SCHOOL OF AVIATION
Lund University

Errors in our understanding of human error: the real lessons from aviation for healthcare

Sidney W. A. Dekker

Technical Report 2003-01

Lund University School of Aviation
Address: 260 70 Ljungbyhed, Sweden
Telephone: +46-435-445400
Fax: +46-435-445464
Email: research@tfhs.lu.se

Main points

Medicine can learn from aviation in dealing with human error, but it is not through counting superficial mistakes or pursuing culprits.

The real problem is not to stop bad doctors from making mistakes that lead to harm, it is preventing good ones from doing so.

Human error does not simply cause failure: there is no single cause, neither for success nor for failure.

Human error, under whatever fashionable label, should be the beginning of an investigation, not its conclusion.

To understand human error we need to capture the systematic connections between human assessments and actions and features of people's tools, tasks and operating environment.

Healthcare publications have recently carried a number of contributions to the debate about human error in medicine[1]. One of these purports to share the lessons that aviation has learned about human error and how to manage it[2], the idea being that medicine could learn vicariously through the false starts, dead-ends and promising avenues that aviation has pursued through the years. The approach suggested by Helmreich, however, is not one of the promising avenues and not representative of the aviation community as a whole. In fact, it propagates what is today known as "the old view of human error"[3], encouraging medicine to retread itself with a counterproductive initiative merely cloaked as progress on safety.

A number of myths dominate the old view of human error[3]. The first is that human error causes failure. Our organised and engineered systems are basically safe—safety threats come from the inherent unreliability of people. The old view sees human error as endogenous, stemming from motivational shortcomings or inherent processing limitations of the human organism[2]. The way to deal with human error is to count the problem and limit the human contribution to potential danger by inserting more technology, more procedures, more supervision, more selection, more training. The old view often leads to what we now call "The Bad Apple Theory"[4]: the system is basically safe, if only it were not for bad apples in it. Getting rid of the bad apples or limiting their spoil on the rest will deal with the human error problem. The model proposed[2] does this through counting errors, coarsely categorizing them, and laying out the various countermeasures necessary to get the counts on each category down.

A case study used to illustrate the proposed model[2] involved a healthy 8-year old boy choked to death on a clogged endotracheal tube while undergoing minor elective surgery. The case study, meant to illustrate how to learn from error with Helmreich's model[2], in fact becomes a re-run of The Bad Apple Theory. Unpacking human error through the model leads to the discovery of multiple failures—nothing more than saying "human error" again. It includes nurses who failed to speak up; a surgeon who continued operating when he should not have; an anaesthetist who failed to intervene. The model simply supplants one failure for the other, while masquerading as explanation. Even when putatively making forays into the organizational antecedents of failure, The Bad Apple Theory prevails. If only the organization had acted earlier on reports about the anaesthetist, this death would not have occurred. There was a bad apple, a weak component, ready to break at any moment in a safe but underinformed system. The problem of human error, however, is not a problem of bad apples at all. It is not an issue of counting their mistakes, or of throwing them out of the basket under whatever label is in fashion.

The myth of counting mistakes

The real problem, says Atul Gawande, is not how to stop bad doctors from harming their patients. It is how to prevent good ones from doing so. Gawande, himself a physician, recounts his medical mistake in the New Yorker[5]. As usual, the story (this one of an emergency tracheostomy gone sour) is one of exceeding complexity, extending deeply into the engineered, organised, social worlds that surround Gawande on his pursuit to safe his patient, dogged in the end by such banal elements as a dim overhead light—the only one available. "Human error", if there were such a thing, is not a question of individual single-point failures to notice or process—not in this story and probably not in any story of breakdowns in patient care. Care that goes sour spreads out over time and in space, touching all the areas that usually make caregivers successful. The "errors" are not surprising brain bloopers. Instead, they are series of actions and assessments that are systematically connected to people's tools and tasks and environment; actions and assessments that often make complete sense when viewed from inside the situation. Were one to trace "the cause" of failure, the causal network would fan out immediately, like cracks in a window, with only the investigator determining when to stop looking because the evidence will not do it for him or her. There is no single cause. Neither for success, nor for failure. Aviation has discovered that we never find cause. We construct it. Labeling certain assessments or actions in the swirl of human and organizational and technical activity as causal, or as "errors" and counting them in some database, is entirely arbitrary and ultimately meaningless.

The body count in medicine ("The US Institute of Medicine estimates that each year between 44000 and 98000 people die as a result of medical errors"[2]) may have mustered political momentum for the patient safety movement, but it has neither a basis in scientific practice, nor a future in productive countermeasures. Counting surface variabilities, as proposed in Helmreich's model[2] is protoscientific at best. Counting does not make science, and intervening on the basis of surface variability is bound to fail because it does not get to underlying genotypes. In fact, when it comes to error, counting is particularly difficult. Human error "in the wild"—that is, as it occurs in natural complex settings—resists tabulation because of the complex interactions, the long and twisted pathways to breakdown and the context-dependency and diversity of human intention and action.

Aviation has learned that defining human error as cause (under whatever labels are still proposed in Helmreich's model[2]: non-compliance, misbehaviour, failures to communicate etc.) is counterproductive. The idea is that 98000 deaths represent the distance we have to go before we reach full safety. Full safety lies somewhere on, or beyond, the horizon, and the 98000 deaths is what lies between us and that goal. This assumption about the location of safety is an illusion, and efforts to measure the distance to it are little more than gauging our distance from a mirage. Safety is right here, right under our feet—not across some distance. Aviation has learned how people in complex systems create safety. They make it their job to anticipate forms of, and pathways toward, failure. They devise strategies that help forestall failure by tailoring their tasks, by inserting buffers, routines, heuristics, tricks, double-checks, memory aids. Aviation has learned that the dominant human contribution to failure occurs because complex systems need an overwhelming human contribution for their safety. Aviation has learned that human error is the inevitable by-product of the pursuit of success in an imperfect, unstable, resource-constrained world. Aviation has learned (for example by automating much of the flying task in the eighties) that trying to eradicate human error can eradicate or compromise human expertise—the most profitable investment in safety any complex system can make.

Learning a new language of failure

Saying what people or organizations failed to do—as proposed in Helmreich's model—does not explain why they did what they did. In fact, it does not explain anything, it only judges them for what they did not do. Pointing out what people should have done is "Monday morning quarterbacking" and does not contribute to our understanding of breakdowns in care. In fact, using a judgemental language for failure will fuel the fears and concerns that continue to frustrate error reporting systems[6], stifling the potential to learn. Changing the language used to describe and deal with failure is a lesson that aviation has had to learn. Investigators must now try to avoid the counterfactual reasoning error (saying what people could have done to avoid the outcome) or the data availability error (pointing to those data that, in hindsight, would have revealed the true situation), because none of this explains why people did what they did on the inside of their unfolding situations. It simply substitutes the investigator's reality for the one that surrounded people at the time.

The pressure to identify culprits can be severe. But learning from failure and pursuing culprits, under whatever guise, are fundamentally at odds. Punishing is about keeping our beliefs in a system intact. Learning is about changing them. Punishing is about seeing the culprits as unique parts of the failure. Learning is about seeing the failure as a part of the system. Punishing is about stifling the flow of safety-related information. Learning is about increasing that flow. Punishing is about not getting caught the next time. Learning is about countermeasures that remove error-producing conditions so there will not be a next time. Punishing is about closure, about moving on from the terrible event. Learning is about continuity, the event firmly integrated in what the system knows about itself.

Patterns of failure

The new view of human error does not see human error as the cause of failure. It sees human error as the effect, or symptom, of trouble deeper inside the system. It uses the discovery of human error as the beginning of an investigation, not as its conclusion. Rather than counting mistakes, the new view seeks to probe events more deeply; transcribe them from a context-specific language to a concept-dependent one through various levels of analysis; and finally synthesise across sequences of events to identify patterns of failure.

What are some of these patterns? We in aviation see people getting outwitted by a dynamic world. At the start, their situations are too unpredictable for effective feedforward—it is uncertain which actions will lead to greatest safety until more evidence is gathered—but once the pace quickens, things may become too dynamic for effective feedback. In aviation we see people and organizations drift into failure, where weeks or months or years of unremarkable repetition (for example, taking off with many unresolved issues on the "Hold Item List", a list of things on the airplane that are basically broken) produce an erosion of safety margins. This normalization of deviance over time has taught us that Murphy's law is wrong: What can go wrong usually goes right—and then we draw the wrong conclusion, in which past safety is taken as guarantee for future success.

In aviation, we have seen the introduction of more technology as illusory antidote to the plague of human error. Instead of reducing human error, technology changed it, aggrevated the consequences and delayed opportunities for error detection and recovery. Technology upset traditional human-to-human coordination, undermining traditional strategies for double-checking and monitoring each other's work. In aviation, we have seen the introduction of more procedures as illusory countermeasure against non-compliance. We discovered that the application of procedures is a substantive cognitive activity, where people continually need to balance the rigid application of guidance in situations that call for different measures, with potentially erroneously adapting procedures in the face of complexity. We have discovered that procedural overspecification does not increase compliance but reduces it. In aviation, we have seen The Bad Apple Theory at work. We have seen the procedural straight-jacketing, technology-touting, culprit-extraditing, train-and-blame approach be applied, and invariably stumble and fall. What we have found is that it is a dead end. There is no progress on safety in the old view of human error.

These are the real lessons from aviation. The point in learning about human error is not to find out where people went wrong. It is to find out why their assessments and actions made sense to them at the time. It is not to say what people failed to do. It is to understand why they did what they did, by probing the systematic, lawful connections between their assessments and actions, and the tools, tasks and environment that surrounded them. The various contributions to the healthcare literature put medicine before a basic choice: (1) see human error as the cause of failure, count mistakes and launch restrictive countermeasures on the basis of surface variability, or (2) see human error as a symptom of failure deeper inside the system, and try to understand why assessments and actions made sense on the basis of people's tools, tasks and environment. The choice, really, is between not learning and learning. The choice is between regressing into the counterproductive pursuit of culprits under fashionable and protoscientific labels, or about actually making progress on patient safety.

1 *BMJ* 2000: 320.
2 Helmreich RL. On error management: lessons from aviation. *BMJ* 2000: 320: 781-785.
3 Cook RI, Woods DD, Miller C. *A tale of two stories: contrasting views of patient safety.* Chicago: National Patient Safety Foundation.
4 Dekker SWA. The disembodiment of data in the analysis of human factors accidents. *Human Factors & Aerospace Safety* 2001: 1: ...
5 Gawande A. When doctors make mistakes. *New Yorker* 1999: Feb 1: 40-55.
6 Pietro DA, Shyavitz LJ, Smith RA, Auerbach BS. Detecting and reporting medical errors: why the dilemma? *BMJ* 2000: 320: 794-799.