The Columbia Incident and NASA's Standard for Modeling and Simulation

Team 1: Hisham Bedri, Erin Leidy, Daniela Miao, Jason Ryan, Yi Xue

History of the Incident

On February 1, 2003, for the second time in the history of the space shuttle program, a shuttle was destroyed and its seven crewmembers tragically lost. Sixteen days earlier, *Columbia* had been launched for a science mission. Just 81.9 seconds after launch, a 1.7 pound piece of insulating foam detached from the external fuel tank and struck the shuttle's left wing. This impact punctured the leading edge of the wing, eventually leading the orbiter to disintegrate in the extreme heat of re-entry.[1]

NASA knew about the problem while the shuttle was on-orbit and had to decide how to ensure the safety of the astronauts and the success of the mission. A Debris Assessment Team was set up to determine the extent of the damage, if any, to the shuttle and the danger it posed to the crew. The team initially requested photographic images of the orbiting shuttle from the Department of Defense, but all requests were either denied or lost in administrative channels. The team then turned to a set of modeling tools to determine the extent of the damage. The first, named Crater, modeled debris impacts but was not designed for the type of impact in question. It was only validated for small debris, on the order of 400 times smaller than the debris that was thought to have impacted the shuttle. Two other tools predicted where and at what angles impacts may have occurred, with both models having been developed and validated for different scenarios than that of insulating foam impacting the shuttle.

Previous experience with the Crater tool suggested that it would overestimate damage, and results showing significant damage to the leading edge of the wing were discounted as overly conservative.[2] Impact angle modeling suggested that no impacts occurred at dangerous angles, while the analysis of impact locations suggested it was highly unlikely that any impacts occurred at the leading edge of the wing. Focus was shifted to assessing other areas of the shuttle.[3] Ultimately, the engineering teams determined that there was no risk of flight for the shuttle, but noted that there was significant uncertainty in this result.[4] However, this uncertainty was not passed on to senior managers, who also did not ask detailed questions concerning the technical analyses.[5] In the end, the shuttle was deemed safe for flight and allowed to proceed with re-entry, with disastrous results.[6]

Immediately after the accident, NASA set up the *Columbia* Accident Investigation Board (CAIB) to investigate its causes. Almost six months later, they produced the CAIB report, which presents the causes of the failure and recommendations about what should be done for the continuation of the space shuttle program. The immediate recommendations were mostly related to the direct physical cause of the

---

[1] Columbia Accident Investigation Board, "Report Volume 1," (Washington, D. C.: NASA, 2003), 49.
[2] Columbia Accident Investigation Board, "Report Volume 2," (Washington, D. C.: NASA, 2003), 143.
[2] Columbia Accident Investigation Board, "Report Volume 2," (Washington, D. C.: NASA, 2003), 143.
[3] Ibid., 145.
[4] Ibid., 160.
[5] Ibid., 161.
[6] Ibid., 140-72.

accident. Other recommendations were meant to improve NASA's culture; CAIB determined early on that "the accident was probably not an anomalous, random event, but rather likely rooted to some degree in NASA's history and the human space flight program's culture."[7] NASA failed on several counts to properly address safety, recognize uncertainty, communicate along the right channels, and effectively utilize the technology available to them. A major element was a failure to properly use models during the analysis process. A second report (the "Diaz report") identified action items from the CAIB report that were to be applied to NASA.[8] A major recommendation from this was the development of standards for models and simulations that would "provide uniform engineering and technical requirements for processes, procedures, practices, and methods that have been endorsed as standard for models and simulations developed in NASA programs and projects."[9]

The NASA Standard: Credibility Assessment and Uncertainty Structure Matrix

NASA developed a formal standard for models and simulations (M&S) following the recommendations by CAIB, the generalizations by the Diaz Commission, and the Return to Flight Task Group Final Report. The goal of the standard is to improve the conveyance of uncertainty and credibility of a model to decision-makers (and others) unfamiliar with the technical details of the process. In this section, we analyze how the standard performs with regards to the credibility, salience and legitimacy (CSL) framework discussed in Cash et al.[10] in terms of the standard's two major components: the Credibility Assessment Scale and the Uncertainty Structure Matrix.

The Credibility Assessment Scale focuses on "ensuring that the credibility of the results from models and simulations (M&S) is properly conveyed to those making critical decisions."[11] The fundamental premise is based on the concept that, the more rigorous the key M&S processes are, the greater the credibility of the M&S results. The general architecture is shown in Figure 1, with distinct aspects of M&S rigor divided into seven separate categories, and the degree of rigor ordered into four different levels. Similarly, the uncertainty structure matrix (Figure 2) was created to meet the requirements for performing uncertainty quantification (UQ), and reporting to decision makers the extent and results of UQ activities. The uncertainty structure matrix includes six columns for the canonical elements of UQ, with five rows describing increasingly more rigorous levels of actions to quantify and manage uncertainty.[12] The NASA Standard provides criteria for each level of rigor in the UQ and credibility assessment methods.

Credibility

Credibility involves the scientific adequacy of the technical evidence and depends on the specific disciplinary practices that are employed. Here, Blattnig et al.[13] caters directly to credibility by constructing the Credibility Assessment Scale. However, they recognize that developing a single scale to evaluate all models and simulations is infeasible since there are many different types of M&S and

---

[7] Columbia Accident Investigation Board, "Report Volume 1," 9.
[8] Al Diaz et al., "A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report," (NASA, 2004).
[9] NASA, "Standard for Models and Simulations," (Washington D. C.: NASA, 2008), 3.
[10] D. Cash et al., "Knowledge Systems for Sustainable Development " *Proceedings of the National Academies of Sciences (PNAS)* 100, no. 14 (2003).
[11] Steve R. Blattnig et al., "Towards a Credibility Assessment of Models and Simulations," in *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (Schaumburg, IL2008).
[12] Lawrence L. Green et al., "An Uncertainty Structure Matrix for Models and Simulations," in *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* (Schaumburg, IL2008).
[13]Blattnig et al., "Towards a Credibility Assessment of Models and Simulations."

organizational cultures. As a result, the goal of this scale is specifically restricted to "evaluation of credibility of M&S results for high consequence decisions within a hierarchical organizational structure."[14]

On the other hand, the credibility of M&S results is also closely related to uncertainty quantification (UQ) and uncertainty management (UM) practices. In decision-making processes based upon technical performance metrics, uncertainties in M&S introduce risk and decrease the credibility of an assessment. By undergoing the process to objectively determine the extent of the potential uncertainties that might be considered for decisions improves the credibility of the assessment process.

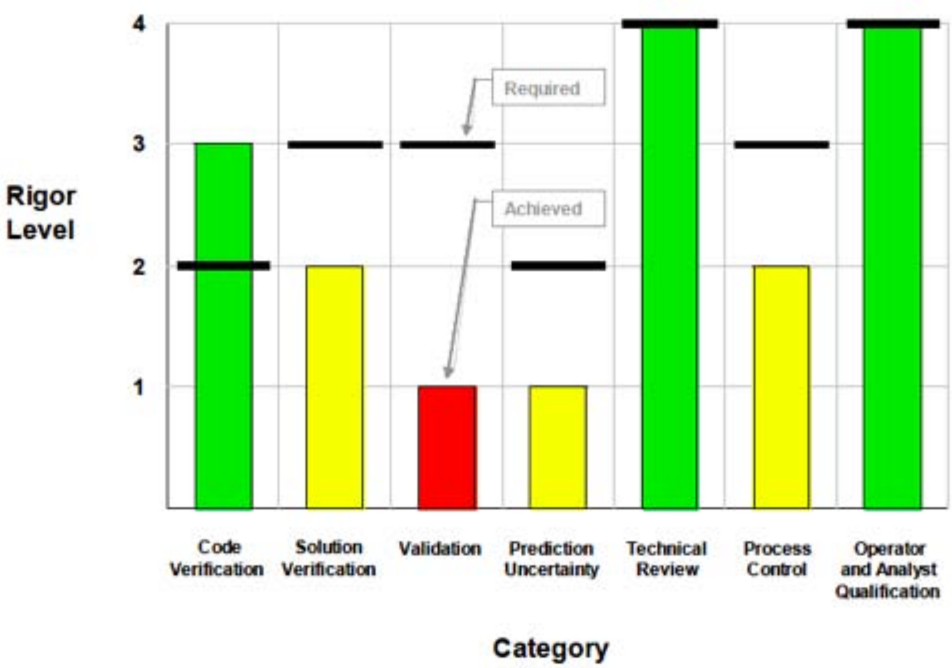

Figure 1. Sample Reporting based on the Credibility Assessment Scale.

	Canonical Elements					
	Element 1	**Element 2**	**Element 3**	**Element 4**	**Element 5**	**Element 6**
Level 1				M&S current state		
Level 2	M&S current state				M&S current state	
Level 3		M&S current state		M&S future state	M&S future state	
Level 4	M&S future state	M&S future state	M&S current and future state			
Level 5						M&S current and future state

Figure 2. Generic Uncertainty Structure Matrix for an M&S.

[14] Ibid.

Salience

Salience, as defined by Cash et al., is dependent upon the utility of specific M&S results to a decision maker in a given situation. The Credibility Assessment Scale and Uncertainty Structure Matrix both achieve salience by providing a common language for the M&S practitioners to communicate the most important and objective contributors to decision makers. Specifically, through standard reporting formats with fixed categorization, decision makers are able to quickly evaluate the results and ask key questions with respects to specific areas.

Legitimacy

By separating rigor assessment into different categories and evaluating each component independently, a wide spectrum of stakeholders is taken into consideration. Fairness and legitimacy are established through identifying individual points of failure in the assessment process and focusing on the specific areas where M&S improvements are needed. For instance, an M&S application may score high on a comprehensive credibility scale but particularly low in a specific category (such as validation in Figure 1). Under such circumstances, it is crucial that merits are given accordingly and weaknesses in specific areas are addressed promptly.

Nevertheless, the caveat here is that through fine-grained categorization, salience is compromised on an abstract level, especially for a decision maker without subject matter expertise in M&S. Without the necessary background knowledge in M&S, a decision maker may have trouble understanding the relative weight of these different rigor categories and struggle with synthesizing a well-informed decision based on the report.

Application of the NASA Standard

The NASA standard itself does not provide much guidance on how to perform a credibility assessment or uncertainty quantification. A search of the literature reveals five examples where some element of the standard was applied to an M&S process. Unfortunately, three of the five papers do not report details on the execution of these assessments, only reporting results.[15,16,17] A fourth discusses the applicability of the standard to its current modeling processes, but again does not report the details of the process or its results.[18] Only one paper was located that provided details on how the assessment was performed.[19] Here, the creation of a credibility assessment matrix was generated through the use of the Delphi method (an interview methodology for obtaining inputs from subject matter experts). In this method, subject matter experts are asked for their assessment of the model; the scores are tabulated and comments anonymized before being sent back to reviewers. The process continues until scores converge. As demonstrated in the

---

[15] Danny Thomas et al., "The Unique Aspects of Simulation Verification and Validation," in *IEEE Aerospace Conference* (Big Sky, MT2009).

[16] Aaron L. Morris and Leah M. Olson, "Verification and Validation Plan for Flight Performance Requirements on the CEV Parachute Assembly System," in *21th AIAA Aerodynamic Decelerator Systems Technology Conference and Seminar,* (Dublin, Ireland2011).

[17] Bradford Cowles, Dan Backman, and Rollie Dutton, "Verification and validation of ICME methods and models for aerospace applications," *Integrating Materials and Manufacturing Integration* 1, no. 2 (2012).

[18] Arturo Avila, "JPL Thermal Design Modeling Philosophy and NASA-STD-7009 Standard for Models and Simulations – A Case Study," in *41st International Conference on Environmental Systems* (Portland, OR2011).

[19] Jaemyung Ahn and Olivier de Weck, "Pilot Study: Assessment of SpaceNet 1.3 with NASA-STD-(I)-7009," in *Pilot Study: Assessment of SpaceNet 1.3 with NASA-STD-(I)-7009* (Cambridge, MA: Massachusetts Institute of Technology, 2007).

paper by Ahn and de Weck, this suggests that the creation of the credibility matrix is a highly subjective process, and it is assumed the uncertainty matrix follows similarly.

Conclusions

> *Nobody is smart enough to avoid all problems. That sliver of fear, the knowledge that the universe is out there waiting for the least lapse in attention to bite, is motivation that just might help you avoid catastrophe. Or perhaps not.... The first principle of a successful high reliability organization is to be "preoccupied with failure."* -Wayne Hale, NASA engineer and deputy chief of the Flight Director Office for Shuttle Operations during Columbia flight [20]

In response to the *Columbia* accident (STS-107), NASA conducted an investigation[21] and implemented a new standard for modeling and simulation. The latter action was primarily in response to the improper use of various modeling and simulation tools in determining the safety of the Shuttle during re-entry. However, the results from these analyses were not the sole cause of NASA's incorrect decisions.[22] While there were significant errors in the application of models outside of their validated areas, concerns about their use were not properly conveyed to individuals in decision-making capacities.

Science and technology do not exist in a vacuum; policy will always be influenced by more than the data on-hand. In the case of NASA, the organizational structure and the use of un-credible models led to incorrect decisions. Extrapolating from this experience, we can say that the culture of an organization heavily affects how technical information is utilized. In order to make the most of technical information (no matter how limited it is), the correct organizational culture must be in place. The credibility assessment and uncertainty quantification matrix do provide good examples of how researchers can qualify the credibility and legitimacy of their models to decision makers, but this means little when decision makers do not understand the significant of the results or do not see the results in the first place. The standard would perhaps have codified the concerns of the engineering team in the inappropriate use of the models, but would not have forced decision-makers to actually take note of and investigate the repercussions of these issues. As such, it is the opinion of this team that the standard will not prevent another accident from happening on its own.

The standard may, however, be more effective in other domains, as there are no elements of the standard that limit its application to only the aerospace domain. Properly conveying the credibility, validity, and uncertainty inherent in a model or simulation is a challenge in a variety of domains. The elements of the NASA standard provide a clear and concise method of defining these properties of a model. While the standard provides little guidance on how this should be done, establishing goals for the analyses is an important first step and preserves the standard's applicability across domains. The standard may not have addressed all elements of the *Columbia* accident, but in environments with organizational structures the encourage the questioning of data and the examination of technical details, these standard and its components could become important tools for generating dialogue about the applicability of the models and their results.

---

[20] Wayne Hale, "After Ten Years: Working on the Wrong Problem," http://waynehale.wordpress.com/2013/01/13/after-ten-years-working-on-the-wrong-problem/.

[21] Columbia Accident Investigation Board, "Report Volume 1."; Board, "Report Volume 2."

[22] ———, "Report Volume 2," 151.

Readings for Class

1. Marcia S. Smith, "NASA's Space Shuttle Columbia: Synopsis of the Report of the Columbia Accident Investigation Board," Congressional Research Service, 2003.
2. Blog posts by Wayne Hale: http://waynehale.wordpress.com/2013/01/13/after-ten-years-working-on-the-wrong-problem/; all posts beginning with "After Ten Years"
3. Martin J. Steele, "The NASA Standard for Models and Simulation,"
4. Lawrence L. Green, Steve R. Blattnig, et al., "An Uncertainty Structure Matrix for Models and Simulations," AIAA-2008-2156, 2008.

Discussion Questions

1. It is difficult to find a standard that addresses models and assessments across a broad range of domains; do you believe that there is a standard that can work? (Or, what do you think are the problems with the NASA standard? Are there problems with it?)
2. If the correct path to address concerns about the credibility or uncertainty of a model is outside of the organizational structure, what avenues do you believe you can take? What avenues should exist?
3. It has been argued by many that this accident was inevitable, due to the sheer complexity of the system. This is a not uncommon sentiment.[23] Do you feel that this accident could have been averted given the use of the NASA standard?

[23] Charles Perrow, *Normal Accidents: Living with High-Risk Technologies* (Princeton, N.J: Princeton University Press, 1984).

References

Ahn, Jaemyung, and Olivier de Weck. "Pilot Study: Assessment of SpaceNet 1.3 with NASA-STD-(I)-7009." In *Pilot Study: Assessment of SpaceNet 1.3 with NASA-STD-(I)-7009*. Cambridge, MA: Massachusetts Institute of Technology, 2007.

Avila, Arturo. "JPL Thermal Design Modeling Philosophy and NASA-STD-7009 Standard for Models and Simulations – A Case Study." In *41st International Conference on Environmental Systems*. Portland, OR, 2011.

Blattnig, Steve R., Lawrence L. Green, James M. Luckring, Joseph H. Morrison, Ram K. Tripathi, and Thomas A. Zang. "Towards a Credibility Assessment of Models and Simulations." In *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. Schaumburg, IL, 2008.

Board, Columbia Accident Investigation. "Report Volume 2." Washington, D. C.: NASA, 2003.

Cash, D., W. Clark, F. Alcock, N. Dickson, N. Eckley, D. Guston, J. Jäger, and R.Mitchell. "Knowledge Systems for Sustainable Development " *Proceedings of the National Academies of Sciences (PNAS)* 100, no. 14 (2003): 8086-91.

Columbia Accident Investigation Board. "Report Volume 1." 49. Washington, D. C.: NASA, 2003.

Cowles, Bradford, Dan Backman, and Rollie Dutton. "Verification and validation of ICME methods and models for aerospace applications." *Integrating Materials and Manufacturing Integration* 1, no. 2 (2012).

Diaz, Al, Scott Hubbard, Vicki Novak, Ghassem Asrar, Julian Earls, and Jim Kennedy. "A Renewed Commitment to Excellence: An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report." NASA, 2004.

Green, Lawrence L., Steve R. Blattnig, Michael J. Hemsch, James M. Luckring, and Ram K. Tripathi. "An Uncertainty Structure Matrix for Models and Simulations." In *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*. Schaumburg, IL, 2008.

Hale, Wayne. "After Ten Years: Working on the Wrong Problem." http://waynehale.wordpress.com/2013/01/13/after-ten-years-working-on-the-wrong-problem/.

Morris, Aaron L., and Leah M. Olson. "Verification and Validation Plan for Flight Performance Requirements on the CEV Parachute Assembly System." In *21th AIAA Aerodynamic Decelerator Systems Technology Conference and Seminar,*. Dublin, Ireland, 2011.

NASA. "Standard for Models and Simulations." Washington D. C.: NASA, 2008.

Perrow, Charles. *Normal Accidents: Living with High-Risk Technologies*. Princeton, N.J: Princeton University Press, 1984.

Thomas, Danny, Alexia Joiner, Wei Lin, Michael Lowry, and Tom Pressburger. "The Unique Aspects of Simulation Verification and Validation." In *IEEE Aerospace Conference*. Big Sky, MT, 2009.