Outside the Lines

"Problems cannot be solved by thinking within the framework in which the problems were created." – Albert Einstein

by Ira J. Rimson and Ludwig Benner, Jr.

Ten-to-the-Minus-Ninth: Satisfaction or Satisfiction?

"[S]cience is not a matter of unattainable perfection, as the histories of Galileo, Copernicus, Pasteur, Einstein and others will attest. Science depends not on speculation but on conclusions verified through experiment and real world data. Verification is more than computer simulations — whose conclusions mirror the assumptions built into the model. Irrespective of repeated assertions regarding 'scientific analysis,' there is neither robust analysis, nor is analysis science."

— Adapted from James Schlesinger, "The Theology of Global Warming." *The Wall Street Journal*, August 8, 2005, p. A10.

Suppose you are safety manager for a new system. Your customer specifies a catastrophic mishap rate of not more than 1×10^{-9} aggregated operating hours. There will eventually be a total of 1,000 systems, operating at an average (allowing for maintenance downtime, etc.) of 12.5 hours per day, 320 days per year. If the first catastrophic mishap amongst all systems occurs at 10,000 hours' aggregated fleet operating time, what is the minimum calendar time that must elapse before the next catastrophic mishap, for the system to remain compliant with the specification?

a) 2.5 years
b) 25 years
c) 250 years
d) None of the above
e) This is an absurd question, because there's no way anyone can demonstrate failures prospectively by extrapolating from a single data point (or even two or three).

Ten reviewers came up with eight different answers, ranging from 90 days to 500 years. You can try your hand at it, but response "e" is acceptable and probably the most accurate.

Peter Ladkin's Op-Ed article, "Taking Software Seriously,"[1] in the May-June 2005 issue of *JSS*, exposed the system safety community's blind acceptance of an insidious sophistry that has been perpetuated for almost half a century. He addressed the issue of product safety specifications that cannot be demonstrated, much less attained. We compromise our efforts to develop reasonably safe systems by uncritically accepting safety objectives that are unreasonable, undemonstrable and impossible to achieve, and then offer "analyses" designed to support those predetermined untenable objectives to corporate or government decision makers.

For software (but not so limited, as we shall see), Dr. Ladkin points out that "…the very best statistical-testing regime will find those faults that lead to failure at a rate equal to or more frequent than 100,000 operating hours." Yet existing standards require that "…someone will have had to demonstrate a dangerous failure rate of at most one failure in one hundred million [or] one billion hours of operation (the fabled ten-to-the-minus-nine rate)…." If you produce only one unit, the unwary customer is led to believe that it may have a "dangerous" failure every ~114,000 years. If you produce 10,000 units, a dangerous failure would presumably be acceptable every 11.5 years; 100,000 units — annually; and so forth *reductio ad absurdum*. Let's look at a real-world example of what happens when we examine this statistical contrivance critically.

The Boeing 737, in all its variants, is the most prolific transport airplane in history. A total of 4,384

[1] *Journal of System Safety*, Vol. 41 No. 4, May-June 2005, pp. 11-12.

were built, of which 3,815 were active as of July 2005.[2] Forty-seven fatal mishaps had occurred since the 737's operational introduction in February 1968, in ~76 million flights.[3] ("Flights" is considered to be a more accurate measure of risk exposure than flight hours.) A fatal mishap can be considered to be demonstrated evidence of a catastrophic failure involving both the specific aircraft and the air transportation system within which it operates. On a statistical basis, the Boeing 737 series has been involved at a fatal catastrophic system failure rate approximating 6.2×10^{-7} flights[4]. Assuming the average B-737 flight to be two hours, the fleet's hourly fatal catastrophic system failure rate is ~1.25×10^{-8} flight hours.[5]

Section 25.1309 of the Federal Aviation Regulations[6] states: "(b) The airplane systems and associated components, considered separately *and in relation to other systems*, must be designed so that — (1) The occurrence of any failure condition which would prevent the safe flight and landing of the airplane is extremely improbable…".[7] (Emphasis added.)

That requirement is meaningless, absent definition of "Extremely Improbable," but never fear — the Feds have developed extra-regulatory administrative techniques to compensate for syntactically deficient regulations. At the FAA, it's called an Advisory Circular ("AC" for short). FAR §25.1309 has its own dedicated document, currently AC25-1309.1A, issued June 21, 1988. The purpose of ACs is to inform the public of what the FAA *really* means by its legally bloviated[8] regulations, and how to provide what the agency *really* wants. Thus, AC25-1309.1A suggests (in a process as slow and painful as dental extraction without anesthesia):

- at ¶7.b., that, "The severities of failure conditions may be evaluated according to the following considerations: (1) Effects on the airplane…; (2) Effects on the crewmembers…; and (3) Effects on the occupants…."
- at ¶7.d.(3): "Catastrophic failure conditions must be extremely improbable."
- at ¶9.e.(3): "Extremely improbable failure conditions are those so unlikely that they are not anticipated to occur during the entire operational life of all airplanes of one type."[9]
- and finally, at ¶10.b.(3): "Extremely improbable failure conditions are those having a probability on the order of 1×10^{-9} or less."

Eureka! There it is, prominently located in paragraph 10 on page 15 of a 19-page, 17-year-old non-regulatory document: a real quantitative definition of the government's criterion for achieving operational safety of transport airplanes. Yet the Boeing 737 type, the largest fleet in the world, has an *actual* operational safety record that falls short of the FAA's certification standard by more than an order of magnitude.

> *"Why have specifications been established that are impossible to attain, or 'design to' requirements that cannot be tested?"*

Does that make the 737 an unsafe airplane? No.

Have catastrophic failures recurred in the 737 series? Yes.[10] Has the FAA done something about all of the sources of its catastrophic failure rate? The FAA prides itself on issuing Airworthiness Directives when, "(a) An unsafe condition exists in a product; *and* (b) That condition is likely to exist or develop in other products of the same type design."[11] After the fact.

Is FAR §25.1309 an absurd specification? According to our mathematic approximations, the 737 fleet has accumulated about 150 million flight hours in 37 years of operation. Had it met the requirements of AC25.1309A, ¶10.b.(3), it shouldn't have had *any* "Extremely Improbable" failures yet. In fact, had it met the requirements of AC25.1309A's ¶9.e.(3), the 737 fleet shouldn't ever have *any* "Extremely Improbable" failures during the entire operational life of *all* versions of the airplane. Reasonable? Rational? Attainable? You be the judge.

---

[2] http://www.planemad.net/data/list/Boeing/737/ — not counting the crashes since July 2005.

[3] http://www.airdisaster.com, data as of July 16, 2005.

[4] For purposes of this discussion, the aircraft is assumed to have a role in all system failures. We acknowledge that fatal mishaps are not all initiated solely by a system failure specifically within the instant aircraft. Likewise, not all catastrophic airplane subsystem failures result in fatal mishaps.

[5] Don't quibble about the numbers; they have been assumed *arguendo*.

[6] Part 25 of Title 14, U.S. Code of Federal Regulations, establishes criteria that must be demonstrated for transport airplanes prior to being certified by the FAA. Section 25.1309 thereof relates to assuring the function of "equipment, systems, and installations."

[7] A unique characteristic of system safety engineering is that it considers human operators as a coequal operating system (or associated component) to hardware; thus, within the requirements of FAR §25.1309, human performance deficiencies from design assumptions must be considered both separately and in relation to other systems.

[8] For definition, see http://www.worldwidewords.com.

[9] In anticipation of semantic arguments on the meaning of the word "type," see the following definition from FAR Part 1 – Definitions and Abbreviations: "*Type:* (2) As used with respect to the certification of aircraft, means those aircraft which are similar in design. Examples include: DC-7 and DC-7C…." Thus the extended 737 family, including derivatives, qualifies as "one type."

[10] United Airlines at Colorado Springs and USAir at Pittsburgh, *e.g.*

[11] FAR §39.1.

In his 1987 book, *Managing Risk*[12], Dr. Vernon Grose exposed in print the fallacy of the quest for "…a single number that wraps up all the agony of alternatives into one focused numeral."[13] He compares actuaries, whose predictions form the bases for estimating future insurance losses, with systems designers, who try to predict how and when their systems might fail. Actuaries work from an extensive objective database from which they derive objective probabilities. On the other hand, most systems designers lack objective data derived from operational experience and face the necessity to comply with irrational requirements that specify the desired outcome *before* examining the data. In the absence of objectively verified data, designers' decisions have become reliant on subjective analyses of unverified data that are accepted uncritically, and Probabilistic Risk Assessments cobbled together from whatever data can be found that bears any resemblance to the system at hand.[14]

The late Nobel Laureate economist Herbert A. Simon coined a term to define a utilitarian criterion by which acceptable decisions are frequently made: that they "satisfice"; *i.e.*, they "…obtain an outcome that is good enough."[15] Satisficing action can be contrasted with maximizing action, which seeks the biggest, or with optimizing action, which seeks the best. An old Russian proverb observes that, "Better is the enemy of good enough." Right enough!

Why have specifications been established that are impossible to attain, or "design to" requirements that cannot be tested? That kind of sophistry leads to adopting wild guesses, guesstimates and ill-conceived PRAs, followed by specious risk-prevention initiatives. Irrational specifications beget irrational decisions, based at best on imagination instead of robust data. Rather than achieving Simon's criteria for "satisfice," they encourage "satisfiction." As Ladkin argued: "Best simply to give them up."

Acknowledgments

The authors are grateful to the following reviewers who contributed comments on the drafts: J. Arlin Cooper, John M. Covan, Andrew L. Fuller, Vernon L. Grose, Peter B. Ladkin, Michael Murphy and Ronald J. Stupak.

---

[12] Englewood Cliffs, Prentice Hall (ISBN 0-13-551110-0).

[13] Id., Chapter 23, p. 260.

[14] In the authors' experience, few Probabilistic Risk analysts bother to verify the accuracy or relevance of the data they use to predict the likelihood of future occurrences from past events.