

Five Genotypes of Failure in the Columbia Accident

David Woods
The Ohio State University
woods.2@osu.edu

Draft
9-2-03

Five Patterns

The accident report identifies a variety of contributors to the accident. These factors have been seen before in other accidents. Focusing on the general patterns or 'genotypes' (Hollnagel, 1993) present in this particular accident help guide the process of organizational change.

Classic patterns abstracted from other accidents and data appear to be present in this accident, including:

- Drift toward failure as defenses erode in the face of production pressure.
- An organization that takes past success as a reason for confidence instead of investing in anticipating the changing potential for failure.
- Fragmented problem solving process that is unable to see the big picture.
- Failure to revise assessments as new evidence accumulates.
- Breakdowns at the boundaries of organizational units.

1. The basic classic pattern in this accident is—*Drift toward failure as defenses erode in the face of production pressure.*

My colleague captured the heart of Columbia when he commented on other accidents:

If anything is unreasonable, it is the requirement to be both efficient and thorough at the same time – or rather to be thorough when with hindsight it was wrong to be efficient.

Erik Hollnagel 6-12-2001

Hindsight bias, by oversimplifying the situation people face before outcome is known, often hides tradeoffs between multiple goals. The analysis in the CAIB report provides the general context of shrinking margins on production goals (the tradeoff space getting squeezed tighter) and how that pressure created strong incentives to move forward and look askance at potential disruptions to schedule.

Goal tradeoffs sometimes are explicitly considered or noted. On the other hand, the tradeoffs often become decided gradually as pressure leads to a focus on some goals and blocks seeing the tradeoff with other goals. This process usually happens with conflicts between acute goals like production/efficiency taking precedence over chronic

goals like safety. It isn't safety is ignored but with little time/resources available good intentions start to become seen as a reasonable substitute for active investment of energy on the chronic goals.

The dilemma of production/safety conflicts is: If organizations never sacrifice production pressure to follow up warning signs, they are acting much too risky. On the other hand, if uncertain "warning" signs always lead to sacrifices on acute goals, can the organization operate within reasonable parameters or stakeholder demands?

Whenever the organization is under tightening production pressure, proactive search for safety side effects is warranted. This doesn't happen in many organizations and didn't happen here either.

The paradox of production/safety conflicts is: it is precisely at points of intensifying production pressure that extra safety investments need to be made in the form or proactive searching for side effects of the production pressure and in the form or re-assessing the risk space—safety investments are most important when least affordable.

This genotype points toward several constructive issues:

How does an organization monitor for drift and its associated signs?

This tasks a safety organization and points toward some indicators they should use to monitor the organization's model of itself, how it is vulnerable to failure, and the potential effectiveness of the countermeasures it has adopted. New work has begun to make progress on how to do this.

How does production pressure create or exacerbate tradeoffs between some goals and chronic concerns like safety?

This tell us that effective organizations need to recognize when side effects of production pressure may be increasing safety risks and have some way to add investments to safety issues at the very time when the organization is most squeezed. For example, how does an organization note an increasingly tight box or squeeze (no margin; focus on creating margin; deciding what had been requirements are just options that can be foregone)?

Creating a new safety organization and culture at NASA means providing support for these functions: (1) detecting signs of increasing organizational risk (e.g., the adaptations to meet the schedule squeeze); (2) having the means to search for side effects on riskiness and the resources or authority to make extra investments in safety at precisely these times when it appears least affordable (including sacrificing schedule, production, and efficiency to control organization's level of riskiness; (3) having a means to recognize when extra investments need to be made and where to control rising signs of organizational risk.

2. Another genotype in Columbia is that *an organization takes past success as a reason for confidence instead of digging deeper to see the underlying risks* or investing in anticipating how changes afoot may affect the potential for failure.

One component in the drift process is the interpretation of past “success”. The absence of failure is taken as positive indication that hazards are not present or that countermeasures are effective. An organization usually is only able to change its model of **itself** unless and until sufficient evidence accumulates that demands revising the model. This is a guarantee that the organization will tend to learn late (i.e., revise its model of risk), that is, only after serious events occur. Effective organizations assume their model of risks and countermeasures is fragile and even seek out evidence that they should revise and update this model. To seek out such information means the organization is willing to expose its blemishes. They do not assume their model is correct and then wait for evidence to come to their attention (for to do so will guarantee an organization that acts riskier than it desires).

The drift toward failure in the Columbia accident is a process that retrenches and reinforces a mis-assessment that foam strikes do not pose risks to orbiter safety (but pose only a maintenance and turn around issue). The missed opportunities to revise and update the organization’s model of the riskiness of foam events seem to be consistent with what I have found in other cases of failure of foresight.

I have described the discounting of evidence as “distancing through differencing” whereby those reviewing new evidence or incidents (e.g, STS-27R) focus on differences, real and imagined, between the place, people, organization and circumstances where an incident happens and their own context. By focusing on the differences, they see no lessons for their own operation and practices or only narrow well bounded responses.

Ominously, this *distancing through differencing* that occurred throughout the build up to STS-107 can be repeated in the future as organizations and groups look at the analysis and lessons from this accident and the CAIB report). Others in the future can easily look at the CAIB conclusions and deny their relevance to their situation based on emphasizing differences (e.g., the technical topic is different, the managers are different more dedicated and careful about safety, the specific deficiency has been addressed). This is one reason avoiding hindsight bias is so important—when one starts with the question, how could they have missed what is now obvious—one is enabling future distancing through differencing rationalizations.

The distancing through differencing process that contributes to this breakdown also indicates ways to change the organization to promote learning. One general principle which could be put into action is—do not discard other events because they appear on the surface to be dissimilar. At some level of analysis, all events are unique; while at other levels of analysis, they reveal common patterns. To focus on common patterns not surface differences requires shifting the analysis of cases from surface characteristics to deeper patterns and more abstract dimensions. Each kind of contributor to an event then can guide the search for similarities.

To step back more broadly, NASA needs a mechanism to generate new evaluations that question NASA’s own model of the risks it faces and the countermeasures deployed. Such review and re-assessment can help NASA find places where it has underestimated the potential for trouble and revises its approach to create safety.

A quasi-independent group is needed to do this— independent enough to question the normal organizational decision making but involved enough to have a finger on the pulse of the organization (keeping statistics from afar is not enough to accomplish this).

3. A third genotype present is: *Fragmented problem solving process that is unable to see the big picture.*

Once again there are parallels in many other cases.¹ Even more important, NASA has its own successful contrast case in Mission Control's record of analyzing and handling anomalies.

Overall, I am struck by how over time, people and groups there was a fragmented view of what was known about the strike and its potential implications. There was no place, artifact, or person who had a complete and coherent view of the analysis of the foam strike event (note a coherent view includes understanding the gaps and uncertainties in the data or analysis to that point). Examples of fragmentation include that no one noted the varying estimates of size and changing categorization as in or out of family. It is striking that people used what looked like technical analyses to justify previously reached conclusions, instead of using technical analyses *to test tentative hypotheses* (e.g., CAIB, p. 126 1st column). People were making decisions about what does or does not pose risk on very shaky or absent technical grounds, and critically, *they couldn't see their decisions rested on shaky grounds* (e.g., the memos on p. 141, 142 illustrate the shallow, off hand assessments posing for and substituting for careful analysis).

The breakdown or absence of cross-checks is also striking. Cross checks on the rationale for decisions is a critical part of good organizational decision making. Yet no cross checks are in place to detect, question or challenge the specific flaws in the rationale, and no one noted that cross-checks were missing.

The engineers didn't recognize the big picture independent of the details provide by any analysis, analysis tool, or new images—the evidence already placed the situation outside the boundary conditions for safe operation. When you are forced to use as the only available analysis tool a model not designed to predict under these conditions, when the event is hundreds of times the scale of what the model is designed to handle, when the uncertainty bounds are so large and there is no way to reduce the uncertainty(email on p. 151-152), then it is basic engineering to realize that being outside the boundaries is not a good place to be. When being outside the analyzed boundaries is confused with not being confident enough to give a definitive answer, what has happened to engineering judgment—in this circumstance there is a very definitive answer, you don't go outside the boundaries and if events force you there, you better start getting data and analysis results fast.

¹ For an example see the first case discussed in Chapter 4 of Behind Human Error. D.D. Woods, L. Johannesen, R.I. Cook and N. Sarter. *Behind Human Error: Cognitive Systems, Computers and Hindsight*. Crew Systems Ergonomic Information and Analysis Center, WPAFB, Dayton OH, 1994. (order at <http://iac.dtic.mil/hsiac/SOARS.htm#Past>)

Seasoned aircraft pilots and ship commanders well understand need for this ability to capture the big picture and not to get lost in a series of details. The issue is how to train for this judgment (which is why it seems so odd that the groups that do train to do this hardly became involved in assessing the anomaly).

This finding leads to ideas for re-defining the kinds of anomalies to be practiced and who participates in those simulation training sessions.

Note that shrinking budgets lead to pressure to reduce training investments (the amount of practice, the quality of the simulated situations, and the number or breadth of people who go through the simulations sessions can all decline).

The value of such training depends critically on designing good sets of anomalous situations and how they unfold. It would be a critical for NASA's safety organization to lead, monitor and change these scenarios. Plus these simulated cases can be used to get information about risks and risk countermeasures that can be part of the proactive approach to making risk part of organizational decision making.

The fragmentation of problem solving also illustrates Weick's (1999) points that high reliability organizations exhibit a "deference to expertise", "reluctance to simplify interpretations", and "preoccupation with potential for failure" none of which were in operation in NASA's organizational decision making leading up to and during STS-107.

Safety organizations play an important role in ensuring that adequate technical grounds are established and used in decision making. They can hold operational units "accountable" for having appropriate technical basis. They can lead the process of defining norms for grounding risky decisions (including the decision that things are not risky) that will be critical for the future .

For example, judgments about risk should always be carried forward with the technical basis which was used to reach that judgment or conclusion. By including the basis and boundary conditions for the judgment with the decision, cross checks are enabled and rationalizing away warning signs lessened.

Given this information is bundled, the organization also needs a cooperative problem solving structure where different people or groups are brought into the process in a way that they can effectively challenge judgments on technical bases. Plus a review process is needed for groups to re-evaluate whether the judgment and its technical basis remains appropriate (or ever was appropriate) given new information or changing circumstances down the road.

4. The fourth genotype that appears in Columbia is a *Failure to revise assessments as new evidence accumulates*.

I first studied this pattern in nuclear power emergencies 20 plus years ago.² What was interesting in the data then and continues in more recent results is how difficult it is to revise a mis-assessment or to revise a once plausible assessment as new evidence comes in.

If revision only occurs when positive, clear cut new facts are salient, there is a grave risk of an organization acting too risky and finding out only from near misses, serious incidents, or actual harm. Instead, revising assessments of risks needs to be an ongoing process. The crux is to notice the information that changes past models of risk and calls into question the effectiveness of previous risk reduction actions without having to wait for clear cut positive evidence. What indicts NASA organizationally is that the correct diagnosis of production/safety tradeoffs and useful recommendations for organizational change were noted in 2000 yet followed much too slowly (The March 13 2000 Mars Climate Orbiter report nicely lays out how the pressure for production and to be 'better' on several dimensions led to management accepting riskier and riskier decisions and this report recommended organizational changes to make proactive risk part of day to day management).

Research consistently shows that revising assessments successfully requires a new way of looking at previous facts. We provide this "**fresh**" view:

- (a) by bringing in people new to the situation
- (b) through interactions across diverse groups with diverse knowledge and tools,
- (c) through new visualizations which capture the big picture and re-organize data into different perspectives.

One constructive action is to develop the collaborative inter-changes that generate fresh points of view or that produce challenges to basic assumptions (again a process that we have observed in the mission control's responses to anomalies).

One can also pursue creating data displays that help show where are the boundaries given past analyses and available data so that people can see when circumstances or organizational decisions are pushing the system to the edge of the envelope (this idea is something that Jens Rasmussen one of the pioneers of the new results on error and organizations has been pushing for 20 years³).

The newly empowered safety organization can play a key role in bringing these processes and tools into the day to day management process.

² D.D. Woods, J. O'Brien, and L.F. Hanes. Human factors challenges in process control: The case of nuclear power plants. In G. Salvendy, editor, *Handbook of Human Factors/Ergonomics*, Wiley, New York, 1987.

³ Rasmussen J. (1990). The role of error in organizing behavior. *Ergonomics*. 33: 1185-1199. Rasmussen, J. Risk Management, Adaptation, and Design for Safety. In B. Brehmer and N.-E. Sahlin (Eds.) *Future Risks and Risk Management*. Kluwer Academic, Dordrecht, 1994.

5. Finally, Columbia brings to the fore another genotype: *Breakdowns at the boundaries of organizational units.*

The Board notes how a kind of catch 22 was operating where the people charged to analyze the anomaly are unable to generate any definitive traction and where the management is trapped in a stance shaped by production pressure that sees such events as turn around issues. This seems to emerge only at boundaries of different organizations that do not have mechanisms for constructive interplay. It is here that we see the operation of Weick's generalization that in risky judgments we have to defer to those with technical expertise (assuming we can get those resources engaged in the event).

This pattern points to the need for mechanisms for creating effective overlap across different organizational units and to avoid simply staying inside the chain of command mentality (though such overlap can be seen as inefficiencies when the organization is under severe cost pressure).

This issue is of particular concern to many organizations since communication technology has linked together disparate groups as a distributed team. This capability for connectivity is leading many to work on how to support effective coordination across these distributed groups (e.g., in military command and control).

Two Contrasts in NASA Experience

For NASA parallels on safety/production tradeoff, the accidents reports on the Mars series of failures are quite illuminating. The March 13 2000 Mars Climate Orbiter report nicely lays out how the pressure for production and to be 'better' on several dimensions lead to management accepting riskier and riskier decisions. The contrast is the FBC Task Final Report (of March 2000) which is an apologia for the administrator's Faster, Better, Cheaper policy and unintentionally captures the tradeoff or double bind of maintaining high safety (mission success in this case) while under production pressure.

For NASA parallels on coordinated anomaly response, the contrast is with how mission control works successfully. In Mission Control problem solving there are clear lines of responsibility to have a complete, coherent view of the evolving analysis vested in the relevant flight controllers and in the flight director. Mission control also has mechanisms to keep different people in the loop (via monitoring voice loops, for example) so that all are up to date on the current picture of situation. Mission control also has mechanisms for correcting assessments as analysis proceeds whereas in this case, the fragmentation and partial views seemed to lead to less error correction.⁴

⁴ For studies of how mission control handles anomalies see E.S. Patterson, J.C. Watts-Perotti, D.D. Woods. Voice Loops as Coordination Aids in Space Shuttle Mission Control. *Computer Supported Cooperative Work*, 8, 353—371, 1999, and J. Watts Perotti and D.D. Woods. A Cognitive Analysis of Anomaly Response in Space Shuttle Mission Control. CSEL 97-TR-02, The Ohio State University, Columbus OH, March 1997. Prepared for NASA Johnson Space Center.

How to bring safety into every day organizational decision making?

Understanding how contributors to the Columbia accident represent general patterns of breakdown in complex systems can help guide organizational change and empower safety in NASA and other organizations when economics and schedules are tight.