

AC 2008-539: THE LOSS OF THE SPACE SHUTTLE COLUMBIA: PORTAGING THE LEADERSHIP LESSONS WITH A CRITICAL THINKING MODEL

Robert Niewoehner, U.S. Naval Academy

CAPT Rob Niewoehner, USN, PhD is Director of Aeronautics at the US Naval Academy. Prior to joining the Naval Academy faculty, he served as a fleet F-14 pilot, and then as an experimental test pilot, including Chief Test Pilot for the F/A-18 E/F Super Hornet, throughout its development.

Craig Steidle, US Naval Academy

Rear Admiral Craig Steidle, USN (ret) holds the Rogers Chair of Aeronautics at the U.S. Naval Academy. In uniform, RADM Steidle served as a combat A-6 pilot, test pilot, F/A-18 Program Manager, F-35 Program Manager, and Vice Commander of the Naval Air Systems Command. He subsequently served as Associate Administrator of NASA for Deep Space Exploration.

Eric Johnson, US Naval Academy

Midshipman, US Navy. Midn Johnson is a graduating senior.

The Loss of the *Space Shuttle Columbia*: Portaging the Leadership Lessons with a Critical Thinking Model

Abstract

Business schools have long valued case studies as a tool for both broadening a student's perspective, and provoking them to deeper consideration of complex situations. The challenge with case studies is assuring the portability of the lessons; we don't expect students to see situations imitating those they've studied, hence the goal must instead be habits of mind and principles of action which the student can portage to the circumstances of their professional lives. This paper evaluates the suitability of Richard Paul's Critical Thinking model as a template for evaluating engineering enterprise thinking habits and organizational behavior, using the *Columbia* Accident Investigation Board (CAIB) report¹ as a case study. Specifically, the authors sought to answer the following questions: "Does the Paul model of Critical Thinking provide a beneficial vocabulary and construct for evaluating complex technological case studies?" and, "Does the structure of Paul's model enhance the portability of the lessons?" The authors determined that with minor refinement, Paul's model provides a powerful vocabulary for complicated case study analysis, and that familiarity with the model provides students and case study participants with both a mechanism for analysis and a means for portaging lessons to other professional situations and organizations.

Introduction

The analysis and evaluation of our thinking as engineers requires a vocabulary of thinking and reasoning. The intellect requires a voice. Richard Paul and his colleague, Linda Elder, from the Foundation for Critical Thinking, have proposed a critical thinking model documented in various sources^{2,3,4}, including over a dozen *Thinkers' Guides* that apply this model to diverse disciplines including science⁵ and ethics⁶. Their *Thinker's Guide to Engineering Reasoning* specifically adapts Paul's model to the intellectual work of engineers, exemplifying the questions that experienced engineers ask of themselves and others.

In the paper below, the Paul model is first summarized using definitions drawn from Paul, Niewoehner, and Elder's *Engineering Reasoning*.⁷ This summary includes brief discussions of our approach in introducing the Paul model and vocabulary to this class of students. Next, the findings of the CAIB report are summarized for those who are not familiar with its contents. Importantly, we do not seek to re-analyze the CAIB's findings or recommendations, nor further excoriate those whose mistakes may have contributed to the mishap. We cannot improve on what we regard as a masterful contribution to the literature describing high technology organizations. No, it is instead the Paul model which is under examination. Our question was solely whether the Paul model was adequate to the purpose of opening the CAIB report and its complexities to our undergraduate students in ways that they could retain and apply.

A Critical Thinking Model For Engineering

Engineers and scientists are quite comfortable working within the context of conceptual models. We employ thermodynamic models, electrical models, mathematical models, computer models or even physical models fashioned from wood or clay. Paul, Niewoehner and Elder apply

a model to the way in which engineers think, an architecture whose purpose is aiding the analysis and evaluation of thought, that we might improve our thought.

The analysis and evaluation of our thinking as engineers requires a vocabulary of thinking and reasoning. The model that follows is not unique to engineering; indeed, its real power is its portability, adapting to *any* domain of life and thought. In so far as an engineer masters the rudimentary skills of critical thinking in the context of engineering, they have really appropriated the skills of life-long learning for whatever domain of learning their professional and personal lives lead them.

We need a definition of Critical Thinking. Several would serve our purposes, but we are particularly fond of David Moore's:

“Critical Thinking is a deliberate meta-cognitive (*thinking about thinking*) and cognitive (*thinking*) act whereby a person reflects on the quality of the reasoning process simultaneously while reasoning to a conclusion. The thinker has two equally important goals: coming to a solution and improving the way she or he reasons.”⁸[italics in original]

The metacognitive piece of Moore's definition is vital. If we consider many technical systems today, such as an airplane, the health of the aircraft's systems are constantly monitored. The flight control computers do not simply process the next surface deflection, they are also constantly asking themselves, “Am I healthy?” Critical thinking includes not only the process by which we gain a conclusion, it is also the process by which we ask, in parallel, “Is my thinking healthy?” This

Figure 1 depicts Paul's model. The goal, at the bottom, is the mature *thinker*, whose thinking skills and ethical dispositions act in concert, as evidenced by intellectual traits/virtues. The *Elements of Thought* comprise the tools by which we *analyze* intellectual work, our own and others, taking it apart to understand its constituent parts. *Intellectual Standards* are the criteria against which we *evaluate* the quality of intellectual work. Specifically, the model identifies the vital *questions* we should be asking ourselves and others. *It's all about the questions!*

Effective Teams Manifest Intellectual Traits/Virtues

The engineer does not work in isolation, but in the context of enterprises, cultures and communities, each of which represents divergent interests and perspectives. Furthermore, no engineer can claim perfect objectivity; their work is unavoidably influenced by strengths and weaknesses, education, experiences, attitudes, beliefs, and self-interest. They avoid paths they associate with past mistakes and trudge down well worn paths that worked in the past. The profession engineer must cultivate personal and intellectual virtues. The leader must both model and foster these traits with those they lead.

These virtues are not radically distinct from those sought by any maturing thinker, regardless of the discipline. They determine the extent to which we think with insight and integrity, regardless of the subject. The engineering enterprise does however pose distinct questions for the engineer in pursuit of such virtue.

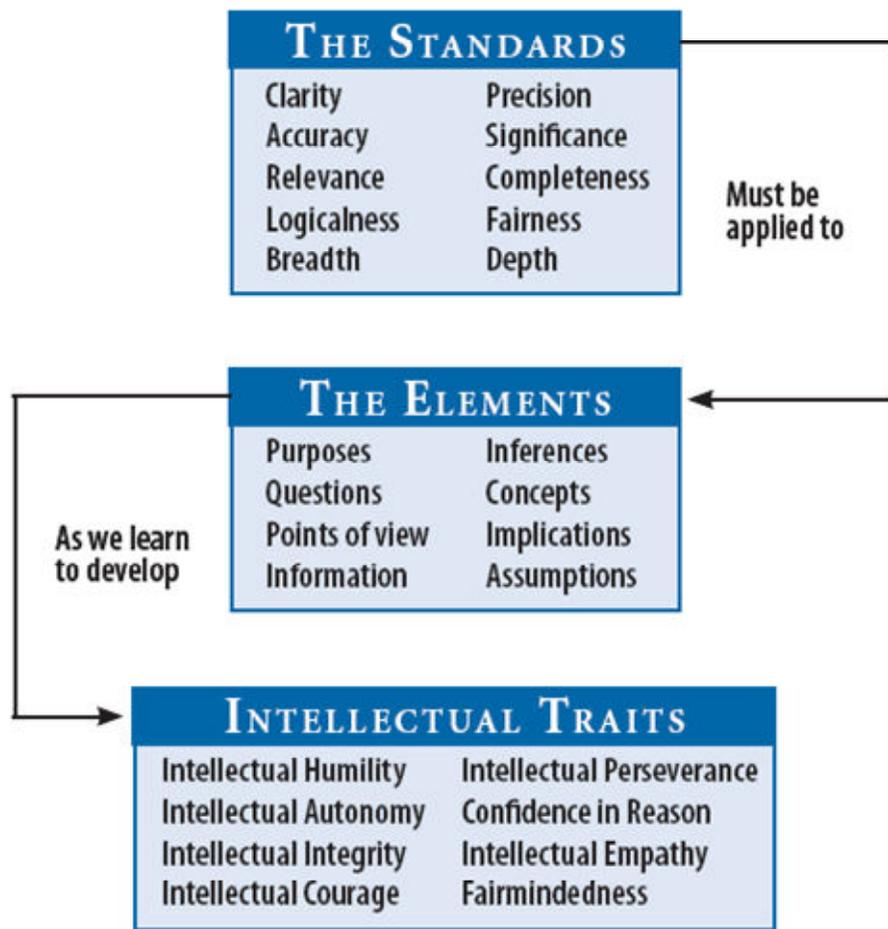


Figure 1: Richard Paul's Critical Thinking Model.⁹

- *Intellectual humility* admits to ignorance, frankly sensitive to what you know and what you do not know. It implies being aware of your biases, prejudices, self-deceptive tendencies and the limitations of your viewpoint and experience.
- *Intellectual courage* is the disposition to question beliefs about which we feel strongly. It includes questioning the beliefs of our enterprise culture and any sub-culture to which we belong, and a willingness to express our views even when they are unpopular (with management, peers, subordinates or customers).
- *Intellectual empathy* is awareness of the need to actively entertain views that differ from our own, especially those with which we strongly disagree. It entails accurately reconstructing others' viewpoints and to self-consciously reason from premises, assumptions, and ideas other than our own.
- *Intellectual integrity* consists in holding ourselves to the same intellectual standards you expect others to honor (no double standards).
- *Intellectual perseverance* is the disposition to work our way through intellectual complexities despite the frustration inherent in the task.

- *Confidence in reason* is based on the belief that one's own higher interests and those of humankind at large are best served by giving the freest play to reason. It means using standards of reasonability as the fundamental criteria by which to judge whether to accept or reject any proposition or position.
- *Intellectual autonomy* is thinking for oneself while adhering to standards of rationality. It means thinking through issues using one's own thinking rather than uncritically accepting the viewpoints, opinions and judgments of others.
- *Fairmindedness* is being conscious of the need to treat all viewpoints alike, without reference to one's own feelings or vested interests, or the feelings or vested interests of one's friends, company, community or nation; implies adherence to intellectual standards without reference to one's own advantage or the advantage of one's group.
- *Intellectual Curiosity* motivates intellectual perseverance (above), and manifests itself as discontentment with unanswered questions. *Curiosity* does not explicitly appear in Paul's lists of intellectual virtues, though tacitly lauded in several of Paul and Elder's papers. We include it here because it explicitly appears in the CAIB report multiple times, and none of the other traits above adequately capture this vital trait.

The intellectual traits/virtues were introduced to students in a Technical Leadership seminar using a workshop format. Individuals within groups of 3-4 were assigned a trait which they then studied briefly from the *Engineering Reasoning Guide*¹⁰ and then explained to their teammates. Successful rounds of this reciprocal teaching were conducted until the list of traits was covered. Students were then asked to write down a vignette illustrating how they had personally witnessed the positive contribution of one of the traits to a team on which they'd served, and likewise one vignette exemplifying how a deficit in one trait had adversely affected a team. The entire class was then polled to nominate particularly noteworthy stories for the entire class. We've conducted similar workshops on this topic with undergraduate seniors in several contexts. By the time they're twenty, students have no shortage of applicable experiences from which to draw, whether athletic, academic, or extra-curricular, exemplifying virtue's relevance.

All Thinking Builds Upon Eight Fundamental Elements

All thinking entails eight fundamental elements, whether it is about engineering, philosophy, cooking, sports, or business. These eight elements express eight questions that we can pose about any intellectual activity or subject. The eight elements, and their use in analyzing a document, were introduced by asking students to write out the purpose, point of view, data, etc. for the CAIB report.¹¹ These were then discussed Socratically as a class. The below summarizes/paraphrases students' responses. Note that these questions and this activity work with *any* topic in *any* field.

Q- What was the **purpose** of the CAIB?

A- The CAIB sought to identify the causes of the *Columbia's* loss and recommend actions for the resumption of U.S space flight activity.

Q- What **questions** did the CAIB principally try to answer?

A- What caused the loss of Columbia? What contributory factors may have been present? What actions should NASA and the U.S. government take in the future to reduce the likelihood of future mishaps.

Q- *What point of view did the CAIB represent?*

A- The CAIB was composed of senior engineers and leaders representing the military, government, academia, and industry. The report acknowledged other points of view, including the NASA workforce and astronaut office, the U.S. Congress, the aerospace industry.

Q- *What did the CAIB assume?*

A- All accident investigations take for granted that all accidents have causal factors traceable to both physical and cultural factors, and that understanding those factors can lead to improved safety in future operations. Additionally, the failures of complex systems are commonly traced to the complex interaction of many cultural and technological features surrounding that system. From the outset, the CAIB assumed that the answers wouldn't be simple. Additionally, they assumed that their recommendations would be taken seriously and would form the basis for both a return to flight and the future vitality of U.S. space activities.

Q- *What information did the CAIB report?*

A- The CAIB report is very expansive in the nature of the information reported. It describes the history of the Space Shuttle Program, including the varying political/budgetary climates in which it was conceived and operated over 30 years time. Additionally, it reports specific technical details of the *Columbia's* last flight and data from other previous flights bearing on the incident. It includes detailed transcripts of relevant team interactions (meetings, presentations, email) during the months leading to the accident. It analyzes the results of experiments conducted by the board to better understand the failure mechanism. Finally, the report details over a hundred pertinent "findings" and several dozen recommendations.

Q- *What are the most significant concepts upon which the report rests?*

A- The span of the report is very, very broad, including U.S. space policy and spending, program management, materials science, organizational behavior, government/contractor relations, flight mechanics, among many others. Particularly important concepts include risk management and accepted risk, failure trees, organizational behavior, safety, and leadership.

Q- *What did the CAIB conclude?*

A- The CAIB concluded that the shuttle's loss was directly attributable to a breach in the left wing, caused by foam shed from the external tank during the shuttle's ascent. That breach allowed a hot jet of air into the left wing's structure which burned through the structure, causing its failure. Tragically, the loss of foam was acknowledged by NASA as a persistent problem, but not viewed as a threat to an orbiter's safety. Consequently, the board concluded that the accident was attributable as much to poor organizational and leadership practices, as it was to foam. "It is the view of the Columbia Accident

Investigation Board that the *Columbia* accident is not a random event, but rather a product of the Space Shuttle Program's history and current management processes."¹²

Q- *What are the implications of the CAIB?*

A- The CAIB provided a foundation for the return to shuttle service two years after the publication of their report, reestablishing U.S. confidence in manned space flight, and providing the means for resumption of the International Space Station's construction.

Engineering Reasoning Applies Intellectual Standards

Universal intellectual standards must be applied to thinking whenever one is interested in checking the quality of reasoning about a problem, issue, or situation. To think professionally as an engineer entails having command of these standards. The standards are not unique to engineering, but are universal to all domains of thinking. They may however have particular meaning or significance which is contextual or disciplinary. While there are a number of universal standards, we focus here on some of the most significant to engineering. Unlike the elements above, this list is not necessarily comprehensive and lists found in Paul's work do not always agree in detail.

Importantly, students must be explicitly introduced to the notion of intellectual standards. High school and undergraduate students seem to recognize only two standards: "Did I get the right answer?" and "Am I done?" Defining intellectual standards, and helping students see that they are universal, helps them understand that good intellectual work is characterized by more than the *right* answer.

Clarity

Clarity is the gateway standard. If a statement is unclear, we cannot determine whether it is accurate or relevant. In fact, we cannot tell anything about it because we don't yet know what it is saying. "Could you elaborate further on that point?" "Could you express that point in another way?" "Could you give me an illustration or example?"

Accuracy

A statement can be clear but not accurate, as in "Most creatures with a spine are over 300 pounds in weight." "Is that really true?" "How could we check that?" "How could we find out if that is true?" "What is your confidence in that data?"

Precision

A statement can be both clear and accurate, but not precise, as in "The solution in the beaker is hot." (We don't know how hot it is. "Could you give me more details?" "Could you be more specific?") Engineers commonly express precision in quantitative terms associated with the calibration of our instrumentation. We can't lose sight however that precision is also qualitative, bearing on the precision of our prose.

Relevance

A statement can be clear, accurate, and precise, but not relevant to the question at issue. A technical report might mention the time of day and phase of the moon at which the test was

conducted. This would be relevant if the system under test was a night vision device. It would be irrelevant if it had been a microwave oven. "How is that connected to the question?" "How does that bear on the issue?"

Depth

A statement can be clear, accurate, precise, and relevant, but superficial. For example, the statement "Radioactive waste from nuclear reactors threatens the environment," is clear, accurate, and relevant. Nevertheless, it lacks depth because it treats an extremely complex issue superficially. (It also lacks precision.) "How does your analysis address the complexities in the question?"

Breadth

A line of reasoning may be clear, accurate, precise, relevant, and deep, but lack breadth (as in an argument from either of two conflicting theories, both consistent with available evidence). Broad thinking suggests questions such as: "Do we need to consider another point of view?" "Is there another way to look at this question?" "What would this look like from the point of view of a conflicting theory, hypothesis or conceptual scheme?"

Logical Validity

When we think, we bring a variety of thoughts together into some order. The thinking is "logical" when the conclusion follows from the supporting data or propositions. The conclusion is "illogical" when it contradicts proffered evidence, or the arguments fail to cohere." Does this really make sense?" "How does that follow from what you said?" "But before you implied this and now you are saying that, I don't see how both can be true."

Fairness

Fairness is particularly at play where either a problem has multiple approaches (conflicting conceptual systems), or conflicting interests among stake-holders. Fairness gives all perspectives a voice, while recognizing that all perspectives may not be accurate or equally valuable.

The following three standards are not found either in Figure 1 above, nor in Paul and Elder's writing. We have included them in our teaching because they have frequently caught our attention as defects in the work of our undergraduates.

Concision

The days are well past when great oratory meant hours, or great literature necessarily included chapter-long depictions of the field at Waterloo or the implements of the New England Whaling trade. Abraham Lincoln was derided for demeaning the fallen by his brevity at Gettysburg; his partner on the podium later confessed that the President had said more in several minutes than he had said in an hour. Concision does not connote short for brevity's sake (the sound bite), but rather an economy of thoughts whereby the thinking is deep and significant, and clarity is enhanced by the economy of words and/or images. In the hours building to the loss of the Space Shuttle *Challenger*, engineers understood the peril faced by launching at extremely low temperatures. Yet, they buried their management in insignificant detail such that their message was missed; their signal was obscured by self-generated noise.¹³

Suitability-

Suitability applies largely to our written and oral communications, seeking to be “fitting”, “appropriate”, or “suited to the purpose.” Suitability entails selecting right tone and presentation for the audience. It is seldom easy to craft our speech or writing to squarely address the interests, knowledge, and abilities of our audience/readers.

The *Columbia* Accident Investigation Board (CAIB) Report

The general facts surrounding the loss of the space shuttle *Columbia* on the morning of February 1, 2003, are well known. A piece of insulating foam broke away from the external fuel tank seconds after launch, puncturing the leading edge of the orbiter’s left wing. The crew then spent fifteen days on orbit conducting a host of very successful science experiments, unaware that their spacecraft had been catastrophically damaged. On re-entry, hot gas tore through the interior structure of the wing, leading to wing failure, disintegration of the vehicle, and the death of the crew.

Unfortunately, the board’s findings on organizational behavior have not been as broadly discussed. The technical story is fascinating; the CAIB’s discussion of organizational behavior is heart-rending. The real meat lies here for those who lead or will lead technical organizations, because it’s a tragic story of bright, devoted, hard-working professionals whose leaders allowed the team’s thinking to stray adrift, killing seven of their friends and scattering an irreplaceable national asset across the Southwestern United States. We regard the CAIB report as required reading for all leaders in high technology enterprises not because of what they might learn about the threat of insulating foam to spacecraft, but rather the threat that uncritical thinking poses to even the nation’s most successful, talented and hard-working teams. We’ll briefly summarize the organizational piece for those unfamiliar with this second most disturbing facet of the *Columbia* mishap.

The CAIB’s most severe criticism of NASA sprang from their observation of the strong similarity between the loss of *Columbia* and the loss of *Challenger*. Neither the loss of foam (*Columbia*), nor O-ring erosion (*Challenger*), were new issues; both had been observed on numerous prior flights. In both mishaps, technical team members raised grave concerns about the safety of the mission during the week prior to each orbiter’s loss. In both events, leadership dismissed team member concerns, focused on keeping the schedule, and blithely inferred that past minor issues with O-rings/foam would remain minor. The “echoes of Challenger” led the CAIB to entitle an entire chapter, “History as Cause: *Columbia* and *Challenger*.”¹⁴

Surely in these grand tragedies we have the grist of poignant lessons for future leaders. Our issue as engineering educators is modeling consideration of the board’s findings in such a way that students can extract lessons about how to think about thinking in organizational contexts, rather than simply reiterating criticism of the actors’ mistakes.

The pages that follow are extracted directly from the CAIB Report, Chapter 6, “Decision Making at NASA.” They summarize a very lengthy section 6.3, “Decision-Making During the Flight of STS-107,” which detailed the substance of multiple meetings and extensive correspondence within and between program teams as decisions were made regarding the

condition of *Columbia* during its final mission. The left column is verbatim from the report; our italicized remarks are to the right note, with vocabulary from Paul's model underlined.

We've chosen this section for emphasis because it describes the dysfunction of a specific team, involving small meetings and personal communications, rather than the report's broader treatment of the dysfunction of an entire agency or U.S. space policy. The team setting is more accessible to the undergraduate who can more readily imagine themselves in a team setting than executive management, and it is for that setting that we seek to first prepare them.

Summary: Mission Management Decision Making ¹⁵

Discovery and Initial Analysis of Debris Strike

In the course of examining film and video images of Columbia's ascent, the Intercenter Photo Working Group identified, on the day after launch, a large debris strike to the leading edge of Columbia's left wing. Alarmed at seeing so severe a hit so late in ascent, and at not having a clear view of damage the strike might have caused, Intercenter Photo Working Group members alerted senior Program managers by phone and sent a digitized clip of the strike to hundreds of NASA personnel via e-mail. These actions initiated a contingency plan that brought together an interdisciplinary group of experts from NASA, Boeing, and the United Space Alliance to analyze the strike. So concerned were Intercenter Photo Working Group personnel that on the day they discovered the debris strike, they tapped their Chair, Bob Page, to see through a request to image the left wing with Department of Defense assets in anticipation of analysts needing these images to better determine potential damage. By the Board's count, this would be the first of three requests to secure imagery of Columbia on-orbit during the 16-day mission.

Clear recognition of the need for better data.

Upon learning of the debris strike on Flight Day Two, the responsible system area manager from United Space Alliance and her NASA counterpart formed a team to analyze the debris strike in accordance with mission rules requiring the careful examination of any "out-of-family" event. Using film from the Intercenter Photo Working Group, Boeing systems integration analysts prepared a preliminary analysis that afternoon. (Initial estimates of debris size and speed, origin of debris, and point of impact would later prove remarkably accurate.)

Excellent initial inferences based upon scant preliminary data. "out-of-family" meant out of NASA's experience base.

As Flight Day Three and Four unfolded over the Martin Luther King Jr. holiday weekend, engineers began their analysis. One Boeing analyst used Crater, a mathematical prediction tool, to assess possible damage to the Thermal Protection System. Analysis predicted tile damage deeper than the actual tile depth, and penetration of the RCC coating at impact angles above 15 degrees. This suggested the potential for a burn-through during re-entry. Debris Assessment Team members judged that the actual damage would not be as severe as predicted because of the inherent conservatism in the Crater model and because, in the case of tile, Crater does not take into account the tile's stronger and more impact-resistant "densified" layer, and in the case of RCC, the lower density of foam would preclude penetration at impact angles under 21 degrees.

Gut-based judgment replaces engineering analysis. Inaccurate inference based on invalid logic, and unsubstantiated assumptions.

On Flight Day Five, impact assessment results for tile and RCC were presented at an informal meeting of the Debris Assessment Team, which was operating without direct Shuttle Program or Mission Management leadership. Mission Control's engineering support, the Mission Evaluation Room, provided no direction for team activities other than to request the team's results by January 24. As the problem was being worked,

Unchallenged working assumptions. Conspicuous lack of intellectual curiosity on the part of leadership.

Shuttle managers did not formally direct the actions of or consult with Debris Assessment Team leaders about the team's assumptions, uncertainties, progress, or interim results, an unusual circumstance given that NASA managers are normally engaged in analyzing what they view as problems. At this meeting, participants agreed that an image of the area of the wing in question was essential to refine their analysis and reduce the uncertainties in their damage assessment.

Some team-members continued to recognize the inadequacy of the data.

Each member supported the idea to seek imagery from an outside source. Due in part to a lack of guidance from the Mission Management Team or Mission Evaluation Room managers, the Debris Assessment Team chose an unconventional route for its request. Rather than working the request up the normal chain of command – through the Mission Evaluation Room to the Mission Management Team for action to Mission Control – team members nominated Rodney Rocha, the team's Co-Chair, to pursue the request through the Engineering Directorate at Johnson Space Center. As a result, even after the accident the Debris Assessment Team's request was viewed by Shuttle Program managers as a non-critical engineering desire rather than a critical operational need.

Insufficient clarity regarding the extent of team-member's discomfort with lack of imagery (data).

When the team learned that the Mission Management Team was not pursuing on-orbit imaging, members were concerned. What Debris Assessment Team members did not realize was the negative response from the Program was not necessarily a direct and final response to their official request. Rather, the "no" was in part a response to requests for imagery initiated by the Intercenter Photo Working Group at Kennedy on Flight Day 2 in anticipation of analysts' needs that had become by Flight Day 6 an actual engineering request by the Debris Assessment Team, made informally through Bob White to Lambert Austin, and formally through Rodney Rocha's e-mail to Paul Shack. Even after learning that the Shuttle Program was not going to provide the team with imagery, some members sought information on how to obtain it anyway.

Leadership canceled photo requests because:
a) inaccurate assumptions of the imaging capability,
b) inaccurate assumptions regarding value of photos
d) unwillingness to disrupt mission to inspect orbiter (confused purpose)¹⁶
e) inaccurate assumption that rescue was infeasible.¹⁷
These assumptions were accepted as fact.

Debris Assessment Team members believed that imaging of potentially damaged areas was necessary even after the January 24, Mission Management Team meeting, where they had reported their results. Why they did not directly approach Shuttle Program managers and share their concern and uncertainty, and why Shuttle Program managers claimed to be isolated from engineers, are points that the Board labored to understand. Several reasons for this communications failure relate to NASA's internal culture and the climate established by Shuttle Program management, which are discussed in more detail in Chapters 7 and 8.

Some perseverance displayed by those willing to circumvent bureaucratic obstacles.

A Flawed Analysis

An inexperienced team, using a mathematical tool that was not designed to assess an impact of this estimated size, performed the analysis of the potential effect of the debris impact. Crater was designed for "in-family" impact events and was intended for day-of-launch analysis of debris impacts. It was not intended for large projectiles like those

Other parts of the report attribute this behavior to lack of intellectual courage on the part of team-members, and lack of empathy on the part of management.

Inaccurate conclusions based on unjustified extrapolation of assumptions. The tool's

observed on STS-107. Crater initially predicted possible damage, but the Debris Assessment Team assumed, without theoretical or experimental validation, that because Crater is a conservative tool – that is, it predicts more damage than will actually occur – the debris would stop at the tile’s densified layer, even though their experience did not involve debris strikes as large as STS-107’s. Crater-like equations were also used as part of the analysis to assess potential impact damage to the wing leading edge RCC. Again, the tool was used for something other than that for which it was designed; again, it predicted possible penetration; and again, the Debris Assessment Team used engineering arguments and their experience to discount the results.

severe predictions were dismissed not on the basis of logic, but on a history which showed that foam had never previously been a safety of flight issue.

As a result of a transition of responsibility for Crater analysis from the Boeing Huntington Beach facility to the Houston-based Boeing office, the team that conducted the Crater analyses had been formed fairly recently, and therefore could be considered less experienced when compared with the more senior Huntington Beach analysts. In fact, STS-107 was the first mission for which they were solely responsible for providing analysis with the Crater tool. Though post-accident interviews suggested that the training for the Houston Boeing analysts was of high quality and adequate in substance and duration, communications and theoretical understandings of the Crater model among the Houston-based team members had not yet developed to the standard of a more senior team. Due in part to contractual arrangements related to the transition, the Houston-based team did not take full advantage of the Huntington Beach engineers’ experience.

A new support team failed to admit when they were over their heads (Intellectual humility).

At the January 24, Mission Management Team meeting at which the “no safety-of-flight” conclusion was presented, there was little engineering discussion about the assumptions made, and how the results would differ if other assumptions were used.

*Unchallenged assumptions.
Lack of intellectual curiosity.*

Engineering solutions presented to management should have included a quantifiable range of uncertainty and risk analysis. Those types of tools were readily available, routinely used, and would have helped management understand the risk involved in the decision. Management, in turn, should have demanded such information. The very absence of a clear and open discussion of uncertainties and assumptions in the analysis presented should have caused management to probe further.

*Imprecise information.
Inadequate intellectual perseverance and curiosity.*

Shuttle Program Management’s Low Level of Concern

While the debris strike was well outside the activities covered by normal mission flight rules, Mission Management Team members and Shuttle Program managers did not treat the debris strike as an issue that required operational action by Mission Control. Program managers, from Ron Dittmore to individual Mission Management Team members, had, over the course of the Space Shuttle Program, gradually become inured to External Tank foam losses and on a fundamental level did not believe foam striking the vehicle posed a critical threat to the Orbiter. In particular, Shuttle managers exhibited a belief that RCC panels are impervious to foam impacts. Even after seeing the video of *Columbia*’s debris impact, learning estimates of the size and location of the strike, and noting that a foam strike with sufficient kinetic energy could cause Thermal Protection System damage, management’s level of concern did not change.

Insufficient intellectual perseverance and curiosity.

The opinions of Shuttle Program managers and debris and photo analysts on the potential severity of the debris strike diverged early in the mission and continued to diverge as the mission progressed, making it increasingly difficult for the Debris Assessment Team to have their concerns heard by those in a decision-making capacity. In the face of Mission managers’ low level of concern and desire to get on with the mission, Debris Assessment Team members had to prove unequivocally that a safety-of-

*Insufficient intellectual fairness.
Confused purpose
(emphasis was justifying the safety of the next mission in lieu*

flight issue existed before Shuttle Program management would move to obtain images of the left wing. The engineers found themselves in the unusual position of having to prove that the situation was unsafe – a reversal of the usual requirement to prove that a situation is safe.

of recovering the current mission).

Other factors contributed to Mission management’s ability to resist the Debris Assessment Team’s concerns. A tile expert told managers during frequent consultations that strike damage was only a maintenance-level concern and that on-orbit imaging of potential wing damage was not necessary. Mission management welcomed this opinion and sought no others. This constant reinforcement of managers’ pre-existing beliefs added another block to the wall between decision makers and concerned engineers.

*Sociocentric blindness.
No breadth of inquiry.
No cultivation of dissenting points of view.*

Another factor that enabled Mission management’s detachment from the concerns of their own engineers is rooted in the culture of NASA itself. The Board observed an unofficial hierarchy among NASA programs and directorates that hindered the flow of communications. The effects of this unofficial hierarchy are seen in the attitude that members of the Debris Assessment Team held. Part of the reason they chose the institutional route for their imagery request was that without direction from the Mission Evaluation Room and Mission Management Team, they felt more comfortable with their own chain of command, which was outside the Shuttle Program. Further, when asked by investigators why they were not more vocal about their concerns, Debris Assessment Team members opined that by raising contrary points of view about Shuttle mission safety, they would be singled out for possible ridicule by their peers and managers.

Insufficient intellectual courage.

A Lack of Clear Communication

Communication did not flow effectively up to or down from Program managers. As it became clear during the mission that managers were not as concerned as others about the danger of the foam strike, the ability of engineers to challenge those beliefs greatly diminished. Managers’ tendency to accept opinions that agree with their own dams the flow of effective communications.

*Sociocentric blindness.
No cultivation of dissenting points of view.*

After the accident, Program managers stated privately and publicly that if engineers had a safety concern, they were obligated to communicate their concerns to management. Managers did not seem to understand that as leaders they had a corresponding and perhaps greater obligation to create viable routes for the engineering community to express their views and receive information. This barrier to communications not only blocked the flow of information to managers, but it also prevented the downstream flow of information from managers to engineers, leaving Debris Assessment Team members no basis for understanding the reasoning behind Mission Management Team decisions.

Deficient Intellectual Fairness/Empathy

The January 27 to January 31, phone and e-mail exchanges, primarily between NASA engineers at Langley and Johnson, illustrate another symptom of the “cultural fence” that impairs open communications between mission managers and working engineers. These exchanges and the reaction to them indicated that during the evaluation of a mission contingency, the Mission Management Team failed to disseminate information to all system and technology experts who could be consulted. Issues raised by two Langley and Johnson engineers led to the development of “what-if” landing scenarios of the potential outcome if the main landing gear door sustained damaged. This led to behind-the-scenes networking by these engineers to use NASA facilities to make simulation runs of a compromised landing configuration. These engineers – who understood their systems and related technology – saw the potential for a problem on landing and ran it down in case the unthinkable occurred. But their concerns never reached the managers on the Mission Management Team that had operational control over Columbia

Here’s a team that showed perseverance, running their questions to ground by end-running the bureaucracy. Their ad hoc study simulating landing with a blown tire showed the crew would survive, so they allayed their own concern.

A Lack of Effective Leadership

The Shuttle Program, the Mission Management Team, and through it the Mission Evaluation Room, were not actively directing the efforts of the Debris Assessment Team. These management teams were not engaged in scenario selection or discussions of assumptions and did not actively seek status, inputs, or even preliminary results from the individuals charged with analyzing the debris strike. They did not investigate the value of imagery, did not intervene to consult the more experienced Crater analysts at Boeing's Huntington Beach facility, did not probe the assumptions of the Debris Assessment Team's analysis, and did not consider actions to mitigate the effects of the damage on re-entry. Managers' claims that they didn't hear the engineers' concerns were due in part to their not asking or listening.

This is a catalog of what's already been said.

The Failure of Safety's Role

As will be discussed in Chapter 7, safety personnel were present but passive and did not serve as a channel for the voicing of concerns or dissenting views. Safety representatives attended meetings of the Debris Assessment Team, Mission Evaluation Room, and Mission Management Team, but were merely party to the analysis process and conclusions instead of an independent source of questions and challenges. Safety contractors in the Mission Evaluation Room were only marginally aware of the debris strike analysis. One contractor did question the Debris Assessment Team safety representative about the analysis and was told that it was adequate. No additional inquiries were made. The highest-ranking safety representative at NASA headquarters deferred to Program managers when asked for an opinion on imaging of Columbia. The safety manager he spoke to also failed to follow up.

Deficient Intellectual Courage, Curiosity, and Perseverance.

Summary

Management decisions made during *Columbia's* final flight reflect missed opportunities, blocked or ineffective communications channels, flawed analysis, and ineffective leadership. **Perhaps most striking is the fact that management – including Shuttle Program, Mission Management Team, Mission Evaluation Room, and Flight Director and Mission Control – displayed no interest in understanding a problem and its implications.** Because managers failed to avail themselves of the wide range of expertise and opinion necessary to achieve the best answer to the debris strike question – “Was this a safety-of-flight concern?” – some Space Shuttle Program managers failed to fulfill the implicit contract to do whatever is possible to ensure the safety of the crew. In fact, their management techniques unknowingly imposed barriers that kept at bay both engineering concerns and dissenting views, and ultimately helped create “blind spots” that prevented them from seeing the danger the foam strike posed.

The most damning line in the report expresses dismay at the want of intellectual curiosity regarding implications [Emphasis added.]

The real tragedy- the Point of View of the crew and their families didn't intrude (Intellectual Empathy and Fairness). The focus on keeping the program schedule (a confused purpose) trumped ensuring the safety of the mission in progress.

Because this chapter has focused on key personnel who participated in STS-107 bipod foam debris strike decisions, it is tempting to conclude that replacing them will solve all NASA's problems. However, solving NASA's problems is not quite so easily achieved. Peoples' actions are influenced by the organizations in which they work, shaping their choices in directions that even they may not realize. The Board explores the organizational context of decision making more fully in Chapters 7 and 8.

Here the board hints at implications of their findings, yet to be discussed.

Throughout the above, vocabulary from all three parts of the Paul model—Standards, Elements, and Traits—are applicable to understanding the team’s thinking.

Another meaty paragraph, found in chapter 7, “The Accident’s Organizational Causes,” holistically evaluates the NASA leadership culture, and provides another condensed opportunity for applying the same methodology.

Conditioned by Success: Even after it was clear from the launch videos that foam had struck the Orbiter in a manner never before seen, Space Shuttle Program managers were not unduly alarmed. They could not imagine why anyone would want a photo of something that could be fixed after landing. More importantly, learned attitudes about foam strikes diminished management’s wariness of their danger. The Shuttle Program turned “the experience of failure into the memory of success.” Managers also failed to develop simple contingency plans for a re-entry emergency. They were convinced, without study, that nothing could be done about such an emergency. The intellectual curiosity and skepticism that a solid safety culture requires was almost entirely absent. Shuttle managers did not embrace safety-conscious attitudes. Instead, their attitudes were shaped and reinforced by an organization that, in this instance, was incapable of stepping back and gauging its biases. Bureaucracy and process trumped thoroughness and reason.¹⁸

-Managers failed to follow the data through to the full range of implications (breadth).

*- Though the foam was discussed repeated in team meetings, no decision-maker demanded, “Can you prove that Columbia has not been harmed?” This was **the question at hand**, but was not asked.¹⁹*

- A fact, “foam hasn’t hurt us badly yet,” became an tragically inaccurate conclusion, “foam is harmless.”

- An ungrounded assumption, “the crew can’t be rescued,” was confused as an inference which then justified inaction.

- Intellectual curiosity cited as an indispensable attribute of a solid safety culture.

- The organization was not metacognitive; it was not thinking about its thinking.

-In sum, the organization wasn’t thinking critically.

Student Observations

The notes above are our observations, reflecting both our familiarity with the Paul model and a lengthy history in Aerospace program management and experimental flight test. Could undergraduates see the same things we do in this report? Could we coach undergraduate in how experienced leaders might monitor the thinking of their organizations?

A Technical Leadership seminar introduced the Paul model to upper-class undergraduate engineers as discussed earlier. These students had no previous exposure, and two 75-minute periods were devoted to introducing the model, weeks prior to opening the CAIB. Latter in the semester, teams within the section were assigned to lead discussions of successive chapters through the CAIB report over a series of class meetings.

Chapter 7 of the CAIB report is perhaps the most difficult to grasp by an uninitiated student, because it attempts to examine the organizational culture and the impact of this culture on the accident. Because of this complexity, we have chosen to present the students’ critical thinking process (with questions), and the resulting outcomes in their analysis of this chapter. From a technical leadership teaching perspective, we hoped that students might grasp the underpinning organizational causal factors, and the influence organizations’ cultures have on performance.

In order to assist the students in their analysis of this complex subject, we offered the Paul model as the candidate guide. From our point, we were to answer the question; “Does the Paul model of critical thinking provide a beneficial vocabulary and construct for evaluation of complex technical case studies”. The following is the students’ assessment guide for one of the most complex chapters in the CAIB report. These are the questions *they* composed to lead their peers in discussing this chapter.

Purpose

What purpose does chapter 7 serve in understanding the Columbia accident?

- Organization structure and climate
- Safety lapse
- Lessons learned?

Questions

What is the key question the Board tries to answer in chapter 7?

What caused the deterioration of a safety culture within NASA?

What models of safety can NASA learn from?

Point of View

Whose perspective is chapter 7 written with respect to?

- CAIB point of view
- Safety within SUBSAFE, Naval Reactors, Aerospace Corp
- Safety within NASA

Assumptions

What did the CAIB assume in chap 7 to explain the organizational flaws?

- Past success justified future success
- Reduced budget and increased launch schedule placed high pressures on NASA
- NASA assumed adequate safety measures were in place.

Implications and Consequences

What were the consequences from the implications in chapter 7?

- Faster, better, cheaper ideology
- Complacent with safety procedures
- No open communications
- Must be proven right, rather than wrong
- Over confidence in ability
- Public opinion and Congressional funding

Information and Data

What information is provided to derive the implications?

- Previous NASA programs and reviews
- SUBSAFE, Naval Reactors, Aerospace Corp
- Safety Structure
- Critical Item List

Inferences and Conclusions

What can be inferred and concluded from the data in chap 7?

- Highly complex systems require effective communications
- Need for Independent Technical Engineering Authority
- Integration of data for engineers, managers, safety personnel
- Safety and Mission Assurance need to be independently funded
- Encouraging minority opinion and questioning

Concepts

What concepts does the Board focus on?

- Commitment to safety culture
- Importance of communication
- Emphasis on lessons learned
- Lack of integration of data and communication

The desire was to have the student appreciate the influence of the complex cultural factors, recognize them, and to suggest how to effect change, in view of what they learned. The conclusion of faculty and students alike was that the elements of thought provided the tools by which effective intellectual analysis could be directed. As can be seen, the model stimulated questions which focused analysis of this subject. It was striking to see that the student conclusions and concepts were in sync with the recommendations, findings, and conclusions of the CAIB. The students chose to use this methodology for their end of semester oral presentation to the faculty and classmates.

Conclusion

The goal here hasn't been preparing students for decisions identical to those faced by the Space Shuttle Program. The first goal is instead fostering a recognition that organizations must not only think, but that they must also think about their thinking. A learning organization is necessarily meta-cognitive, both thinking and thinking about its thinking. This is true both for the team and the team-member.

But in order to think about their thinking, they must also recognize the key questions they're to ask themselves. Richard Paul's model suggests broad classes (genera) of questions that critically thinking teams and team-members will ask themselves. Lastly, the teams will recognize and foster members' growth in intellectual virtues: demanding integrity, honoring humility, cultivating fairness, praising empathy.

As we surveyed the CAIB report for ourselves, and the passage above is but a sample, we found that their broadest board findings all fit within the model's bounds, once we'd added "Intellectual Curiosity" to the list of traits. Some findings did not fit due to the specialized discussion, such as those pertaining to centralized vs. decentralized organizations, or particulars with respect to safety management. These are surely beyond the goals of such a model.

More importantly, the model provided students with a ready point of entry into a complicated story with numerous interwoven sub-plots. It permitted them to recognize the

necessity of not only thinking, but thinking about thinking (metacognition). It permitted ready identification of broad classes of common organizational errors and the challenges facing leaders, without being mired in the details of NASA's particular errors. This latter is what we hope they might portage. We do expect greater facility with the model as support for the Paul model seeps into a broader range of our institution's academic programs (two core freshmen courses representing two humanities departments have now embraced the model). In 2011, we will be better able to report the insights gained by students who've spent four years handling it in various contexts.

Acknowledgement

Extracts of Paul, Niewoehner and Elder's *Engineering Reasoning* are used with permission from the Foundation for Critical Thinking.

¹ Gehman, H.W. et. al., *Columbia Accident Investigation Board, Report Volume 1*[CAIB], August 2003.

² Paul, R.W. and Elder, L., *Critical Thinking: Tools for Taking Charge of Your Professional and Personal Life* [CT1], Prentice-Hall, Upper Saddle, NJ, 2002.

³ Paul, R.W. and Elder, L., *Critical Thinking: How To Prepare Students For A Rapidly Changing World* [CT2], Foundation for Critical Thinking, Sonoma, CA, 1995.

⁴ www.criticalthinking.org [accessed 15 Dec 2007]

⁵ Paul, R.W. and Elder, L., *A Miniature Guide to Scientific Thinking*, Foundation for Critical Thinking, Sonoma, CA, 2003.

⁶ Paul, R.W., Niewoehner, R.J., and Elder, L., *A Miniature Guide to Engineering Reasoning*, Foundation for Critical Thinking, Sonoma, CA, 2006.

⁷ *Engineering Reasoning* above.

⁸ Moore, David T. *Critical Thinking and Intelligence Analysis*, Joint Military Intelligence College, Occasional Paper 14, May 2006, pg. 2.

⁹ Adapted from *Engineering Reasoning*, pg 24. Used with permission of the Foundation for Critical Thinking.

¹⁰ *Engineering Reasoning*, pgs. 6-8.

¹¹ *Engineering Reasoning*, pgs 12-13.

¹² CAIB, pg. 21.

¹³ Tufte, Edward R. *Visual Explanations*, (Graphics Press: Cheshire CT, 1997), pg. 45ff.

¹⁴ CAIB, pg. 195ff.

¹⁵ CAIB, pgs 166-170.

¹⁶ CAIB, pg. 153.

¹⁷ CAIB, pg. 147. quote of Mission Manager Team Chairperson, "I don't think there's much we can do..." Another manager was quoted on pg. 160: "if there was severe damage to the tiles, '*nothing could be done.*'" [italics in original] The CAIB addressed this question in the subsequent section, finding that rescue by *Atlantis* would have been feasible, had the question been asked rather than the assumption accepted as fact.

¹⁸ CAIB, pg. 181.

¹⁹The CAIB expresses this as a lamentation: "Organizations with strong safety cultures generally acknowledge that a leader's best response to unanimous consent is to play devil's advocate and encourage an exhaustive debate. Mission Management Team leaders failed to seek out such minority opinions. Imagine the difference if any Shuttle manager had simply asked, 'Prove to me that *Columbia* has *not* been harmed.'" pg. 192.