO principle: efficiency-thoroughness trade-off*. Farnham: Ashgate.

Hollnagel, E., Woods, D.D. & Leveson, N. (Eds.) (2005). Resilience Engineering: *Concepts and Precepts*. Aldershot: Ashgate.

Hollnagel, E., Pariès, J., Woods, D.D. & Wreathall, J. (Eds.) (2011). Resilience *Engineering in Practice*: A Guidebook. Farnham: Ashgate.

Woods, D.D. (2012). *Personal Communication*, July 2012.

Woods, D.D. & Cook, R.I. (2002). Nine Steps to Move Forward from Error. *Cognition, Technology, and Work*, 4(2): 137-14

DFS Deutsche Flugsicherung GmbH
Unternehmenssicherheitsmanagement
Am DFS-Campus 10
63225 Langen

Telefon +49 06103 707-4111
Telefax +49 06103 707-4196
E-Mail safetyletter@dfs.de
Internet www.dfs.de