Podcast 40 - Safety 2

Welcome to the 40th episode of the 737 Talk. This week we'll be going outside of the 737 and discussing safety culture and how the outlook on it is changing, not just in aviation but across other safety critical industries too. We'll be looking through a white paper "From Safety-I to Safety-II by professors Erik Hollnagel, Robert L Wears and Jeffrey Braithwaite, as well as "A tale of two safeties" by Erik Hollnagel which we will of course provide links for over on b737talk.com.

We have been very good in aviation over the years at retrospectively looking at incidents or accidents that have occurred and analysing causes to hopefully learn and prevent future occurrences. This cause-effect model has been prevalent over a long time span, far pre dating the advent of powered flight and we have perhaps clung on to this even though we may have been missing perhaps a better way of looking at safety. That is not to say that this approach doesn't still have its place.

One of the problems with always associating safety and its structures with incidents or accidents is the fact that we have potentially ignored giving our attention to the things that go right. These things become habit and we don't consciously notice the fact of smooth operation as our subconscious expects it to happen, that way it is no surprise when it does, which naturally reduces our attention to those events, or perhaps non-events as we may see them.

How we have perceived this smooth operation previously is that things went well because the system, the interaction between people and technology, worked as it should because nothing untoward occurred. This area is where we are now looking to shine a light and safety 2 shows us the importance of doing that.

Our focus on what goes wrong is re-enforced by regulators and authorities and where our focus on the data we have is concentrated. An example cited in the report is Frankfurt airport where in 2011 there were a total of 490,007 movements with only 10 infringements of separation and 11 runway incursions. This corresponds to an event ratio of approximately 2 in every 100,000 movements. But why did all the others go right? Where is our data on those?

The focus on failures conforms to our stereotypical understanding of what safety is and how it's managed. It is the principle of 'find and fix' to avoid those things that have gone wrong that defines our approach in what the authors term Safety 1.

Consequences of this outlook is that safety is competing for resources with core business meaning that investment in safety is seen as a cost, sometimes making it hard to justify or sustain. The other consequence is the limited learning experiences we have to draw on. Just imagine what we could perhaps have learned from the other 489,986 successful movements in Frankfurt in 2011.

Strategies to achieve these efficient movements would benefit the core business too, thus taking that element of competing for resources away and the amount of data available would mean almost endless learning opportunities would be available.

At the moment though, there are no regulations in place to look at this data with few methods to help us study how human and organisational performance succeeds. Breaking down a tradition can also be hard and creating a successful method or tools to test a new theory will always have teething issues and of course critics, rightly so as what is science without criticism and alternative thought processes.

The maintenance of keeping accidents and incidences to as low a rate as possible is termed Safety 1. ICAO defines safety as "the state in which harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and risk management".

The philosophy of Safety 1 gives a binary view of work and activities where they have either acceptable outcomes or unacceptable outcomes. How we block the transition from one to the other is where the focus is. When the cause is found, it is eliminated and then the outcome should be fewer future events.

In airlines safety 1 philosophy assumes systems work because they are well designed and maintained, procedures are complete and correct and people behave as they are trained to do so. This leads to an expectation of compliance. Human performance variability is therefore seen as a threat and thus safety 1 aims to constrain these variabilities. To do this we have selection, barriers of various kinds, SOPs, rules, and regulations. These are historic defences, justified in the past, but the dramatically changed work environment calls for a different approach.

Management of safety using safety 1 is very much reactive and this has its associated problems. In aviation airlines struggle to keep up with the sheer number of incident reports required by law meaning that even if we only analyse the most serious events, there may still be insufficient time to understand and respond to what happened. In order for this style of management to be effective the airline must be able to recognise and respond very quickly. This can be impossible in an unknown event and also lead to inaccurate responses due to the speed required.

So, what is safety 2 then? Essentially, we are looking at a changing of philosophy from what goes wrong, to what goes right. We are facing complex systems in our work lives now and just like we now have the pilot competencies that can be applied to multiple scenarios we are looking at a safety system that can succeed under varying conditions, which will keep the number of intended successful outcomes as high as possible.

Those sub conscious type activities we spoke of earlier now need to be delved into and understood. Safety 2 assumes systems work because people are able to adjust what they do to match the conditions of the working environment by adapting to overcome design flaws and functional glitches and intervening before a situation becomes an incident or accident event. This is positive performance variability key for safety productivity and shows safety 2 acknowledges the fact that systems are not completely understood and that changes are frequent and irregular.

Staff are seen as an asset rather than a threat, a way we'd all like to see ourselves I'm sure, with the ability to adjust to the actual conditions a strength. In safety 1 we focus on constraining performance variability which will of course have the adverse effect of doing the same where it is desired under the safety 2 philosophy where that same performance variability produces so many successful outcomes.

The remit of safety 2 management is to first acknowledge the existence of performance variability, monitor it and then control it by either dampening it where it is heading in the wrong direction or actively encouraging it if it is heading in the right direction.

Management objectives for safety2 are to ensure that as much as possible goes right. This means a proactive approach with adjustments made before something happens. Early responses will save time and require smaller effort because the consequences of the event will have had less time to develop and spread. That all sounds very nice and easy but in order for that to happen we need to be able to foresee what could happen and have the means to do something about it.

To achieve this, events need to be analysed together looking for patterns and relations across them rather than a 'find and fix' on an individual basis. This will help develop our understanding of a system and how it works in the real world.

The problem of this system could of course be the fact that we prepare for an event that is seen as an acceptable certainty, but that never happens. The report cites the predicted H1N1 flu pandemic of 2009 where governments stockpiled vaccines ensuring they were ready for its impact. This never came and preparations where time and resources were spent ended up being in vain. However, calculated risk is where resources should be focused and the alternative of not being ready when a serious event rears its head will undoubtably be more expensive in both the short and the long run. This is perhaps well illustrated with the lack of pandemic planning that afflicted much of the world when Covid 19 was thrust upon us.

Proactive safety management requires effort up front into thinking about what could possibly happen, to prepare responses, allocate resources, and make contingency plans. Sounds familiar, in inclement weather for example we would discuss the possibility of windshear, our reaction if we had it, what we would expect of our PM and the possibility of delaying the take-off or arrival. All these measures would be far more conducive to a safe outcome than flying off into the storm knowing we have procedures, and then discussing how well we coped with the event after it happened.

It can be more difficult to remain proactive when a series of small-scale events occur in our dynamic work environment causing workload to mount. To try and maintain our proactive environment the report references numerous suggestions including.

1. Learn from both what goes right as well as what goes wrong. Things go well not just because people follow procedures but also because they make sensible adjustments dependant on the day. Look for these adjustments and try to learn from them.
2. Look at what happens regularly focusing on how often things happen rather than how serious they are. Being proactive is easier with something that is frequently occurring

and a small improvement in everyday performance may count more than a one-off large improvement.

3. Give yourself time to reflect and learn from what you have observed. On a small scale, we get quiet periods in the cruise where you could talk to your captain or first officer about what you have noticed and learn from the experience together. On a larger scale an organisational culture needs to be in place where these experiences can be shared. As pilots and crew this could be done through our yearly get togethers for SEPs or tech refreshers where time could be given to discuss times where operations are observed as particularly high performing and look together for the reasons why.

4. Our old friend armchair flying comes into play again here. The report mentions thinking of, or even making a list of, undesirable situations and imagine how they might occur and how you could prevent them or respond to them as they are happening. One technique I find useful is flashcards. You can write scenarios, and they don't just have to be specific to aircraft failures, they can be those little things we deal with daily such as passenger problems, security threats, sick crew members etc. Then take another set of cards and write different times throughout your day. For example, on the way to work, in the crew room, on the ground, during pushback, on take-off, you get the idea. Now pick a card from each pile and run the scenario in your head, discuss it with colleagues and come up with your proactive way of dealing with these events leading to a safe outcome.

We have discussed a lot of theoretical stuff here, with some practical examples, but how are airlines going about putting this into practise. American Airlines has been leading the way on this one and has produced a couple of white papers which we'll put links too over on b737talk.com. Their data collection and further progress has been somewhat hampered by the pandemic, but they've still managed to produce some very useful data and a model that others could perhaps follow.

American Airlines Learning and Improvement team, or LIT, was a new team that tasked itself with implementing the theories of safety 2 into the airline environment. The word Learning was utilised as the team believe every interaction has the power to generate new knowledge.

The team reached out to Erik Hollnagel who suggested separating this programme from current methods and creating their own language with his research as guidance, specifically the resilience assessment grid model which we'll come to later. A further recommendation of devising a method of data separated from the current TEM model was also suggested.

Over a 15 month the timeline the team had 4 objectives:

1. Language and model development
2. Data collection and analysis
3. Solidify program structure and methods
4. Program implementation and data dissemination throughout AA and the aviation industry

The AA LIT teamed up with Ohio State University allowing them to draw on cutting edge theoretical developments. This relationship also benefited the university with its research as it got to access data on Work-as-done rather than Work-as-imagined.

We'll give you a Dr Steven Shorrock definition of that. "Work-as-imagined is both the work that we imagine others do and the work that we imagine we or others did, do, or would do, in the past, present, or future. Work-as-done is actual activity – what people do. It takes place in an environment that is often not as imagined, with multiple, shifting goals, variable and often unpredictable demands"

The AA team utilised the LOSA observation framework but with its own trained observers giving a new and separate data stream to the existing framework. An AA Data Collection Tool specifically designed with safety 2 was utiulised. LIT was aimed at complementing current models such as safety 1, TEM and training philosophy rather than replacing them.

The team also expanded the collection method to include pilot interview sessions called "Shop Talk". These were conducted during downtime at recurrent simulator training or at other scheduled periods. Approximately 30 minutes in length, they provided an opportunity to delve into potentials deeper by examining how the crews analyzed and dealt with unique circumstances such as diverting to another airport, approach go arounds, emergency situations and other instances where the proficiencies could be observed. Following guidance from Klein et al. (1989) Critical Decision Method and Flanagan's (1954) Critical Incident Technique, questions were designed as open-ended probes to spark dialogue to better understand how pilots think and make decisions. Confidentiality of data and anonymity of participants was ensured.

Hollnagel's resilience assessment grid highlights Anticipate, Monitor, Respond and Learn with Anticipate being knowing what to expect, monitor being knowing what to look for, Respond, knowing what to do and learn knowing what has happened.

From these were born the AA LIT model which became Plan, Coordinate, Adapt and Learn. All were termed potentials and then broken down further into observable proficiencies for data collection.

- Coordinate was added to better capture the inherently cooperative aspects of safe airmanship, both within the flightdeck and outside of it. It also reflects the understanding that the LIT is rooted in the front-line work of pilot crews.

- Monitor was removed as a potential in its own right, as it was recognized as a vital component of Coordinate.

- Adapt replaced Respond. Respond implies reactivity, a product behavior which is less considered and less deliberate. Adapt more accurately describes pilot performance, particularly as it relates to the active management of multi-factor trade-offs. and

- Plan replaced Anticipate. While anticipation is an expert skill, recognizing the proficiencies demonstrated by pilots who leverage anticipation to devise plans that create resilient potential better encapsulates the work and mission of all pilots.

Early on, the team struggled with how to model Leadership and Communication, as they were clearly essential among the crews who exhibited positive performance yet seemed difficult to discretely observe or quantify. Eventually, the concepts of Lead and Communicate were removed as directly observable potentials but retained as overarching principles that were present in performance observations. The team continued to assert these concepts as critical to the LIT project and envisioned them as the glue that held the four potentials together.

This model isn't linear but sits with resilience at its centre with the complexities of their interconnectedness and how they spark resilience being further studied.

A proof-of-concept study looked at a particular part of the data involving ATC threats. Three primary factors stood out among the LOSA narratives: the crew's experience level, their knowledge of the theatre of operation, and, perhaps most interestingly, the crew's willingness to say 'Unable' in response to an ATC instruction that was too challenging or placed the crew in a precarious situation. Exploration of what factors positively and negatively contributed to this kind of open communication between pilots readily yielded some teachable best practices. Satisfied that examining data for "what went right or well" yielded leverageable insights, the LIT program was approved by AA's Flight Safety Leadership.

Data analysis of the first nine months of observations is encouraging with experienced AA pilots, AA senior leadership, and aviation experts agreeing that the emerging patterns within LIT data are logical and reflect their lived experiences as flight crew members.

As an example of how this can be incorporated into our training, early ideas include giving examples of resilient performance, case scenarios observed during LIT rides, and insights from LIT "Shop Talk" conversations out to the pilot workforce. Recurrent human factors and quarterly distance learning are also targeted as future platforms to expand LIT principles.

This process also has the potential for introducing change into company procedures and systems. A common theme among interview responses was that a semi-structured conversation between the PF and PM would be beneficial to knowledge discovery and sharing, sensemaking (individual and joint), and collegiality between roles. AA doesn't currently offer a post-flight debrief opportunity but is looking at minimum, to offer support and direction to pilots who wish to engage in a professional, performance-reflective conversation either pre- or post-flight.

Something that is industry wide is the impending retirements of some very experienced crew with the associated influx of new crew. LIT recognises this and seeks to find opportunities to learn from these vastly experienced professionals and use that to provide teaching tools for cadets. LIT also seeks to answer questions of where and how learning takes place on the flightdeck and how it can be facilitated in a light, spontaneous manner, without the tedious

formalisation which we've all been subjected too at times. The findings will contribute to training scenarios in realistic simulation syllabus development while providing fresh material to Human Factors (HF) training courses and Captain's Leadership and Mentoring modules currently mandated by the Federal Aviation Administration (FAA). One of the ultimate goals of the program is to help foster a more extensive and robust learning culture among pilots themselves.

Since the original white paper we just discussed we have seen Covid slow operations across the world. AA recognised the importance of this team to the airline and even through these troubled times has invested in further observers with more promised in the future.

LIT have used this time to refine their language and coding system as well as to lay the groundwork for a more widely recognized construct to view Safety-II within commercial aviation, sharing data throughout the industry. Data has been validated and "cleaned" with some removed due to the fact coding for this area was still in its infancy when collected.

Interestingly an insight that came out of this process was the reinforcement that current LOSA observers with their TEM and safety 1 outlook were not suitable to do the safety 2 observations requiring around 6 months to recalibrate their outlook to look instead for resilient performance during flightdeck observations.

After the initial analysis of the proficiencies during the cleaning effort, it became apparent that the Coordinate potential dominated the dataset: of the four potentials, Coordinate was recorded nearly 50% of the time with the most prevalent Coordinate proficiencies being:

- Briefs or gives new info/update to other member or adds info to build SA (24%) and
- Ask other crew member for input or assistance (12%)

This overwhelming presence of Coordinate was not entirely worrisome as it was congruent with the team's understanding and past research that safe and successful flight operations require a great deal of coordination to routinely achieve successful outcomes. The team theorized that refinement of the two proficiencies in question would allow for more nuanced, and thus more valuable, coding by considering when pilots brief each other and/or ask for assistance from internal vs. external resources. The group also realized that some aspects of briefing could be considered within the Plan potential when it involved returning to revise the original plan based on new information.

Fine tuning of the language for the proficiencies with some separated should give more refined data going forward. Deliberate decisions were made during the iterative development of the new codebook as to which observable examples belonged to which proficiencies.

The codebook is broken down by potentials, the associated proficiencies, and pressures. Each proficiency has examples listed to assist observers in coding observations. While not all-encompassing, the examples serve as a solid foundation for observers to code their observed flightdeck behaviour. This has significantly improved the ease of data collection, standardization, and allows for a more refined data analysis. When I mentioned pressures there that refers to things such as weather impact, ATC, aircraft mechanical, cabin etc.

Shop talks were not carried out during 2020 due to Covid protocols but began again in January 2021. During the analysis of the original first six months of Shop Talks, LIT realized that more structure was needed in the discussions: each facilitator had asked questions on the same subject in slightly different ways, resulting in a lack of consistency in responses. The decision was made to create a standard template to ensure the exact same questions were asked of each participant. Care was taken though to ensure the loose structure of the questions encouraged responses from the participant.

Responses has proven to be challenging due to the flexible format and the variability of follow up questions, making each Shop Talk session slightly varied. However, it was decided the data collected would be more valuable than the risk of data analysis complexity.

In conjunction with AA's training department, LIT leadership decided to explore four avenues of inquiry relating to captains: captain's leadership training, characteristics of leaders, advice for new captains, and challenges faced by captains.

To summarise the findings for classroom training. Captains believe they would learn best in one 8-hour, highly interactive and engaging, low-tech, all-encompassing leadership class with a small class size that they attend immediately upon upgrading. This should focus on things such as who to call, forms to complete, decision making strategies and people skills.

Highly practical, flightdeck based scenarios of real events both mundane and extraordinary, accident/event reports, and experiences of other pilots in the class, especially those new to AA were the preferred topics.

Leadership wise, pilots appreciate leaders who remain up-to-date on protocols and take them seriously, who "walk the walk", lead by example, and are self-aware of their skill. Good leaders communicate confidently yet tactfully, have a sense of humility, initiate congenial small talk, and create a focused yet relaxed atmosphere.

Advice wise. Respondents would advise new CAs to be methodical in stressful situations and to remember their responsibility as a leader to manage the tempo of work. As a word of caution, experienced CAs want new CAs to equate complacency with risk.* New CAs should internalize how important their job is and strive for excellence over and above what the rules and regulations stipulate.

And finally, on the challenges faced CAs find themselves responsible for being aware of, as well as integrating into their work, a great many fast-paced changes in organizational as well as aviation knowledge and information. Sometimes, this new information seems to arrive without context. Current CAs feel pressure to make the right decisions, beyond reproach, constantly and consistently. Nevertheless, they would advise new CAs to leverage teamwork to create success, and that new CAs need not internalize the weight of all decisions made.

Training surveys were also sent out to crews to review the data focused on looking at PF/PM relationship and resilient performance with questions including Do you think the crew performs better when you are pilot flying or pilot monitoring? Why?

To summarize the results, AA crews answered overwhelmingly in favour of learning during cruise and preflight, with the possibility for these conversations to carry over during the layover.

Most crews who completed the survey preferred the role of PM due to their ability to see a bigger picture, thus, gaining greater situational awareness and being able to reduce the workload of the PF. Many of the respondents also noted that their communication is better as PM, keeping the crew on the same page, building resiliency, and having more opportunity to trap threats and mitigate errors. Interestingly, captains surveyed were over 5 times more likely to favour a role as PM, versus FOs who were only 1.7 times as likely. This "desired" crew pairing differs substantially from the observed flightdeck resilient behaviour captured by the AA proficiencies, which shows that CAs as PF and FOs as PM exhibit the most resilient behaviour during actual line flying.

In conclusion, with the foundational work and struggles surmounted, the AA team has confidence that the LIT language and coding scheme is relevant, reliable, and appropriate given the strengths and constraints of the system. LIT's processes to capture the data have been jointly designed and refined with input from leading academic supporters, top-level scientists from NASA, and leaders in other industries.

The knowledge captured is being used to develop more vigorous and comprehensive leadership and mentoring training for American Airlines pilots. This enhanced training is in direct support of the ultimate goal of LIT, which remains to encourage and stimulate a better learning culture in AA's flightdecks, across the entire spectrum of operations.

Through the Learning and Improvement Team's efforts, AA and APA members have successfully provided a new and capable process for the aviation industry to capture data related to resilient behaviour.

This now tested and successful method of gathering data in relation to resilient performance and safety 2 will no doubt begin to be looked at by airlines across the world, hopefully leading to further data creation and analysis giving us a different outlook on safety in our industry. It is in its infancy and perhaps faces a few more tweaks but is yielding promising results. The proof will be in convincing other airlines across the world to implement these systems in a post covid, cash strapped environment.

We believe there are merits to both systems in aviation, as shown by American running them complimentary to each other. The fact we have lacked perspective on one side due to its evolution as our workplace, and world, became more complex leaves us playing catch up but gives us the opportunity to present learning in a more relevant and interesting way thus encouraging higher participation levels and retention in the workforce.

We hope you enjoyed this podcast on what we think is an interesting subject to our industry. We have referenced multiple papers throughout which we will provide links for on the website, and all are worth a read as we have only taken parts to give you a brief summary. We'll return to a more 737 based podcast next time but until then fly well, and be safe.