5-1-2021

Evaluating the Use of High-Fidelity Simulator Research Methods to Study Airline Flight Crew Resilience

Chad L. Stephens

Lawrence J. Prinzel

Daniel Kiggins

Kathryn Ballard

Jon Holbrook

Evaluating the Use of High-Fidelity Simulator Research Methods to Study Airline Flight Crew Resilience

Chad L. Stephens[1], Lawrence J. Prinzel III[1], Daniel Kiggins[2], Kathryn Ballard[1], Jon Holbrook[1]

NASA Langley Research Center[1]
Hampton, VA

San Jose State Research Foundation[2]
San Jose, CA

As it evolves, aviation will continue to require integration of a wide range of safety systems and practices, some of which are already in place and others that are yet to be developed. New concepts in system safety thinking have emerged to consider not only what may go wrong, but also what can be learned when things go right during commercial flight operations. Taken together, these complementary perspectives form a more comprehensive approach to system safety thinking that can help to recognize and preserve the resilient performance capabilities currently provided by humans. A need exists, however, for research methods to enable better understanding of the human contributions to aviation safety. NASA's System-Wide Safety Project supports research on using flight simulation methods to study operator resilience and safety-producing behaviors. Building on prior NASA efforts investigating procedural non-adherences during area navigation standard terminal route arrivals, a high-fidelity commercial aviation line operational simulation (LOS) experiment has been designed to study how flight crews anticipate, monitor for, respond to, and learn from expected and unexpected disturbances during these operations. A diverse set of LOS scenarios were developed to simulate highly realistic, complex, but routinely encountered operational situations. Each scenario provided multiple opportunities to collect data on how flight crews manage threats and errors, as well as novel opportunities to observe resilient and safety-producing behaviors. The experimental design, implications for the study of safety-producing behaviors using simulation, and considerations for airline pilot training will be discussed.

New innovative technologies and operational concepts will be required to meet the ever-increasing global demands on air transportation. The NASA System-Wide Safety (SWS) Project is focused on how future aviation advances can meet demand needs while maintaining today's ultra-safe system safety levels. Aviation safety as it evolves shall require new ways of thinking about safety, integrating a wide-range of existing and new safety systems and practices, creating and enhancing tools and technologies, leveraging the access to system-wide data and data fusion, improving data analysis capabilities, and developing new methods for in-time risk monitoring and detection, hazard prioritization and mitigation, safety assurance decision-support, and in-time integrated system analytics (Ellis, Krois, Davies, & Koelling, 2020). To meet these needs, the SWS project has developed research priorities including In-time System-wide Safety Assurance (ISSA) and development of an In-time Aviation Safety Management System (IASMS; Ellis et al. 2020). As part of this effort, the concepts of "resilience" and "productive safety" are being studied (e.g., Hollnagel, 2016). Traditional approaches to aviation safety have focused on what can go wrong and how to prevent it. Another approach to thinking about system safety should reflect not only "avoiding things that go wrong" (protective safety) but also "ensuring that things go right" (productive safety) that together enable a system to exhibit resilient performance.

The SWS project has focused on development of domain-specific safety monitoring and alerting tools, integrated predictive technologies, and adaptive in-time safety threat management (Ellis et al., 2019). Significant research challenges include how to identify data sources and indicators for in-time safety critical risks, how to analyze those data to detect and prioritize risks, and how to optimize safety awareness and safety action decision support. One focus area within the SWS Project is understanding how to evaluate and measure resilient performance and productive safety and application for in-time safety assurance and safety management systems. The research outcomes are intended to both significantly expand the knowledge base of resilience engineering through empirical data collection and analysis, and also help to inform ISSA and IASMS for traditional and emerging operational concepts, such as Advanced Aerial Mobility (AAM; e.g., National Academy of Sciences, 2020).

The challenges associated with ISSA and development of IASMS are significant even for existing air transportation system operations where work-as-imagined and work-as-done can actually be compared. These challenges include collecting productive safety in-time data, granularity of data types and measurement, need for new analytical methods, issues of identifying in-time productive safety metrics and indicators, and potential approaches toward quantification of resilient performance indices. On-going SWS research is focused on application of these concepts for ISSA and design of IASMS, and a test case for this effort concerns non-adherence of area navigation standard terminal arrival route (RNAV STAR) procedures used at major airports. Through initial focused research efforts (i.e., understanding productive safety through test case of non-adherences of RNAV arrivals), the benefits shall provide for a more comprehensive system-wide safety research approach.

Alternative and Complementary Approach for Risk and Safety Management

Stewart, Matthews, Janakiraman, and Avrekh (2018) analyzed aircraft flight track data for more than 10 million flights into 32 domestic airports, which revealed that only 12.4% of flights fully complied with the vertical and lateral profiles on published arrivals. Building on that NASA research, Holbrook et al. (2020) further examined safety producing behaviors during RNAV STAR by collecting data from pilots, from mainline and regional airlines, and terminal radar approach control (TRACON) air traffic controllers. Interviews were conducted to understand how they "anticipate", "monitor" for, "respond" to, and "learn" (Hollnagel, 2014) from routine disturbances during RNAV arrivals into Charlotte Douglas International Airport (KCLT). As reported by Holbrook et al. (2020), different data sources resulted in different estimates of the frequency of RNAV STAR non-adherences at KCLT, ranging from 30% (TRACON controller estimate) to 43% (pilot estimate) to 84% (estimate based on flight track data specific to KCLT arrivals collected for Stewart et al., 2018).

These previous findings highlight how published procedures can be misaligned with routine and safe operations. The reasons for the misalignment and interpretation of these findings cannot be addressed with traditional approaches to risk and safety management. An alternative and complementary approach for risk and safety management is necessary to explore whether non-adherences may reflect desired safe behaviors. Examples from aviation operations such as those reported by Stewart et al. (2018) and Holbrook et al. (2020) indicate that the definition of safety should reflect not only “avoiding things that go wrong”, but also “ensuring that things go right”. Global demands on air transportation drive increasingly complex operations, and to maintain safety, humans in the system continuously adjust their work to match their operating conditions (Hollnagel, 2014).

Proposed Human-in-the-Loop (HITL) Flight Simulation Study

Currently, the resilience engineering literature and very limited empirical published research have focused almost entirely on the conceptual aspects of productive safety. There remains a need to systematically collect empirical data to explore the practical application of these concepts toward improving aviation safety. The current research study is intended to meet that need with identification and collection of data sources and indicators for in-time safety critical risks, examination of how to analyze those data to detect and prioritize risks, and specifications for how to optimize safety awareness and safety action decision support toward development of ISSA and IASMS, for both traditional and future aviation concepts of operations.

Research Questions

The central research question is, “how do commercial airline pilots manage routine contingencies and productive safety during RNAV arrivals?” The present study was designed to identify and capture real-world operational behavior through replication of known actual line operational events that have occurred at KCLT in which observable resilient behavior had been described. A “structured observation” methodology (Gray, 2013) was chosen for this study to allow for careful observations of specific behaviors in a setting that is more structured than the settings used in naturalistic and participant observation. A structured observation approach, when combined with other methodologies (e.g., interviews) has significant benefits, particularly in comparison to naturalistic observation, by providing for costs, time, access, safety, and validity controls needed to meet the experimental objectives.

A challenge in studying such events in line operations is the limited data that can be collected and coded for the purposes of productive safety research for ISSA and ISAMS, or that is unavailable for collection or analysis for various logistical, procedural, or regulatory reasons. The proposed research study’s primary objective is to obtain a comprehensive data set of identified candidate measures, in order to facilitate anticipated data science efforts and to help better understand the phenomena of interest.

Experimental Design Considerations

The proposed study will investigate how pilots respond to expected and unexpected disturbances during RNAV arrivals. Boeing 737NG rated professional commercial pilots from a major airline will be recruited to perform multiple RNAV STAR arrivals and data will be collected with regard to how they anticipate, monitor for, respond to, and learn from routine disturbances during RNAV arrivals into KCLT. The purpose of the present paper is to describe the methodology for this experimental study of productive safety that requires high-fidelity simulation of commercial aircraft line operations and scenario constructions that enable collection of these data types. A validation of the scenarios was performed with carefully screened participants in an off-site high-fidelity flight simulation facility that enacted substantial

COVID-19 participation protection protocols. The objective was to assess the efficacy of the scenarios for the study in the NASA high-fidelity simulators with airline participants. The remainder of the paper describes the testing methodology that shall be utilized in the planned NASA study.

Study Test Participant Considerations

Twelve (12) 737-800 Part 121 commercial airline flight crews (24 pilots) shall be the test participants for the study. The pilots shall be recruited to serve as a flight crew in respective roles (Captain, First Officer) and have familiarity with KCLT RNAV arrivals. The selection of 737-800 pilots is to ensure high familiarity with the NASA Langley Research Center Cockpit Motion Facility (CMF) 737-800 high-fidelity simulator (Figure 1).

Figure 1. NASA 737-800 High-Fidelity Full-Motion Simulator

Flight Simulation Scenario Considerations

Seven scenarios were constructed based on current KCLT RNAV arrivals. The scenarios were designed to simulate anticipatable and un-anticipatable "routine" disturbances that are well-documented for occurrence in Aviation Safety Reporting System (ASRS) and Aviation Safety Action Program (ASAP) reports, airline crew reports, and known high frequency events that occur during KCLT operations (e.g., weather, traffic). Each of the experimental scenarios included documented events in which flight crews have exhibited resilient performance in response to the disturbance(s) encountered. The scenarios each have two events that were designed to present an "opportunity" for the flight crew to evince "anticipate", "monitor", "respond", and/or "learn" behaviors.

An important component of each of the scenarios is the ecological validity of the emulation of the scenario to replicate real-world line operations, including dispatch releases, weather reports (e.g., TAF, METAR, WSI), required and unanticipated cabin calls, dispatch communications (e.g., ability to contact dispatch including data communications and ACARS and text-to-speech capability), live air traffic control (TRACON, APPROACH), and other aspects, often neglected in research studies but critical to replicating actual operating conditions. Essentially, the scenarios and simulation environment are designed to provide high fidelity re-creation of commercial line operation arrivals into KCLT based on the following event categories: (a) energy management; (b) traffic compression and high flows; (c) convective weather; (d) unanticipated tailwind; (e) autoflight issues; (f) icing conditions and ice crystal icing; (g) system caution-level events; (h) wake encounter during arrival descent; (i) ATC and/or pilot clearance errors (e.g., hearback/readback error); and (j) high workload.

Each scenario begins after top-of-descent (TOD) and the aircraft is positioned on the RNAV arrival track with operationally appropriate attitude and airspeed. Pilots shall be provided with a detailed synopsis of the scenario before scenario start (including time to conduct a detailed arrival briefing and any FMS entries, etc., nominally completed prior to TOD in line operations). Dispatch paperwork has been created for the simulated RNAV arrivals. Since all airlines have different paperwork, our dispatch paperwork is a conglomeration of different airline formats with all the required information included. The dispatch release, Notices to Airmen (NOTAMS), Terminal Area Forecasts (TAF), Meteorological Terminal Aviation Routine (METAR) report, terminal weather, Automatic Terminal Information Service (ATIS), and other typically available flight/weather information shall be provided.

The pilots will perform as a flight crew in respective roles at the airline. All airline company hardcopy (plastic) checklists and other normally available documents shall be provided. All pilots shall also utilized their company supplied tablet which contains all necessary plates, aircraft reference manuals, Quick Reference Handbook (QRH), etc. All required company arrival briefings, standard calls, all standard operating procedures, etc., that would nominally be conducted in actual line operations shall be conducted in the study.

For this study, the key research questions for characterizing and measuring productive safety for in-time system-wide safety assurance include: (a) What data can and should we collect and analyze to understand existing productive safety capabilities?; (b) How can we measure the productive safety capability of a system?; and (c) How can productive safety support safety assurance of emerging systems?

Scenario Validation

The research study was designed based on a "structured observation" methodological approach combined with dialog, interview, and observer-based rater data and analyses. One goal of the research is to develop a system-level framework/taxonomy to understand operator's resilient performance, and develop organization-level strategies that promote recognition and reporting of resilient performance. Therefore, key to this work is understanding what those data are, how to collect them, and how to utilize them for in-time system-wide safety assurance and emergent risk prediction. The study attempts to contribute to the development of new metrics based on quantification and measurement of behaviors that support resilient performance through, in part, conducing high-fidelity simulation of "work-as-done" in traditional commercial airline operations.

The preliminary check-out validation of the scenarios was conducted at the Boeing Miami B737-800 flight simulator facility with active commercial airline pilots highly experienced with RNAV arrivals into KCLT. Using COVID-19 screening and safety protocols, four commercial airline flight crews were required to don personal protective equipment that limited the realism in simulating actual line operations in addition to the inability to fully implement all the scenario aspects that are critical to the study of productive safety. Despite these limitations, the validation flight crews confirmed the capability of these scenarios to achieve the experimental objectives. Based on SME preliminary assessment, the flight crews that were highly responsive early in the scenarios and exhibited resilient behaviors (i.e., monitor, anticipate, respond, learn) were better able to address potential threats well before they emerged to become significant hazards and additional data analyses are ongoing to confirm this finding. Review by subject matter experts and post-hoc discussions with the pilots confirmed the efficacy of the scenarios to enable simulation study of both productive safety and more traditional threat and error management (e.g., line operations safety audit type observable behaviors).

Conclusions

A challenge intended to be addressed in this research study (and project, more generally) is that data from system, observer, and operator sources are rarely (if ever) all available for the same set of events. Thus, there is little opportunity to explore the integration of factors that contribute to operators' resilient performance. Additionally, existing methods for measuring resilient performance are immature. Safety monitoring, prediction, and mitigation technologies based only on hazards and risks will address an incomplete picture of safety. Furthermore, the low frequency of undesired outcomes may impact the temporal sensitivity of safety assessments, challenging the notion of "in-time" mitigation. Building a more thorough, data-rich, and representative understanding of safety is needed to achieve NASA's vision of in-time system-wide safety which includes developing methods to enable the systematic study of productive safety. The high-fidelity simulation research study described here is an important step forward toward this goal for aviation safety, and the NASA team is prepared to begin data collection impacted by the unprecedented situation the pandemic has presented and has impacted so many involved in scientific study involving human participants.

Acknowledgments and Notes

This work was funded by NASA's System-Wide Safety Project, under the Aeronautics Research Mission Directorate's Aviation Operations and Safety Program. Dr. Lawrence (Lance) Prinzel is the corresponding author, but has co-primary authorship with Mr. Chad Stephens who contributed significantly to the paper.

References

Ellis, K., Krois, P., Davies, M., Koelling, J. (2019). In-Time System-Wide Safety Assurance (ISSA) Concept of Operations. ARC-E-DAA-TN74516-1. Moffett Field, CA: NASA Ames Research Center.

Gray, D. E. (2013). Ethnography and participant observation (Chapter 17). Doing research in the real world. London: Sage.

Holbrook, J., Prinzel, L. J., Stewart, M. J., & Kiggins, D. (2020). How do pilots and controllers manage routine contingencies during RNAV arrivals? In Proceedings of the 11th International Conference on Applied Human Factors & Ergonomics.

Hollnagel, E. (2014). Safety-I and Safety-II: The Past and Future of Safety Management. Farnham, UK: Ashgate.

Hollnagel, E. (2016). Resilience Engineering. erikhollnagel.com/ideas/resilienceengineering.html. Accessed March 1, 2021.

National Academies of Sciences, Engineering, and Medicine 2020. Advancing Aerial Mobility: A National Blueprint. Washington, DC: The National Academies Press. https://doi.org/10.17226/25646

Stewart, M., Matthews, B., Avrekh, I., Janakiraman, V. (2018) Variables influencing RNAV STAR adherence. IEEE/AIAA. Proceedings of the 37th Digital Avionics Systems Conference (DASC). London, UK.