5-1-2021

How Do Different Knowledge Frameworks Help Us Learn From Aviation Line Observations?

Randall J. Mumaw

Dorrit Billman

Jon Holbrook

HOW DO DIFFERENT KNOWLEDGE FRAMEWORKS HELP US LEARN FROM AVIATION LINE OBSERVATIONS?

Randall J. Mumaw
San Jose State University Foundation
San Jose, CA

Dorrit Billman
NASA Ames Research Center
Moffett Field, CA

Jon Holbrook
NASA Langley Research Center
Hampton, VA

Human performance includes actions that increase safety, as well as actions that can reduce safety. Ensuring safety in complex dynamic operations like commercial aviation depends on the ability to institute appropriate responses based on what is learned from flightcrew performance and the contexts in which it occurs. To do this systemically at the organization level requires collecting data on flightcrew performance, developing effective approaches to analyzing those data, and understanding how to translate what has been learned into policies, procedures, and practice. Systematic observation of front-line operators is a vital source of human performance data. Much has been learned from such observations, including methodological principles. Most observations have been based on a framework focused on managing safety challenges and the ensuing unsafe events. A complementary perspective focuses on flexibility and actions that promote continued safe and effective operation. We consider lessons learned about observational methods from an established framework focused on undesired actions and how these might be extended for a framework focused on desired actions.

Holbrook (2021) makes a case to expand our approach to aviation safety to include the ideas introduced through the concept of Safety II (Hollnagel, 2014), sometimes characterized as "what goes right." Safety II is focused on understanding and documenting how operators (for aviation, the flightcrew) successfully manage the high workload, variable operating conditions, goal trade-offs, limitations on procedures, unexpected threats and disruptions, and other normally occurring challenges to achieving mission objectives safely and efficiently.

The more traditional approaches to safety—that focus on the rare things that go wrong—have not focused on the extensive knowledge and skills that pilots bring to the job to achieve safe, effective operations despite these challenges. In aviation, the term "airmanship" is often a reference to these hallmarks of professionalism and skill. Unfortunately, few have articulated the elements of airmanship, which makes it more difficult to provide to less-experienced pilots through training. One pathway to capturing how flightcrews create safe and efficient flights is through structured observations in a real or realistic setting.

Structured observations can reveal a range of performance dimensions within the current operational context. Further, pilot training, standard operating procedures (SOP), and equipment design need to be held up to the actual demands of line operations; what is sometimes called "work as done" (as opposed to "work as imagined"; see Hollnagel, 2017). Specifically, observation of practitioners in a real or realistic operational setting can provide a deeper understanding of:

- the operational demands, such as pace, time-pressures, forced trade-offs, pop-up tasks from the cabin crew, time zone shifts and their effects on sleep, and workload.
- the demands from operational variability, such as noise in the communications, unfamiliar accents in ATC English, changing weather and visibility, shifts in safety margins, and changing runway conditions and wind direction.
- the threats to efficient and safe operations, such as slam-dunk approaches, thunderstorms, and strong tailwinds.
- the limits of the engineered solutions, such as procedures that may need to be adapted, system alerts that confuse more than clarify an equipment issue, and unexpected autoflight mode changes and consequences.
- the types of knowledge and skill that pilots call on to operate safely and effectively, such as detailed airport knowledge regarding how it typically operates, and flight path management methods that create room for downstream flexibility.
- how flightcrews manage potentially conflicting operational goals, including which goals are prioritized and how safety is maintained when meeting efficiency goals.

Two Frameworks for Guiding Flightcrew Observation

Currently, opportunities to observe flightcrews exist: Airlines routinely put check airmen in the flight deck (real or simulated) as assessors of SOP and operational policies, and airlines observe flightcrew performance through LOSA (Line Operations Safety Audit, see ICAO, 2002), which uses the TEM (Threat and Error Management) perspective to capture work as done. These evaluations focus on the threats present in the operational environment, any wrong decisions and erroneous actions taken by pilots and others, and if/how those get linked to an undesired outcome. For example, a pitot tube becomes clogged by ice while in flight, it generates a set of alerts, the airspeed indication decreases rapidly from the air data system failure, and the pilot, not aware that airspeed indications are invalid, pitches the nose down to regain airspeed. Certainly, it is important to understand how the airplane presents the pitot failure and the pilot's poor response. More broadly, when these types of events occur, it is important to understand pilot response. This is the traditional approach to safety: **How are undesired outcomes produced and how do we prevent or manage those situations?**

The other framework for observation is a "Safety II"-oriented assessment (e.g., American Airlines, 2020). Safety II, as a complement to Safety I, asks instead: **How do flightcrews/pilots manage the variability, disruptions, and threats of normal operations to ensure that mission objectives are achieved?** Applications of Safety II attempt to understand and characterize flightcrew/pilot performance that creates capacity, anticipates problems that may be developing, adapts to the needs of the moment, and learns from operational experience. The focus is not on undesired outcomes but on how flightcrew performance manages the normal (and unexpected) variability in the operational environment.

Lessons about Observational Methods

The current observational protocols, shaped in the Safety I framework, have helped establish basic principles that make observational data more reliable and generalizable:

- A high level of domain expertise – The observer is an interpreter of flightcrew actions and verbalizations. The observer needs to understand the operational situation to see what the flightcrew is faced with and how they should respond, for example, to know that a tailwind can threaten their ability to make an altitude restriction. Observers, therefore, need to be highly skilled practitioners.
- An agreement about which aspects of performance are relevant – Observers are limited in what they can capture about flightcrew performance, especially given the need to capture the larger context and the appropriateness of the actions in that context. The observer has to focus on selected decisions and actions that are relevant to the observation goals, for example, for TEM, the flightcrew's ability to identify and manage an operational threat.
- A set of performance standards – For TEM, the observer is working from an understanding of what the typical flightcrew can do. Given the threat that is presented, what is expected of the flightcrew and did they perform at a requisite level? For example, given a severe failure or environmental hazard, there may be limits on what the flightcrew is able to do to mitigate or manage it. Ideally, these expectations of flightcrew performance are shared across observers to ensure that the observers score consistently.
- An agreement about how observations will be aggregated and reported – Finally, the observations need to be aggregated and reported to influence policy, procedures, training, or some other component of airline operations. This understanding of how the data are gathered and reported needs to be shared across the set of observers to ensure a consistent approach to data collection.

Translation into Safety II: Considerations in Collecting Observational Data

Safety II expands the set of relevant behaviors that could be observed, and it, therefore, becomes even more important to maintain standards. Observation focuses on how the entire flight is managed; acknowledging the complexity and variability of line operations, Safety II captures how pilots "make it all work." In some sense, Safety II is trying to capture the "extra" things that pilots do beyond what is prescribed in the SOP, which is a large behavior space. Hollnagel (2014) has identified four broad categories for Safety II behaviors: anticipate, monitor, respond, and learn. These four categories also suggest a temporal perspective.

- Anticipate: anticipating what might happen in the near future and preparing for the range of situations that may occur.
- Monitor: gathering and integrating information about the operational situation, as it develops.
- Respond: responding appropriately to an event that just happened or an unexpected shift.
- Learn: using experience to expand skills and knowledge from the present to be useful in future operational situations.

For example, observations may capture the following

- Pilot Flying (PF) saw that the Pilot Monitoring (PM) was overwhelmed with radio communications with dispatch on taxi out and used the relief pilot to complete the Before Takeoff checklist.

- The crew added a briefing at 18,000 feet because runway conditions had changed significantly during the initial descent.

Gathering useful observational data demands both an effective coding scheme and trained observers. A coding scheme for complex situations, implicitly or explicitly, requires three interdependent aspects: a prompt or trigger that a codable event occurred, a classification system for categorizing the identified events of interest, and criteria that segment the behavior stream or mark the beginning and end of the event. Safety II presents several challenges to forming an effective coding scheme.

Coding Schemes for Safety II

Prompts. Consistent coding for Safety II is difficult because relevant behaviors may be both frequent and not marked by a single, observable event or prompt. Rather, pilot activity is often driven by internal goals and intentions, which may be expressed in a variety of specific actions, as appropriate in context. Thus, there are these concerns about prompts:

- Prompts may be less connected to a single event – There is not always an easily identified failure or error or environmental threat to prompt an observation. The prompts for the relevant behaviors may lie in normal operational complexity: clearance changes, shifts in workload, changing weather. This variability or complexity is not infrequent. Indeed, it is incredibly routine, and therefore, there is a much larger set of relevant behaviors.
- Prompts may not be directly observable – In many cases, a pilot responds to some anticipated event, such as anticipated workload, anticipated flight path deviation, or maybe anticipated limits to performance by the autoflight system in its current configuration. In other cases, a pilot is attempting to better understand the operational environment by seeking or sharing knowledge, or even just encouraging knowledge-sharing.

Because of the absence of unambiguous behavioral cues, classification may require inferring and identifying an intent and associated behavioral indicators. The mapping between intent and behavior may be quite varied. While the importance of behavioral indicators is well-established, its application here may be relatively difficult. Indeed, it is hard to determine how much the intent effectively cues a search for behaviors or behaviors cue the intent for the observer. For example, the cues to intent might come from a statement about a plan that occurs considerably prior to the event. Because many categories of interest are grounded in intentions, we suggest that useful prompts or triggers may be organized around anticipation and intent.

Classification. Classification may rely on a mapping between intentions and the most commonly used mechanisms for satisfying those intentions. The intention categories may be quite general, though more specific than Hollnagel's categories of anticipation, monitoring, responding, and learning. For example, the PF may intend to distribute workload more evenly prior to entering a busy period by changing who makes entries into the flight management system (FMS). The observer should capture the mechanism for distributing workload (the re-assignment of FMS duties) and the intention of that mechanism (redistributing workload to prevent overload at a later time). The intention captures an important Safety II skill (referred to as a proficiency), and there is value in understanding the various ways that pilots can satisfy that intention. That is, other duties could have been re-distributed to accomplish a workload reduction. While it is easier for the observer to latch onto the mechanism, one might argue that the pilot's intent is, at least, equally important to capture. It shows that the pilot recognized the need to manage workload. However, this approach puts the observer in the position of "seeing" intention, and it suggests a data collection form that is intention-oriented.

Segmentation. Setting the event start and stop boundaries may also pose additional challenges. Effective action is based on both stability of intent across changes that require behavioral adjustment and flexibility to change intent with changing conditions. An intent might be formulated some time before any action is taken and might or might not be expressed when formed. An event might take several actions to complete. Thus, there may be ambiguity and tradeoffs in what forms the unit of activity, and if these are very small, how relations among small events to accomplish an intention should be represented.

Observer Requirements

Capturing both the intention and the action places a heavy burden on the observer. The observer now has to "see" the intention. Notably, the pilot may not have been able to articulate the intention in the moment; it may be too well embodied in the actions. However, in order to capture pilot proficiencies, it is critical to provide the intentional context for the actions.

To achieve this as a scoring scheme, observers will need to be facile with the set of intentions (which should be kept relatively small for this reason). In addition, observers will need strong familiarity with the range of mechanisms (behaviors) that can be applied to fulfil the intention. The data collection tool could make explicit many of these links as an aid. The observers might first see behaviors to recognize intent, or see intent to recognize behaviors, or rely on the interplay. In any case, following training, observers must be able to apply the scheme with acceptable reliability.

Uses of Observer Data

The objective of collecting Safety II observations is to influence operations going forward. Capture and analysis of observations might impact operations through changes in training, procedures, or interface.

To Influence Training

The observational data, ideally, allows one to create a mapping between Safety II-type intentions and the specific behaviors that pilots are using during line operations to enact those intentions. An intent can be realized in many ways, and it may be possible to train at a more-generalized level than individual behaviors. Gathering positive (but not proceduralized) behavior might make "higher order" behavior more visible and thus amenable to training; this might include management of multiple goals such as building pilot experience and immediate operational efficiency or prioritizing different briefing topics. For airlines, this articulation adds another layer of description for skilled performance—on top of "technical" skills, which are focused on flight maneuvers, and Crew Resource Management (Helmreich et al., 1999) skills, such as leadership. Finally, observations of line operations will make trainers better aware of operational environment complexities and variability, which can lead to the development of more realistic training scenarios, as well as better feedback and debriefing.

To Influence Procedures and Interface Design

Initially, observational data may point to areas where flightcrews need to adapt existing written procedures due to operational factors. Eventually, these data may identify ways that the procedures should signal intent and meaning of the individual procedure steps. Alternatively, the data may point to a need to rewrite procedures to be better aligned with how system failures present. Similarly, there is the potential for feedback into interface design. While airplane operators do not determine design directly, they have the opportunity to point out to manufacturers that operational realities make it hard to perform necessary tasks effectively. Perhaps the interface requires too many actions to meet performance time limits. The larger point is that airplane operators have to deal with design deficiencies that are revealed by the operational environment.

Summary and Conclusion

The observational methods that have been refined in the context of Safety I provide a valuable foundation for Safety II methods and observer requirements. Important adaptations for a Safety II framework may derive from the importance of anticipation and intent within Safety II.

Acknowledgments

This work was funded by NASA's System-Wide Safety Project, part of the Aeronautics Research Mission Directorate's Aviation Operations and Safety Program.

References

American Airlines' Department of Flight Safety (2020). Trailblazers into Safety-II: American Airlines' Learning and Improvement Team, A White Paper Outlining AA's Beginnings of a Safety-II Journey.

Helmreich, R.L., Merritt, A.C., & Wilhelm, J.A. (1999). The evolution of crew resource management training in commercial aviation. *The International Journal of Aviation Psychology, 9*(1), pp.19-32.

Holbrook, J. (2021). Exploring methods to collect and analyze data om human contributions to aviation safety, Proceedings of the International Symposium on Aviation Psychology

Hollnagel, E. (2014). *Safety-I and Safety-II: The past and future of safety management.* Farnham, UK: Ashgate.

Hollnagel, E. (2017). Why is work-as-imagined different from work-as-done?, Resilient Health Care, Volume 2: The Resilience of Everyday Clinical Work. https://doi.org/ 10.1201/9781315605739.

International Civil Aviation Organization (2002). *Line operations safety audit (LOSA)*, Doc 9803, AN/761, Montreal: ICAO.