

# Flight Test Safety Fact



Published for the Flight Test Safety Committee

## In This Issue

**Your Words Matter**- How we talk when we disagree (outside of the cockpit/control room) is part of culture we haven't mastered.  
**2D Risk Matrix** – Rod Huete presents an op-ed about subjective probability and the 2D risk matrix from FAA order 4040.26C  
**Turbo Talk** – Turbo shares insight regarding the necessity to implement lessons learned and makes a plug for the FTSW.

## Your Words Matter

**Mark Jones Jr.**

Words have the power of life and death. With a single word, you can save the lives of the aircrew or prevent an airplane crash: "Abort!" A word like that normally gets repeated — "Abort! Abort! Abort!" But it only needs to be heard once.

Your words matter. We need to hear them.

In the familiar setting of the cockpit or control room, we have certain protocols for communication. In those situations, there are specific words spoken at specific times. You already know what word to say and when to say it, but that wasn't always true. Can you remember back to your early childhood? You didn't know enough to warn your dad about the car in his blind spot when he started to drift into the other lane. When you learned to drive, though, you gained a perspective that allowed you to contribute meaningfully to the same situation. If you yelled, "Watch out!" your words mattered.

When we start to talk or write about "safety culture," we are in unfamiliar territory. The careful procedures for CRM that apply to cockpit and control room communication don't exist here. When we discuss where safety officers fit on the org chart or what the safety review process should look like, we don't have clear lines to prevent us from drifting out of our lane. When we disagree, we tend to prefer the kind of compromise that doesn't hurt feelings. Discussions about organizational drift, complacency, and feedback can create tension and discomfort. These conversations are not quite the same as the debrief either, and the norms we use in the briefing room may not apply. As it turns out, it's hard to decide if "lives are at stake" and whether or not to blurt something out right there or wait until a more appropriate forum. All of this applies even more at a macroscopic level, and I think there's evidence that we haven't figured this out as a community yet. But this issue presents a difference in opinion, so let's see how we do. With the focus on culture for this year's Flight Test Safety Workshop, maybe we should talk about how a culture that does not have conflict is unhealthy. After three years of FTSF's, this edition is your formal re-invitation to join the conversation...

**Your words matter. We need to hear them.**

## Op/Ed: The Subjective 2D Risk Matrix used in Flight Test

**Rod Huete**

There appears to be a difference of opinion as to whether the probability of occurrence in the Risk Matrix should be based on subjective judgement (grey bearded approach) or based on a known database (e.g., contain numbers).

I remember when I was Chief of Safety at the USAF AFFTC back in the early nineties, and I remember this discussion taking place. It was unanimously agreed that we could not even attempt to assign probability numbers because these numbers are very subjective in Flight Test. It's not like a piece of hardware where you can define the failure modes and assign statistical failure probabilities based on reliability numbers (used in classical systems safety). These probabilities are based on known failures rates and based on corporate history but many times miss the new failure modes that could occur on something that has never been done before, what we call unknowns or unknown unknowns. Is this an obsolete concept from the nineties? Let's see in the discussions below.

Some people may argue that they have failure data based on flight test events (e.g., a database). For example, as Mark Jones states:

It's NOT hard to count how many times Steep Approach resulted in hard landings. This is a starting point for quantitative prob. It's not hard for me to count the number of times a pilot did not guard the "good throttle," and during a throttle chop I initiated; the good throttle was bumped back. Both of these occur much more frequently than the qualitative words suggest. Our computers and data collection are so much better than they were.

The concept above is counter-argued by John Hed:

You say it's "easy to count" how many times there was a hard landing during steep approach. I beg to differ. To get an accurate number, first you have to know or find out how many steep approach projects have happened throughout history (and all companies, etc.). Then you have to find out how many steep approach tests were flown (company and cert or whatever). Then you need to find out how many hard landings there were. You see my point. This would be a HUGE undertaking and the odds of getting an accurate number are pretty low, in my honest opinion. And in the end, the "probability" side of the chart is not always set up with numbers. So it really devolves to "engineering judgement" by someone with a lot of experience based on what they have heard over the years. (Another reason to have a good non-retribution system to report incidents that somehow, we can all tap into. Without that, some companies will never admit to any incidents.)

In flight test we are constantly testing the edge of the technological envelope. Most probabilities you put on hazards occurring are educated, engineering estimates at best, and based on past history, which of course we don't have for new systems. Trying

to put numbers to those probabilities is pretty futile in my mind. And the fact is, that those are not the point. The point is the PROCESS. Analyzing and trying to find all the hazards and then trying to mitigate them to the best of our ability. When you do a risk analysis for a wind-up turn do you really think you can come up with some number like 1.35E-05 or something?

Rod’s note: We have tried in the NASA/FAA Flight Test Safety Database to get companies to report lessons learned (or incidents) to be able to have such database. Unfortunately, companies are not very willing to share what they consider proprietary information. So no matter how complete you may think your database is, you will be missing those incidents that were not reported. Even if you had such a database, you just can’t pick up numbers from a table or from somebody’s definition of those numbers and feel like you have done a good job—that is a false sense of security as far as I am concerned.

Below are examples of 2D Risk Matrices from different sources for further discussions below:

ICAO Annex 19 Manual:  
Safety Risk      Severity

Probability	Catastrophic A	Hazardous B	Major C	Minor D	Negligible E
Frequent 5	5A	5B	5C	5D	5E
Occasional 4	4A	4B	4C	4D	4E
Remote 3	3A	3B	3C	3D	3E
Improbable 2	2A	2B	2C	2D	2E
Extremely improbable 1	1A	1B	1C	1D	1E

Figure 1.

Safety Risk		Severity				
Probability		Catastrophic A	Hazardous B	Major C	Minor D	Negligible E
Frequent	5	5A	5B	5C	5D	5E
Occasional	4	4A	4B	4C	4D	4E
Remote	3	3A	3B	3C	3D	3E
Improbable	2	2A	2B	2C	2D	2E
Extremely improbable	1	1A	1B	1C	1D	1E

Figure 2.

Safety Risk Index Range	Safety Risk Description	Recommended Action
5A, 5B, 5C, 4A, 4B, 3A	INTOLERABLE	Take immediate action to mitigate the risk or stop the activity. Perform priority safety risk mitigation to ensure additional or enhanced preventative controls are in place to bring down the safety risk index to tolerable.
5D, 5E, 4C, 4D, 4E, 3B, 3C, 3D, 2A, 2B, 2C, 1A	TOLERABLE	Can be tolerated based on the safety risk mitigation. It may require management decision to accept the risk.
3E, 2D, 2E, 1B, 1C, 1D, 1E	ACCEPTABLE	Acceptable as is. No further safety risk mitigation required.

Figure 3 – FAA Order 4040.26C.

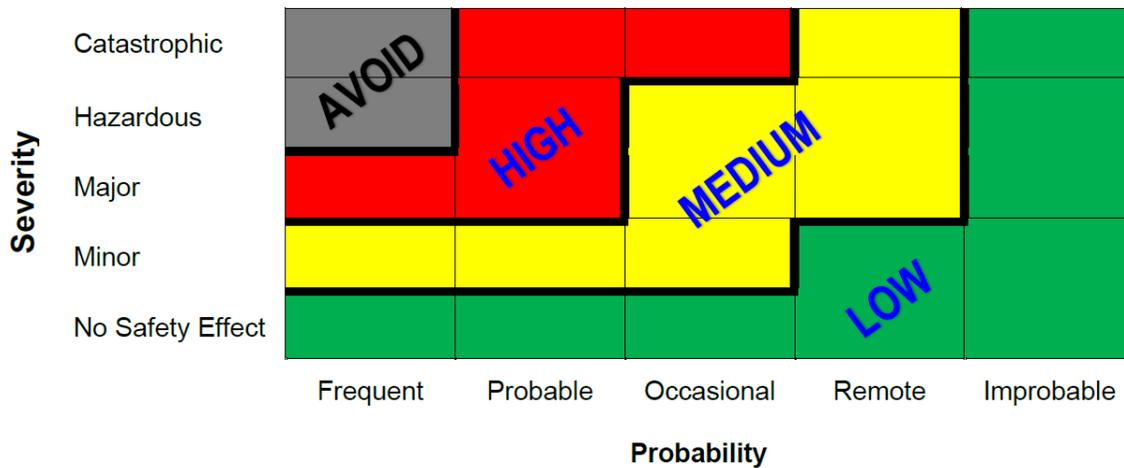


Figure 4 – Another way of describing Figure 3 above.

		PROBABILITY				
		Frequent	Probable	Occasional	Remote	Improbable
SEVERITY	Catastrophic	AVOID	HIGH	HIGH	MEDIUM	LOW
	Hazardous	AVOID	HIGH	MEDIUM	MEDIUM	LOW
	Major	HIGH	HIGH	MEDIUM	MEDIUM	LOW
	Minor	MEDIUM	MEDIUM	MEDIUM	LOW	LOW
	No Effect	LOW	LOW	LOW	LOW	LOW

It may appear that ICAO is using statistical numbers in their matrix but, if you look carefully, all they do is assign numbers to the probabilities and letters to the severity, then they break these combinations into intolerable, tolerable, and acceptable. In the end, figure 1 looks pretty much like figure 3 and 4. Figures 3 and 4 go even further by defining an AVOID category. I personally think the labels of intolerable, tolerable, and acceptable by ICAO are not very practical because in a particular program, you may choose to accept what is tolerable or even intolerable in the ICAO definition. I like the High, Medium, Low definitions for this very reason. This brings us to the argument as to whether you make these assessments BEFORE or AFTER the mitigations are in place. This is one of my favorite subjects and will get into it further down this column.

The question at this point is: What is the value of the 2D Risk Matrix? Like John Hed says and I agree, the value is in the PROCESS of going through this Risk Assessment. Here is how it goes:

1. You determine the Hazards,
2. You determine the consequences of these hazards,
3. You determine the causes for each hazard,
4. You design mitigation for each cause (or consequence),
5. You then design emergency procedures in the case the hazard occurs.

Note: be careful not to mix up what is a hazard and what is a cause or consequence. The best way to avoid this is by having the definitions of hazard/causes/consequences in front of you when you are going through this process (I suggest you use FAA Order 4040.26C for these definitions).

Now, back to what is the value in this process. After you have gone through the process correctly, you should assess whether you have brought down the risk to an ACCEPTABLE level. An acceptable level could be High, Medium, or Low (some tests are just risky). The question now is: When you assign risk levels (High, Medium, or Low) are the mitigations and emergency procedures going to change? The answer is “no” if you were diligent during the process. Then, we might ask, “What is the value of assigning risk levels?”

Here is where my argument of BEFORE or AFTER comes in. The value of assigning risk levels before is simply for your own realization and management visibility. Theoretically, you could come up with mitigations that make all tests Low Risk after the mitigations, then most of the tests would be approved at the lowest levels and management would be missing in the review. Management visibility is a mitigation in itself—it opens the eyes to management, and it forces us to be more diligent in our process. (Translation: do not use boilerplates or just copy and paste.) I emphasize again that the risk assessment at the end is to bring the tests to an acceptable level. Note: Some organizations do their risk assessment at both, before and after the mitigations.

So, back to the main point: should the 2D Risk Matrix use numbers (from some database) or a subjective assessment of the risks. My conclusion is that the use of numbers to assign risk levels is a noble wish if we could have confidence of where those numbers come from, but it is not practical. Realistically, I have more confidence in reaching a reasonable assessment of the risks by depending on corporate judgement (subjective) based on prior experience.

Bottom line: do a diligent and thorough risk assessment for your tests and design credible Test Hazard Analysis (THA) worksheets.

Rodrigo J. Huete

Director Emeritus, Flight Test Safety Committee

## Turbo Talk – Chairman’s Corner

*Art “Turbo” Tomassetti*

Happy New Year to all and I hope 2022 is off to a good start as well as a safe start.

A few administrative things up front. First, the call for presentations is out for our North American Flight Test Safety Workshop. Yes, it is not until May, but now is the time to submit your abstract. If you have done this sort of thing before, great, then consider doing it again. If you have never done a presentation before, OK, just one question: Why Not? This is a great opportunity to share your knowledge and experience. Yes, if you are in or ever have been in the professions of flight test or aviation you have both knowledge and experience. If you think that no one would be interested, you’re wrong. (If it is fear of public speaking no problem, contact me I might be able to help with that.)

Second, we are now taking nominations for the Tony Levier Flight Test Safety Award. You can find out more information on our website for how and what to submit. Leaders, this is part of the job. You owe it to your people to recognize those who do good things and make significant contributions.

In the January podcast I tried to open a discussion about the benefits of simulator mission rehearsal for flight test events. I posed a few questions which included one about how you would operationalize lessons learned. I received a few responses, and thank you to those who took the time to provide. Maybe more will come in and maybe discussions are happening in small groups and teams about the topic. One can only hope. I actually thought this would be an easy one to generate debate. Not on whether doing these types of rehearsals is beneficial (I think that would get nearly universal agreement), but the harder piece to deal with. Operationalizing lessons learned. Sure, we can talk about them in newsletters, podcasts, videos and symposia. But history seems to indicate that may not be enough, as we keep seeing some lessons learned over and over again. While some become policy, requirement, or Standard Operating Procedure there is probably room for more. Talking about lessons learned is actually pretty easy figuring out the right balance between restriction and flexibility, between limitation and latitude.

So your thought exercise for the day: Think of a lesson you have learned or maybe even one you have heard. And just for the exercise let’s keep it aviation related. Is it a good enough one to put into policy or requirement? (If it’s not pick another one). If it is, how exactly would you word it? Assume for this exercise we operate in a world where things like cost, schedule and performance are rather important. If you are so inclined share what you figure out with us.

Until next time: Be Safe, Be Smart and Be Ready.

**Turbo**

## LeVier Award Nominations due 30 March

For more information: <http://flighttestsafety.org/awards/35-awards/information/54-tony-levier-flight-test-safety-award>.

### Subscribe to our Podcast

Rehearsal was the topic of January’s podcast, and Turbo also discusses how to make “Lessons Learned” part of the culture and policy in our organization. It doesn’t take long to subscribe, and it takes even less time to recommend it to a colleague. Please subscribe to the Flight Test Safety Podcast on the [Apple](#) or [Google podcast app](#). You can also navigate directly to the recording in [a web browser](#) and leave comments on these sites.

Podcast: [fscchannel.podbean.com/](https://fscchannel.podbean.com/)

### Contact Flight Test Safety Committee

Art “Turbo” Tomassetti, Chairman  
Susan Bennett, FTSC Administrator  
Society of Flight Test Engineers  
Society of Experimental Test Pilots  
AIAA Flight Test Group

[chairman@flighttestsafety.org](mailto:chairman@flighttestsafety.org)  
[susan@setp.org](mailto:susan@setp.org)  
[edir@sfte.org](mailto:edir@sfte.org)  
[setp@setp.org](mailto:setp@setp.org)  
[derek.spear@gmail.com](mailto:derek.spear@gmail.com)

### Contact Flight Test Safety Fact

Mark Jones Jr, Editor

[mark@flighttestfact.com](mailto:mark@flighttestfact.com)

Website: [flighttestsafety.org](http://flighttestsafety.org)