

CREW STATE MONITORING AND LINE-ORIENTED FLIGHT TRAINING FOR ATTENTION MANAGEMENT

Chad Stephens, Lawrence Prinzel, Angela Harrivel, Ray Comstock, Nijo Abraham, and Alan Pope
NASA Langley Research Center
Hampton, VA

James Wilkerson
Boeing Flight Services
Seattle, Washington

Daniel Kiggins
National Institute of Aerospace
Hampton, VA

Loss of control – inflight (LOC-I) has historically represented the largest category of commercial aviation fatal accidents. A review of worldwide transport airplane accidents (2001-2010) indicated that loss of airplane state awareness (ASA) was responsible for the majority of the LOC-I fatality rate. The Commercial Aviation Safety Team (CAST) ASA study identified 12 major themes that were indicated across the ASA accident and incident events. One of the themes was crew distraction or ineffective attention management, which was found to be involved in all 18 events including flight crew channelized attention, startle/surprise, diverted attention, and/or confirmation bias. Safety Enhancement (SE)-211, “Training for Attention Management” was formed to conduct research to develop and assess commercial airline training methods and realistic scenarios that can address these attention-related human performance limitations. This paper describes NASA SE-211 research for new design approaches and validation of line-oriented flight training (LOFT).

Recent accident and incident data suggests that Spatial Disorientation (SD) and Loss-of-Energy State Awareness (LESA) for transport category aircraft are becoming an increasingly prevalent safety concern in all domestic and international operations (Commercial Aviation Safety Team, 2014a). SD is defined as an erroneous perception of aircraft attitude that can lead directly to a Loss-of-Control Inflight (LOC-I) event and result in an accident or incident. LESA is typically characterized by a failure to monitor or understand energy state indications (e.g., airspeed, altitude, vertical speed, commanded thrust) and a resultant failure to maintain safe flight.

Commercial Aviation Safety Team (CAST) Analysis of LOC-I

A Commercial Aviation Safety Team (CAST) study of 18 LOC-I events determined that issues with flight crew attention were involved in all of the 18 events. CAST created a research “Safety Enhancement” (SE) specifically to address this problem state as identified in the CAST JSAT (Joint Safety Analysis Team) and JSIT (Joint Safety Implementation Team) analyses (CAST, 2014a). It was recommended that the aviation community (government, industry, and academia) should conduct research on methods for understanding the phenomena of flight crew channelized attention, startle/surprise, diverted attention, and confirmation bias. In response, NASA initiated a sub-project under the Airspace Operations and Safety Program (AOSP), “Technologies for Airplane State Awareness”, to address this SE and others. The research described in this paper specifically addresses SE-211, “Training for Attention Management”.

Training for Attention Management. CAST recommended research and training organizations develop methods to detect and measure attention-related human performance limiting states (AHPLS). Furthermore, research organizations should work with industry partners (air carriers, manufacturers, and

commercial training providers) to develop methods and guidelines for creating training scenarios that induce AHPLS and develop and assess potential mitigations to these issues in the training environment. The “detailed implementation plan” for SE-211 (Commercial Aviation Safety Team, 2014b) described two key tasks, assigned to NASA, which were: 1) the development of valid methods to detect and measure AHPLS in pilots; and, 2) the development of methods for creating realistic, high workload scenarios that can induce human performance limitations, including channelized attention, startle/surprise, diverted attention, and confirmation bias.

Scenarios for Human Attention Restoration Using Psychophysiology (SHARP)

The SHARP study was conducted at NASA Langley Research Center in the spring of 2016 and consisted of multiple facets to assess crew state monitoring measures (Harrivel, et al., 2017) and the induction of AHPLS via benchmark tasks and a line-oriented simulation (LOS) scenario. Data collection was performed in the Research Flight Deck in the Cockpit Motion Facility at NASA Langley. The simulator has full-mission, Level D type capabilities and the flight deck emulates a B-787, but with a B-757 aerodynamic model. The crew state monitoring (CSM) data will be used, post-test, in the development of classification methods for detecting AHPLS.

A LOS scenario was designed to provide a high-fidelity simulation of line operations with event sets designed to induce channelized attention and startle/surprise; the CSM and pilot qualitative and quantitative data collected during the event sets will be used for purpose of validating AHPLS classification algorithms during LOFT. This paper describes the LOS scenario results. The results of analyses on the crew state monitoring (CSM) measures captured during the benchmark tasks and during employment in the LOS scenario are reported elsewhere (Harrivel et al., 2017, and Harrivel et al., 2016).

NASA Langley Research Center subject matter experts (SMEs) and line-operational commercial airline pilots with combined experience of more than 30 years designed the LOS event set. The scenario was also developed by reference to FAA Advisory Circular (AC) 120-35D (Federal Aviation Administration, 2015a) which presents the guidelines for the design, implementation, and validation of LOFT. The LOS used a gate-to-gate (from pushback to taxi-in) scenario with multiple event sets designed to induce startle/surprise and channelized attention AHPLS.

Twelve flight crews were paired based on pilot role (Captain, First Officer) and same airline. Each flight crew averaged 22,000 hours of experience with both the B-757 and B-787 aircrafts. The LOFT scenario included debriefing, dispatch paperwork, and other materials and instruction that airlines typically provide for LOFTs (based on two major airlines and manufacturer that had partnered with NASA for this research).

LOFT Scenario Events

Wake Hazard Event. Following the taxi-out, the first major event consisted of a wake encounter which occurred at 700 ft. mean sea level (MSL) after take-off from Runway 36L at Memphis (KMEM). The event created a startle state due to an aircraft roll upset at low altitude. The simulated wake encounter aerodynamic behaviors were verified by SMEs and calibrated by line-operational commercial airline pilots who had each experienced similar low-altitude wakes.

Hydraulic System/Anti-Skid Failure Events. The second major event was a right hydraulic system pressure and antiskid failure approximately 20 nmi. from the LEOOO waypoint on the BBKNG 4 departure. The event set was designed to induce channelized attention on the part of the flight crews by requiring an extensive sequence of checklist items and decision-making considerations (e.g., alternate airports, systems integrity, landing/stopping distances, availability of controls and gear, etc.). The event set provided behavioral indicator checks to determine whether the flight crew was channelized on the basis of: (a) communication patterns and verbiage and attentional management toward other activities (e.g., Air Traffic Control, ATC, responses); and, (b) detection of “proximate” traffic that was also heading to the LEOOO waypoint. The potential incursion traffic was an aircraft that departed from Runway 18C

(the scenario design allowed for both north and south traffic departure flows) and party-line communications were provided that indicated that the traffic was cleared to the LEOOO waypoint at altitude that conflicted with the own ship. The event set was designed to cause a proximate traffic encounter if pilots were channelized in attention, since the traffic was observable and appropriate mitigation responses could be performed (e.g., contact ATC) well before the encounter. The traffic was clearly visible on the navigation display for the entire duration of the event set and SMEs predicted that the traffic should be detected 100% of the time under normal operations and conditions (note: depending on how the flight crews navigated and managed the situation, the incursion traffic could become a Traffic alert and Collision Avoidance System, TCAS, “caution”).

Trailing Edge Flap Asymmetry Event. The third major event was a trailing edge flap asymmetry (TE FLAP ASYM) which occurred after flight crews were directed back to KMEM for approach to Runway 36L following the hydraulic leak. Runway 36C is the longer runway at Memphis, but the scenario had the runway occupied and unavailable due to foreign object debris that was on runway. Because of the weather conditions and poor braking action reported, flight crews had significant cognitive overhead when deciding whether to accept the runway assignment or request to go an alternate airport due to the aforementioned hydraulic leak and antiskid failures. There were significant variations in how flight crews handled the decision and problem-solving and exhibited threat and error management. However, all flight crews eventually accepted an approach to Runway 36L.

During the approach, the trailing edge flap asymmetry event occurred; the flap asymmetric deployment was alerted to the crew on the Engine Indication and Crew Alerting System. The checklist allowed for a flight crew decision to continue the landing based on the flap configuration but most flight crews requested a go-around and executed the missed approach. For those that elected to continue, ATC issued a go-around (traffic was reported on the runway). The event, combined with the existing issues, was designed to induce channelized attention due to the temporal demands and decisional factors that needed to be considered once the event occurred (e.g., electronic checklist, decision to go-around or land, etc.) The amount of cognitive effort was high during the timing of the event (which went caution alert to the go-around and clean-up and climb to Hold), regardless of whether the pilots initiated the go-around and contacted ATC or ATC issued the go-around, to include the subsequent crew coordination, clean-up of aircraft, and discussion on option. The exception were the four flight crews that immediately executed the missed approach after the TE FLAP ASYM caution was presented on engine indication and crew alerting system (EICAS) display (see discussion below).

Missed Approach Event. After initiating the Runway 36L missed approach, the flight crews climbed and then leveled-off at 3000 ft. on the runway heading and then were turned to a heading of 330 and instructed to proceed to the KALIE waypoint to hold at 5000 ft. ATC then gave vectors to return to KMEM Runway 36C (the longer runway that all pilots preferred earlier was now available). Flight crews were provided speed and vectors to the ILS 36C approach. Due to the trailing edge flap asymmetry, the approach speed was significantly higher than normal (186 knots indicated airspeed).

Runway Incursion Event. The Runway 36C runway incursion event was designed to induce startle/surprise. The incursion was triggered by an aircraft that had erroneously crossed the active runway. Because the landing speed is higher than nominal approach, the reaction time to such an event was reduced creating the conditions for a startle/surprise response. The aircraft timing was intended to purposely not cause a collision on the runway but to simulate a Category B runway incursion event. Due to flight crew decisions or timing issues, in a few cases, the runway incursion aircraft was blocking the runway when the aircraft landed; in such cases, the pilots were briefed that the event was not as intended.

ATC Taxi Clearance Event. After the flight crew turned off the runway, ATC instructed the aircraft to hold on the taxiway and contact ground. Ground ATC issued a plausible and almost correct taxi clearance that would require the flight crews to carefully consider the path prior to execution to avoid an error in taxi. Depending upon their exit, they were either given a taxi clearance which crossed a

runway (without a hold short of or clearance to cross the runway in the ATC taxi clearance) or were given a clearance that had a discontinuity (i.e., the cleared route omitted a taxiway). If the flight crews communicated that they had an issue with the clearance, ATC immediately corrected it. It is standard practice for pilots to immediately read-back the clearance to ATC verbatim (which in this case was an intentionally generated ATC error), or 'Roger' or call sign or other (which is not recommended SOP but this would not be marked as an error if done), but they then should review and verify the route on chart. Often, this is done while the aircraft is taxiing, but in this case the aircraft was stopped on taxi-way and there were no temporal pressures to begin taxi until the pilots were ready (due to the runway incursion event ATC had located the aircraft where they were a non-issue for other aircraft and ATC told the pilots they could begin when ready). However, if the flight crew did not identify the error and contact ATC before taxiing, this was not considered as an error; only, if the flight crews did not detect the routing deficiency prior to arriving at the route error was it marked error (recognizing that the original error was ATC).

Discussion

Quantitative and qualitative data were collected from the crews to assess the efficacy of the scenario to illicit realism, training effectiveness, and AHPLS. The qualitative data results for the LOFT scenario evince that the LOFT scenario was rated to be "excellent" / "very good" (82%) with 68% of pilots responding that NASA LOFT scenario was of higher quality than airline LOFT scenarios they had experienced. The NASA LOFT scenario was also judged "very good" to "excellent" for all pilots' responses in comparison of realism to actual commercial flight operations and these hazards encountered on the line.

Startle/Surprise

Wake Encounter Event. The LOFT scenario was found to be highly effective to producing startle/surprise responses for the wake encounter event set - 58% of Captains and 33% of First Officers exhibited behavioral indicators of startle/surprise (based on SME video analyses). Participant pilots rated the wake encounter as 4.5/5 on the Wake Vortex Encounter (WVE) questionnaire (Ahmad et al., 2014) for realism. The WVE data ranged from pilot ratings of 'Minor' (2 responses), 'Major' (18 responses), to 'Hazardous' (6 responses) in effect. Pilots reported that roll angle and roll rate (20 out of 26 responses) was the most significant parameter identifying the disturbances as a wake. Pitch angle and rate (6 out of 26 responses) was also indicated as significant parameter. Pilot comments validated that the simulated wake event was realistic and similar to those operationally encountered.

Runway Incursion Event. The LOFT scenario was also found to be highly effective to producing startle/surprise responses for the runway incursion event set; 42% of Captains and 33% of First Officers displayed behavioral indicators of startle/surprise. Jones and Prinzl (2011) reported on a set of standard dependent measures used in runway incursion research based on the Runway Incursion Severity Index (Federal Aviation Administration, 2015b). The LOFT scenario event was designed to be a "pilot deviation" event (cross hold line on active runway of other traffic) - a Category B type runway incursion - requiring the flight crews to make corrective/evasive action to avoid a collision but was not expected to result in a collision unless the flight crew exhibited poor attention management. Post hoc analysis, based on the FAA Runway Severity Index Rating, of video of the 10 crews who experienced an incursion showed that 4 were rated as Category A events, no Category B, 6 Category C, and 2 Category D events. These data support that these highly experienced flight crews were mostly effective at recognizing and preventing a more serious runway incursion situation.

ATC Taxi Clearance Event. The ATC taxi clearance error event set demonstrated that approximately half of the flight crews accepted the erroneous taxi-in clearance without cross-checking and verification. The error was not that the flight crews read-back of the erroneous ATC clearance and ATC confirmed the read-back, but that the pilots were told to stop on the taxiway after runway turn-off and to contact ground and, therefore, were given ample time to review route before starting taxi again. It

is standard practice for pilots to read-back the clearance upon receiving it, but to then to after review the route on the charts to ensure it is correct (often while taxiing where the pilot-not-taxiing reviews the route on the chart) and that they know where they are going. There were no temporal demands on the pilots, as there often are at major airports, and the event was not meant to be a major safety event although one flight crew had taxied onto the active runway before stopping beyond the hold line before contacting Tower. The results evinced that those pilots that experienced the highest channelized attention and startle/surprise responses previously during LOFT did not review, or did so only cursorily, the taxi-in route before or during taxi; these flight crews only realized the ATC error when they came to the mistake in the route. The results suggest that the effects of startle/surprise and channelized attention can continue after the event even when pilots had substantial opportunity to stop and reset without significant temporal demands.

Channelized Attention

Hydraulic System/ Anti-Skid Event. The first channelized attention event set was highly effective to induce channelized attention. 92% (11/12 flight crews) did not detect the proximate traffic and in several cases, a TCAS ‘caution’ alert was generated due to the attentional focus required by the complex and lengthy electronic checklist.

Trailing Edge Flap Asymmetry. The second channelized attention event set was marginally effective owing largely to the highly variable nature of scenario segment which, to maintain realism, allowed degrees of freedom for pilot responses; as consequence, the trailing edge flap asymmetry and behavioral indicators did not always manifest themselves in the LOFT scenario. 42% (5/12 flight crews) showed evidence of channelized attention. Half of the flight crews did not complete the scenario event set segment as crafted so they did not encounter the event mechanisms designed to induce AHPLS.

Qualitative and Quantitative Pilot Performance

Overall, the flight crews exhibited acceptable threat and error management (e.g., Maurino, 2005) Human Factors Training Manual Doc 9683, and NOTECH or non-technical skills (e.g., Flin et al., 2003) and line/LOS behavioral markers (e.g., Kanki, Helmreich, and Anca, 2010) were found to be “acceptable” to “good” across all the commercial pilots (based on SME video analyses). Pilot technical standards were found to meet the FAA published standards (FAA-S-8081-5F, 2008) and were evaluated against the performance standards for each phase of operations during the LOFT scenario.

Pilot responses to an extensive and detailed final questionnaire provided a wealth of data in terms of current LOFT scenario implementation at airline training centers and substantial information for further work for SE-211. The questionnaire revealed significant and valuable data for how to enhance LOFT scenario and implementations and potential avenues to explore for further scenario development specific to construction of training for attention management scenario and related constructs (confirmation bias, diverted attention, startle/surprise, and channelized attention).

Future Directions

Although not discussed here, analyses are on-going to compare these AHPLS behavioral responses to the CSM classification data. The subjective data suggests that there are a number of potential other or additional opportunities to implement and assess the CSM data for AHPLS, including diverted attention, within the LOFT scenario. Communication analyses (Kanki, Lozito, and Foushee, 1989) are on-going for each event set and the overall LOFT to add additional behavior markers for this characterization/classification. A substantial amount of data cannot be fully described here within the space available, but the results show that LOFT scenarios can be effectively designed to induce AHPLS.

The data suggests that LOFT sessions may have more value if event sets were used with the goal of training pilots to combat AHPLS rather than focus on the event set itself (e.g., training on runway incursion mitigation). Results described in Harrivel et al. (2017) suggest that CSM methods and

approaches may be useful in the validation of event sets and potentially for real-time analysis during LOFT sessions. Harrivel et al. (2017) describe the CSM benchmark classification results and similar analyses that are being conducted.

Future directions include additional LOFT scenario evaluation with events sets designed to induce other AHPLS, including diverted attention and confirmation bias. Airline and major aircraft manufacturer training centers have partnered with NASA to continue to improve the design of LOS design and training methods. The planned efforts include evaluation of scenario event sets and recommended approaches during actual airline training LOFTs for training AHPLS.

References

- Ahmad, N., VanValkenburg, R., Bowles, R., Duparcmeur, F.L., Gloudemans, T., van Lochem, S., and Ras, E. (2014). Evaluation of fast-time wake vortex models using wake encounter flight test data. Atmospheric and Space Environments Conference. Atlanta, GA. June 16-20, 2014.
- Commercial Aviation Safety Team. (2014a). Airplane State Awareness Joint Safety Analysis Team Interim Report. Retrieved from http://www.skybrary.aero/index.php/Commercial_Aviation_Safety_Team_%28CAST%29_Reports.
- Commercial Aviation Safety Team. (2014b). CAST Safety Enhancement SE 211 detailed implementation plan (DIP). Retrieved from <http://www.skybrary.aero/bookshelf/books/2541.pdf>.
- Federal Aviation Administration. (2015a). AC 120-35D - Flightcrew Member Line Operational Simulations: Line-Oriented Flight Training, Special Purpose Operational Training, Line Operational Evaluation. Washington, DC.
- Federal Aviation Administration (2015b), FAA Runway Safety Report 2013-2014. Department of Transportation, Federal Aviation Administration, Office of Runway Safety, Washington DC. May, 2015.
- Flin, R., Martin, L., Goeters, K. Hoermann, J., Amalberti, R., Valot, C., Nijhuis, H. (2003). Development of the NOTECHS (Non-Technical Skills) system for assessing pilots' CRM skills. *Human Factors & Aerospace Safety*, 3, 95-117.
- Harrivel, A. R., Liles, C., Stephens, C., Ellis, K., Prinzel, L., & Pope, A. (2016). Psychophysiological Sensing and State Classification for Attention Management in Commercial Aviation. Paper presented at the American Institute of Aeronautics and Astronautics, SciTech, San Diego, California.
- Harrivel, A. R., Stephens, C. L., Milletich, R. L., Heinich, C. M., Last, M. C., Napoli, N. J., Pope, A. T. (2017). Prediction of Cognitive States during Flight Simulation using Multimodal Psychophysiological Sensing. Paper presented at the American Institute of Aeronautics and Astronautics, SciTech, Grapevine, Texas.
- Jones, D., & Prinzel, L. (2011). Simulator evaluation of runway incursion prevention technology for general aviation operations. NASA-TP-2011-217045. Hampton, VA: NASA.
- Kanki, B., Helmreich, R., & Anca, J. (2010). *Crew Resource Management (2nd Ed)*. Amsterdam: Academic Press.
- Kanki, B., Lozito, S, & Foushee, H.C. (1989). Communication indices of crew coordination. *Aviation, Space, and Environmental Medicine*, 60(1), 56-60.
- Maurino, D. (2005). Threat and Error Management (TEM). Canadian Aviation Safety Seminar (CASS), Vancouver, Canada, 18-20 April 2005.