Chasing the 6-sigma: Drawing lessons from the cockpit culture

Edward J. Hickey, MD,^a Fredrik Halvorsen, MD,^a Peter C. Laussen, MD,^b Guy Hirst, FRAES,^c Steven Schwartz, MD,^b and Glen S. Van Arsdell, MD^a

From the ^aDivision of Cardiovascular Surgery, Department of Surgery, and ^bDepartment of Critical Care Medicine, University of Toronto, The Hospital for Sick Children, Toronto, Ontario, Canada; and ^cRetired British Airways Training Standardization Captain (Boeing 747-400), Crew Resource Management Instructor and Examiner, and Senior Instructor and Examiner on behalf of the Civil Aviation Authority, Guildford, England, United Kingdom.

Received for publication May 9, 2017; revisions received Sept 8, 2017; accepted for publication Sept 19, 2017. Address for reprints: Edward J. Hickey, MD, The Hospital for Sick Children, 555 University Ave, Toronto, Ontario, Canada M5G 1X8 (E-mail: edward.hickey@sickkids.ca).

J Thorac Cardiovasc Surg 2018;155:690-6

0022-5223/\$36.00

Copyright © 2017 Published by Elsevier Inc. on behalf of The American Association for Thoracic Surgery

<https://doi.org/10.1016/j.jtcvs.2017.09.097>

Supplemental material is available online.

In August 2013, a captain and first officer on a British long-haul commercial airliner reported that they had both unintentionally—and simultaneously—fallen asleep midflight.¹ The systems approach of the airline industry to human error encourages such reporting, and the pilots involved did so freely and without repercussion. The commercial airline industry currently functions beyond the 6- σ * level—approximately 2.6 incidents per million takeoffs and landings.² Other high-stakes industries—nuclear power, military aircraft carriers, and air traffic control, for example—have also acquired 6- σ safety,³ through a philosophy toward human error akin to that of commercial aviation: (1) all have developed a preoccupation with failure and therefore engrain a culture of systemic vigilance; (2) all have endorsed and promoted mechanisms for blame-free reporting; and (3) all accept that human error is both ubiquitous and inevitable. These industries, therefore, have all embraced a systems approach to error by focusing on preventing, predicting, recognizing, and rescuing the errors that they anticipate will occur.

HUMAN ERROR IN MEDICINE: THE “PERSONAL APPROACH”

In contrast, the medical profession has been obstinate in its approach to human error. Historically, there has been a reluctance to acknowledge the occurrence of errors, or their

* For confusing reasons that relate to long-term process iteration models, 6- σ actually correlates statistically to 4.5 standard deviations and hence 3.4 events per million, or an event rate of 0.00034%. Commercial aviation exceeds this quality metric. The top pediatric heart surgery centers currently function at ~3.5- σ in terms of patient mortality (~3%).

“Fightplan” review of patient perioperative journey: team “performance rounds.”

Central Message

High-stakes medical personnel would likely benefit from training in error recognition and containment akin to the “crew resource management” that is mandatory for all commercial pilots.

See Editorial Commentaries pages 697 and 699.

See Editorial page 688.

impact.⁴ When errors are exposed, there is frequently a general resistance to transparency regarding the details and circumstances. The reasons for this stem from the fact that the medical profession has generally adopted a personal approach to human error.³ Accordingly, error is considered a shortcoming of a person or small group of individuals with whom responsibility is therefore deemed to rest. Consequently, blame is implied, if not stated. This personal approach to human error is satisfying in many respects; failures are “contained” and accounted for. It provides easy and direct causation for colleagues, patients, and their families. The personal approach also makes for sensationalist journalism. (The press seems comfortable with the phrase and concept of “pilot error” as a frequent factor in air accidents. Simple online searches for surgical error lead to national newspaper headlines describing “scandals” of “bungling surgeons,” “botched operations,” and “baby killers.”)

A fundamental flaw of the personal approach is that it ignores causal factors beyond the individuals; therefore, there is a high likelihood of error recurrence. Surgeon-specific data reporting, such as the UK Cardiac Surgery Database,⁵ although well-intentioned, endorses

TABLE 1. Classification and prevalence of threat and error subtypes observed during simulator studies and direct observation of >3500 commercial airline flight segments⁸

Threats		Errors	
Aviation	Medical correlate	Aviation	Medical correlate
Terrain: 58%	Morphology	Violation of SOP: 54%	Nonadherence to guidelines, SOP
Weather: 28%	Comorbidity	Procedural: 28%	Procedural
Aircraft malfunction: 15%	Equipment	Communication: 7%	Communication
External errors: 8%; air traffic control, ground crew	External factors, ward, administration, etc	Proficiency: 6%	Proficiency, knowledge, or skill
Operation pressures: 8%; fatigue, crew stresses	Operational stressors, fatigue, scheduling, etc	Decision error: 7%	Decision or judgment

SOP, Standard operating procedure.

the personal approach to error management by implying individual accountability and disregarding nonsurgical and institutional factors that affect patient outcome.

In the cockpit, error is a human action or inaction that leads to a reduced safety margin.⁶ Errors are common, ubiquitous, and can be considered inevitable.⁷ Fly-on-the-wall assessments (line operating safety audits [LOSAs]) of >3500 commercial airline flights by trained observers conclude that 80% contain error.⁸ Errors tend to fall into 1 of 5 category types⁸ (Table 1), the most common of which is noncompliance with standard operating procedures (SOPs).⁸ Noncompliance may reflect a cavalier work ethic, contempt for controlling regulations, or misperceptions of personal invulnerability (which, like surgeons, pilots have been shown to exhibit⁹). However, it should be recognized that overenthusiastic introduction of protocols will itself breed noncompliance and disdain for the philosophy of systemic error control. In certain situations, humans may have reasonable judgment regarding when an error can be ignored; however, investigations into intentional noncompliance (by definition “ignored errors”) in the aviation industry raise serious doubts about this general assumption. More than 40% of approach and landing accidents involve intentional noncompliance with an SOP.⁸ Perhaps more importantly, pilots who commit intentional noncompliance errors are 25% more prone to other types of error than pilots who adhere to SOPs.⁸ Therefore, noncompliance represents a general propensity to err.

Procedural errors reflect a true “mistake” in the execution of a certain task (often termed “lapses”), for example, touching the wrong key when entering coordinates or reading the wrong line of data from a chart. As with technical skills in surgery, these procedural errors account for only a minority of factors affecting performance.¹⁰ “Proficiency” errors are the least comforting, as the name implies a personal deficiency in skill or knowledge; however, these may be the most important type of error to acknowledge;

denial of failures in proficiency (a tendency in medicine) completely ignores the huge innate fallibility of humans. (The predominant criticism of individuals involved in the “Bristol heart surgery scandal” of excess deaths after arterial switch was a failure to acknowledge and manage failures [both individual and institutional] in proficiency, rather than the decrements in proficiency per se.)

ERRORS ARE OFTEN TRIGGERED BY AMBIENT THREATS

To paraphrase the Australian Civil Aviation Authority, a threat is anything that takes you away from the ideal day. Strictly speaking, threats are external influences that increase the operational complexity of the planned journey.⁶ They come at the crew, and are the risk factors for errors that may occur. Understanding and mitigating threats are central to the systems approach of threat and error management.

Threats tend to fall into distinct categories⁸ (Table 1). The most common relate to terrain or adverse weather (morphology and comorbidities in medicine). Latent threats, a particularly important type of threat from a system error management perspective, include operational conditions, culture, management structure, or aspects of training that indirectly lead to an increased risk of error.¹¹ The importance of latent threats lies in the fact that unless they are addressed, it is highly likely that errors will recur. To use James Reason’s analogy,³ active failures are like mosquitos; they can be swatted one by one, but keep coming. A better remedy is to drain the swamp (latent condition) from which they breed.

Commercial airline cockpit LOSAs indicate that ~75% of flights face 1 or more threats (range, 0-11; median, 2) and approximately 10% of these threats are mismanaged, therefore leading to an error.

THREAT AND ERROR MODEL

In the cockpit, errors—unless benign, recognized promptly and effectively mitigated—lead to an unintended

CONG

FIGURE 1. Threat and error model proposed by Helmreich.⁸ Immediate threats are factors outside the control of the cockpit crew that act to increase the complexity of the situation and therefore predispose to error occurring. Errors must necessarily be recognized for a rescue attempt to correct for the error. The rescue may completely mitigate the error (inconsequential error). On the other hand, a consequential error is one that leads to an unintended state (either an ineffective or mismanaged rescue attempt, or a completely unrecognized but important error). The unintended state itself might not be dangerous, but may serve as a threat for chains of additional error or unintended states. It is these chains that serve as the stage for amplification of the situation and potential catastrophe. Overarching organizational and cultural factors may serve as latent threats. We propose that the same model holds true in high-stakes medical specialties. Aviation threats have corollaries in medicine, as do error types. Threat and error management strategies, such as crew resource management, serve to (1) predict and manage threats, (2) minimize human error, (3) increase error recognition, (4) improve team coordination and resource utilization during rescues, (5) maximize safety margins during unintended states, and (6) recognize and break cycles of error-unintended states.

aircraft state. Importantly, an unintended state might not actually be a danger at all (eg, a perfectly safe, but different, flying configuration in an aircraft). However, a central premise of the threat-error model described by Helmreich⁸ and adopted by the Federal Aviation Authority is that an unintended state is itself an important threat that significantly increases the propensity for further errors and the occurrence of additional unintended states. Therefore, a chain has started that has a high propensity for self-propagation, leading to progressive loss of safety margins unless recognized and actively broken through crew intervention.

In commercial airline cockpits, 25% of all errors initiate such a chain; 19% lead to an unintended state, whereas 6% lead directly to a second error.⁸ Chains of unintended circumstances and errors are now recognized to set the scene for a major incident (Figure 1). It is extremely rare that an air accident investigation does not identify an upstream chain of threats leading to errors and propagation of unintended states and additional errors. One-third of all

flights contain unintended states, and 5% of landing approaches are considered frankly unstable.⁹ One-third of all unintended aircraft states are considered the end result of a chain from threat leading to error leading to an unintended state⁹ and, in the most extreme situation, loss of situational awareness. The 2009 Air France flight 447 disaster is an excellent example of this¹²; complete loss of situational awareness led a crew of 3 pilots to stall a completely functional aircraft at 38,000 feet. Icing of a pitot tube led to brief and transient [23 seconds] loss of air speed data and autopilot disengagement. The crew responded with some inappropriate manual flight control inputs, which led to an escalation of unintended states, errors, and increasingly unstable flying configurations. The crew became increasingly confused and disbelieving of the instruments. Task-sharing and coordination of roles was poor, and even at the point of impact, 2 pilots were attempting to make opposite maneuvers with the side-sticks (Figure E1).

During highly stressful emergency situations, the absolute priority is to maximize safety margins via the coordinated use of all available resources (“crew resource management” [CRM]). Problem solving is then the second priority. Sensory and cognitive senses become highly distorted in situations of extreme stress or fear, and behavior becomes quite unpredictable¹³; for these reasons, such skills as role allocation, task prioritization, and resource utilization need to be taught and rehearsed.⁹ In Air France 447, SOPs and checklists were ignored for the initial upstream problem.¹² (Checklists and SOPs are central to pilot training from a very early stage. Therefore, unintended cockpit events often trigger immediate use of a checklist or SOPs. Pilots have reported anecdotally that this automatic use of checklists during crisis situations serves a crucial role in providing clarity and focus of attention. Clarity and focus are necessary preludes to the problem-solving stage of crisis management.) Role allocation and coordination between the pilots was poor, and their initial focus was not on maintaining the aircraft within the safe flight envelope. Owing to the extreme stress, the pilots lacked the clarity of mind to use the information available to them and instead relied on “gut instinct.” Sadly, had SOPs been implemented for “loss of airspeed data at high altitude” according to the Air France emergency procedures flight manual,¹⁴ the aircraft would have remained airborne within safe margins while the crew then explored the reasons behind the crisis. Instead, the fully functional aircraft hit the Atlantic Ocean belly first, with a forward velocity of only 107 knots, only 4-1/2 minutes after the emergency began.

CREW RESOURCE MANAGEMENT

The aim of CRM is to train individuals to perform effectively in degraded situations⁹ by focusing on

nontechnical skills (“NOTECHS”^{15,16}), rather than on “rudder-and-stick” piloting skills. During training, emphasis is placed on 4 behavioral indicators: ability to cooperate, management and leadership, situational awareness, and decision making.^{9,16} Crews develop effective cross-checking and support capabilities that reduce confusion and enhance task allocation. Since its inception at a National Aeronautics and Space Administration/industry workshop in 1979,¹⁷ CRM has become a mandatory core component of commercial cockpit training. (There has been a recent expansion of CRM to include noncockpit resources [total resource management], with the realization that errors commonly occur at the boundaries of operation when cockpit crew must interact with cabin crew, engineers, and air traffic control. These occurrences often occur during cockpit communications.) Whereas the efficacy of CRM is difficult to prove, the vast expenditure of energy and resources for CRM training by commercial and military aviation is a testament to its value.¹⁸ Analyses of recent near-catastrophes have demonstrated exemplary CRM skills by cockpit crew and serve as anecdotal support of its merit. (US Airways flight 1549 successfully landed in the Hudson River after a double bird strike shortly after take-off. British Airways flight 38 lost thrust on both engines at an altitude of 720 feet during the final approach to Heathrow Airport and crash-landed 890 feet short of the runway. There was no loss of life in either accident, and both crews were praised for their CRM skills during the crises.) Comparable training has been developed for medical and surgical teams in an effort to learn to function with maximal effectiveness in highly degraded and stressful circumstances, but the impact of such training is hard to measure. However, a large Veterans Affairs study of 182,000 operations investigated the impact of CRM-type training for surgical teams and found a 50% reduction in patient mortality.¹⁹

HIGH-STAKES MEDICINE

Doctors are not pilots, and analogies between the 2 professions can be misconstrued. Nevertheless, we contend that 40 years of aviation safety research provides valuable lessons for how to improve the safety culture, especially for healthcare teams involved in high-stakes surgery and critical care medicine.

In addition to conventional “morbidity and mortality” sessions (analogous to air accident investigations), for more than a decade we have been conducting weekly performance rounds, during which we review every child who passes through our congenital heart surgery service. The process has evolved, but central to the concept is a regular protected time during which all members of the multidisciplinary heart center team can convene. (Our performance rounds involves a weekly session incorporating

surgery, cardiology, critical care, anesthesia, radiology, perfusion, and nursing leaders.) Initially, each child’s surgical journey is represented by a simple slide summarizing the success at various stages during the admission (Figure 2, A). This slide then serves as the substrate for discussions on performance gaps, near misses, or areas for improvement.

A key development was the introduction of a graphical display illustrating the smooth—or otherwise—patient journey. The in-hospital journey is depicted as a “flight” using an infographic that concisely conveys the patient’s recovery through various risk levels. These graphics clearly depict unexpected clinical deviations and escalations in risk (Figure 2, B). Mechanisms are in place to capture all clinical “events” occurring through a child’s journey, which are illustrated temporally on the graphic. Initially, these graphics were generated by hand, but the process has since become automated (Figure 2, C).

Clinical events can be categorized as threats, errors, or unintended clinical states using the very simple Federal Aeronautics Administration threat and error model; potential links between these events are usually simple to ascertain. Reviewing patients in this way provides far more insight into perioperative error and its management.²⁰ The prevalence of threats facing surgical teams (approximately 80%) is remarkably similar to that revealed by cockpit LOSAs. In our experience, one-half of all children experience important clinical errors at some point during their in-hospital journey. Errors themselves are only loosely linked to adverse patient outcomes; however, 20% of all children undergoing congenital heart surgery appear to experience chains of multiple errors and unintended clinical state, often—but not always—triggered by upstream threats. These chains, just as in the aircraft cockpit, are extremely dangerous. Patients whose journeys contain chains of error and unintended clinical states are approximately 10-fold more likely to die or to sustain a major complication, such as stroke.²¹

Errors that occur in the operating room are most often the apical errors that set the scene for subsequent chains,²⁰ and thus extreme vigilance should be exercised for their recognition. Lessons from aviation, as well as studies of NOTECHs in surgical teams, strongly suggest that operating room teams can be trained to maximize operating room prevention and rescue. Apical operating room errors are very strongly linked to failed de-escalation of risk levels through intensive care.²⁰ The intensive care environment is also prone to high error rates, either de novo errors or errors that act to amplify preexisting chains. Amplifying errors and failed de-escalations in risk are associated with 8- to 10-fold increased likelihood of death or brain injury.²⁰ These data are consistent with our previous blinded review of all 261 patient deaths over a 10-year period, in

TABLE 2. Adult congenital heart disease outcomes before and after “performance rounds” implementation in 2012

	2002 to 2012	2012 to present
ACHD surgical mortality	607 (29 deaths) = 4.7%	456 (14 deaths) = 3.0%
Adult Fontan revisions	10 (3 deaths) = 30%	11 (0 deaths) = 0%
ACHD transplant	10 (4 deaths) = ~40%	17 (2 deaths) = 11%
Adult Fontan revision/transplant	7/19 deaths = 37%	2/26 = 8%
Team structure	Dispersed	Very cohesive

“Performance rounds” flight plan review of patients was introduced—among other initiatives—to our adult congenital heart disease program in 2012. The team has become cohesive and outcomes have detectably improved. It is difficult to ascribe the improvements specifically to our performance rounds concept, but nevertheless it is one important component of the programmatic changes that we believe have played a vital role. *ACHD*, Adult congenital heart disease.

None of the physician coauthors of this article—who have a combined experience of 105 years as physicians in critical care or high-stakes surgery—has received any training focused on how to perform best in highly stressful medical emergencies. In contrast, the coauthor who is a commercial airline captain has received regular and mandatory crew resource management training as an integral part of his training and reaccreditation ever since its inception in the early 1980s. Such training is a licensing requirement for all commercial pilots.

It is extremely difficult to quantify improvements in hard outcomes that result from changes in team culture or error management initiative. Nonetheless, we have detected improvements in clinical outcome that we have attributed in part to the implementation of team performance rounds (Table 2). However, irrespective of any potential impact on patient morbidity and mortality, numerous other positive benefits have unexpectedly emerged from our performance rounds review process, including (1) a reinforced sense of individual accountability; (2) rapid resolution of problems and development of action plans; (3) greatly reinforced collective memory for future clinical guidance; (4) dramatic reduction in “corridor gossip” about complications that is highly divisive and detrimental within teams; (5) strong educational component with inclusion of imaging, photos, or literature reviews; and (6) improved cohesion, collegiality, and team-building. Our performance rounds review process has become a powerful clinical tool for patient management planning beyond discharge. The latter point is important because when performance strategies overlap with clinical work they are likely to become significantly better accepted and more effective.

CONCLUSIONS

The safety culture of commercial aviation and other high-stakes industries has evolved not just from analyzing disasters, but also largely from learning and studying error patterns in routine, live flights. Errors are recognized as ubiquitous and inevitable, often prompted by a large number of ambient threats. Errors can be effectively mitigated if recognized promptly and managed appropriately; however, if not effectively mitigated, a chain of errors and

unintended states can emerge, leading to a progressive loss of safety margins. LOSA assessments have permitted characterization of the traits of outstanding captains: (1) they are highly vigilant, (2) they continuously problem-solve, (3) they are highly communicative, (4) they use resources to delegate effectively, and (5) they delegate and empower subordinates to engage in decision management, as opposed to the historically steep power gradient between captains and subordinates. Efforts to detect error rates and patterns in high-stakes medicine have revealed that the same threat and error models are highly relevant to patient morbidity and mortality. LOSA assessment of team functioning will be a key tool for understanding threats, improving error management, and especially improving team performance. Specific training for healthcare teams to enhance nontechnical skills for maximal team performance in crises is an essential missing piece for high-stakes medicine.

Conflict of Interest Statement

Authors have nothing to disclose with regard to commercial support.

References

1. Silverman R. Airline pilots asleep in the cockpit during long-haul flight. *The Telegraph*. September 26, 2013.

2. International Civil Aviation Organization. *State of Global Aviation Safety: Evolving Toward a Risk-based Aviation Safety Strategy*. Montréal, Canada: International Civil Aviation Organization; 2013.

3. Reason J. Human error: models and management. *BMJ*. 2000;320:768-70.

4. Teasdale GM. Learning from Bristol: report of the public inquiry into children’s heart surgery at the Bristol Royal Infirmary, 1984-1995. *Br J Neurosurg*. 2002; 16:211-6.

5. Bridgewater B, Grayson AD, Jackson M, Brooks N, Grotte GJ, Keenan DJM, et al. Surgeon specific mortality in adult cardiac surgery: comparison between crude and risk stratified data. *BMJ*. 2003;327:13-7.

6. Merritt A, Klinec J. *Defensive Flying for Pilots: An Introduction to Threat and Error Management*. University of Texas Human Factors Research Project. The LOSA Collaborative; 2006. Available at: https://flightsafety.org/files/tem_dspt_12-6-06.pdf. Accessed October 8, 2017.

7. Reason J. *Human Error*. New York: Cambridge University Press; 1990.

8. Helmreich RL. Culture, threat, and error: assessing system safety. In: *Safety in Aviation: The Management Commitment*. London: Royal Aeronautical Society; 1999.

9. Helmreich RL, Wilhelm JA, Klinec JR, Merritt AC. Culture, error, and crew resource management. In: Salas E, Bowers CA, Edens E, eds. *Applying Resource Management in Organizations: A Guide for Professionals*. Mahwah, NJ: Lawrence Erlbaum; 2001.

10. Spencer FC. Observations on the teaching of operative technique. *Bull Am Coll Surg.* 1983;68:3-6.
11. Helmreich RL. On error management: lessons from aviation. *BMJ.* 2000;320:781-5.
12. Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile. *Final Report on Accident of Air France Flight 447.* Le Bourget Cedex, France: Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile; 2012.
13. Sinnett S, Spence C, Soto-Faraco S. Visual dominance and attention: the Colavita effect revisited. *Percept Psychophys.* 2007;69:673-86.
14. Federal Aviation Administration. Line Operations Safety Audits: Advisory Circular 120-90; April 27, 2006. Available at: <https://flightsafety.org/files/AC-20120-9011.pdf>. Accessed October 9, 2017.
15. Mishra A, Catchpole K, McCulloch P. The Oxford NOTECHS system: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. *Qual Saf Health Care.* 2009;18:104-8.
16. van Avermaete JAG. *NOTECHS: Non-Technical Skill Evaluation in JAR-FCL. Human Factors Open Day JAA-Project Advisory Group on Human Factors.* Hoofddorp: National Aerospace Laboratory/Nationaal Lucht- en Ruimtevaartlaboratorium; 2005.
17. Ruffell Smith HP. *A Simulator Study of the Interaction of Pilot Workload With Errors, Vigilance, and Decisions.* NASA Technical Memorandum 78482. Washington, DC: National Aeronautics and Space Administration, Scientific, and Technical Information Office; 1979.
18. Helmreich RL. How effective is cockpit resource management training? *Flight Safe Digest.* 1990;9:1-17.
19. Neily J, Mills PD, Young-Xu Y, Carney BT, West P, Berger DH, et al. Association between implementation of a medical team training program and surgical mortality. *JAMA.* 2010;304:1693-700.
20. Hickey E, Pham-Hung E, Nosikova Y, Halvorsen F, Gritti M, Schwartz S, et al. NASA model of "threat and error" in pediatric cardiac surgery: patterns of error chains. *Ann Thorac Surg.* 2017;103:1300-7.
21. Hickey EJ, Nosikova Y, Pham-Hung E, Gritti M, Schwartz S, Caldarone CA, et al. National Aeronautics and Space Administration "threat and error" model applied to pediatric cardiac surgery: error cycles precede ~85% of patient deaths. *J Thorac Cardiovasc Surg.* 2015;149:496-505; discussion 505-7.
22. de Leval MR, Carthey J, Wright DJ, Farewell VT, Reason JT. Human factors and cardiac surgery: a multicenter study. *J Thorac Cardiovasc Surg.* 2000;119(4 Pt 1):661-72.
23. Carthey J, de Leval MR, Reason JT. The human factor in cardiac surgery: errors and near misses in a high technology medical domain. *Ann Thorac Surg.* 2001;72:300-5.
24. Catchpole K, Giddings AE, Hirst G, Dale T, Peek G, de Leval MR. A method for measuring threats and errors in surgery. *Cogn Tech Work.* 2008;10:295-304.
25. Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and error in the operating room: analysis of skills and roles. *Ann Surg.* 2008;247:699-706.
26. von Thaden TL, Wiegmann DA, Shappell SA. Organizational factors in commercial aviation accidents. *Int J Aviat Psychol.* 2006;16:239-61.

FIGURE E1. Threat and error model summarizing the last 4 minutes of Air France flight 447, as described in the official accident report.¹² The flight was proceeding as planned: cruising at 35,000 feet at 540 knots under autopilot control. Poor weather and high cumulonimbus clouds (threat 1 [T_1]) contributed to the icing of a pitot tube, with loss of airspeed data and consequent autopilot disengagement: the major threat (T_2). A latent threat (LT_1) was a design fault, which predisposed to pitot tube icing. Loud, rare, and distressing stall warnings sounded in the cockpit because of the erroneous airspeed data. The junior copilot in right-hand seat took manual control and made a major proficiency error (E_{2P}) by inappropriately pulling back on the stick and climbing. This important error may have been prompted by limited simulator training sessions of low-altitude stalls, during which an appropriate response is to pull the nose up under full throttle (LT_2). The 2 copilots did not refer to or follow loss of airspeed checklists (E_{3V}); had they done so, they would not have climbed in altitude. The preceding errors all contributed to unintended loss of speed and excessively high altitude (U_1). At one point, the pilot flying (PF) became concerned about overspeed and reduced throttle (E_{4J}), likely because training sessions disproportionately focus on the risks of high-altitude overspeed, which tend not to apply to the modern Airbus A330 (LT_3). The progressive climbing and loss of airspeed (U_1) gradually led to a breach in aircraft flight ceiling (U_2). The senior nonflying copilot (PNF) recognized this, but a lack of communication and cross-checking (E_{5C}) led to his requests to descend never actually materializing. This latter error in communication was likely exacerbated by perhaps the earliest protocol violation in the flight: the captain not clearly debriefing and allocating roles before retiring for rest (E_{1V}). The pitot tube deiced and provided accurate airspeed only 27 seconds after it initially malfunctioned. At that time, all instruments were entirely functional, the aircraft was airborne, and the pilots were—briefly—in control. From that point on, therefore, the events were entirely manmade (shaded area). Because of the falling airspeed, the stall warnings restarted (and continued until the end of the flight). The junior copilot (having started to put the nose down and gain speed) immediately responded again inappropriately by pulling the nose up sharply and applying full throttle (TOGA) (E_{6P}). The resulting gradual increasing angle of attack led to gradual loss of kinetic energy despite 100% engine thrust. The usual maximum angle of attack before stall is typically $\sim 20^\circ$; however, soon the angle of attack exceeded 40° with a velocity of ~ 100 knots: the flight envelope had been breached and the aircraft was in a nose-up flat stall (U_3). Throughout these final minutes, there was complete loss of situational awareness. Air accident investigators have suggested that lack of horizons due to nighttime flight may have contributed to disorientation (T_3). Crew resource management techniques were poor (E_{7C}), and perhaps exacerbated by the number of extremely rare and stressful aural stall alarms (T_4). Behavioral science experiments have demonstrated that in highly stressful situations, humans tend to act on aural stimuli preferentially over visual stimuli (perhaps explaining the TOGA response to stall warnings rather than digesting the information on visual displays).¹³ The aircraft was plummeting at $>10,000$ feet/minute. Distressingly, when passing through 10,000 feet, the PF and PNF were fighting with the controls, with the more senior copilot trying to push the nose down and gain speed. A design flaw in the A330 means that neither pilot could feel the other's maneuvers, and they cancelled each other out (LT_4). As they passed through 2000 feet, the terrifying ground proximity warning sounded (an automated voice repeating "pull up...pull up," and the PF declared in despair "But I've been pulling up all the time!") At this moment, the PNF suddenly realized this major error had transcended the entire emergency and that the plane was in a steep nose-up stall. It was too late to regain control, however, and impact followed 4 seconds later. Cockpit voice recorder transcripts are available at <http://www.planecrashinfo.com/cvr090601.htm> (warning: distressing content). ft, Feet; kt, knots; TOGA, "take-off, go-around configuration"; GPW, ground proximity warning; PF (R), pilot flying (31-year-old copilot Bonin, in the right-hand seat); PNF (L), pilot not flying (38-year-old copilot Roberts, in the left-hand seat); A/P, autopilot; LT, latent threat; T, threat; E_V , error (protocol violation); E_C , error (communication); E_P , error (proficiency); E_J , error (judgment).