QUANTIFYING PILOT CONTRIBUTION TO FLIGHT SAFETY DURING AN IN-FLIGHT AIRSPEED FAILURE

Timothy J. Etherington
Rockwell Collins
Hampton, Virginia
Lynda J. Kramer
Randall E. Bailey
Kellie D. Kennedey
NASA Langley Research Center

Accident statistics cite the flight crew as a causal factor in over 60% of large transport fatal accidents. Yet a well-trained and well-qualified crew is acknowledged as the critical center point of aircraft systems safety and an integral component of the entire commercial aviation system. A human-in-the-loop test was conducted using a Level D certified Boeing 737-800 simulator to evaluate the pilot's contribution to safety-of-flight during routine air carrier flight operations and in response to system failures. To quantify the human's contribution, crew complement was used as an independent variable in a between-subjects design. This paper details the crew's actions and responses while dealing with an in-flight airspeed failure.

Accident statistics like Baker (2001) often cite flight crew error as the primary contributor in accidents and incidents in transport category aircraft. However, the Air Line Pilots Association (2011) suggests "a well-trained and well-qualified pilot is acknowledged as the critical center point of the aircraft systems safety and an integral safety component of the entire commercial aviation system." This is generally acknowledged but cannot be verified because little or no quantitative data exists on how or how many accidents/incidents are averted by crew actions. Anecdotal evidence suggest crews handle failures on a daily basis and Aviation Safety Action Program (2003) data generally supports this assertion, even if the data is not released to the public. However without hard evidence, the contribution and means by which pilots achieve safety of flight is difficult to define. Thus, ways to improve the human ability to contribute or overcome deficiencies are ill-defined.

Method

The pilot contribution to flight safety was investigated by experimentally manipulating crew complement (single pilot and crewed conditions) during normal and increasingly challenging non-normal airline operations.

Experiment Design

To assess human performance and safety, the experiment contrasted normal two-crew operations to conditions when one of the crew was absent from the flight deck. If the condition included a temporary absence, it was designated as reduced crew operations (RCO). If the condition included a permanent absence, it was designated as single pilot operations (SPO). The experimental independent variables were crew complement and scenario. The three crew complement conditions were: Two-crew, RCO, and SPO. Two normal scenarios and six non-normal scenarios were flown. The six non-normal scenarios were grouped into three categories; Category A featured failures initially unannunciated with the autopilot available, B featured annunciated failures with autopilot available, and C featured annunciated failures with autopilot not available. Failures were triggered near top of climb (TOC) or top of descent. This paper details one Category A failure, unreliable airspeed. Etherington et al (2016) details the entire experimental matrix and details one Category C failure.

The data shown here is taken from the 18 nominal Two-crew and SPO runs, 6 nominal RCO runs (with the Captain resting), and 6 unreliable airspeed non-normal runs in each of the SPO, Two-Crew, and RCO crew conditions. For the RCO condition, the non-normal started out with one pilot flying and the other resting in the seat, isolated in sight and sound from the cockpit. Two minutes after the failure, the resting pilot returned to flying duties in the cockpit.

Participants

Thirty-six pilots (18 crews), representing five airlines, participated. Each pilot held a current Airline Transport Pilot certificate and was current in the Boeing 737-800 aircraft. Crews were paired by function (Captain and First Officer) and by employer to minimize conflicts in training, standard operating procedures, and crew resource management techniques. Crews were instructed to bring their company's paper and/or electronic charts and checklists with them to further reduce conflicts in standard operating procedures or training.

Apparatus

The research was conducted using the Boeing 737-800 simulator operated by the FAA AFS-440 at Oklahoma City, OK (See Figure 1). The simulator is Level D certified and yet fitted with experimental controls, modifications, and recording capability to support research operations. The fidelity of the simulator and the recording capability were both critical to this research effort. The scenario was an air carrier flight from Denver International to Albuquerque. Dispatch paperwork was provided and constituted the flight release. Simulated weather en-route consisted of convective activity along the mountain range to the west of Denver, and weather and visibility were designed to affect any diversion decisions. Live Air Traffic Control (ATC) and pseudo-pilots provided interactive clearance procedures, realistic pilot workload, and a level of realism to the scenario. Dispatch could be contacted on the radio.

Figure 1. FAA OKC Boeing 737-800 Simulator.

Results

The results detailed here describe the major findings of only one of the Category A failure conditions, unreliable airspeed. This failure emulated an iced-over pitot tube at the cruise altitude of Flight Level (FL) 350 which caused erroneous airspeed readings on the corresponding side. When the pitot tube became blocked, the airspeed indicator then performed like an altimeter such that increasing or decreasing altitude from FL 350 would also appear as an increase or decrease in airspeed. The failure is latent and cannot be detected until the aircraft deviates from the altitude at which the blockage occurred.

At approximately 15 minutes after TOC, the failure scenario was triggered by ATC instructing the crew/pilot(s) to climb to FL370 from FL350. As the aircraft climbed, the airspeed indicated an increase on the failed side. At a difference of 5 knots airspeed, the "IAS DISAGREE" amber warning would appear on both pilots' primary flight displays under the airspeed tape. Eventually the failed side airspeed reached an overspeed condition and the overspeed warning clacker triggered.

Failure Identification

As a Category A failure scenario, inconspicuous symptoms of failures, impending failures, or non-normal conditions were evident in the cockpit before a warning triggered; in this case, airspeed would diverge side to side and the IAS Disagree amber message illuminated before the overspeed warning clacker.

In the total of 21 non-normal runs, only 11 pilots noticed the IAS Disagree light before the overspeed warning clacker started. Only 33% of the SPO pilots recognized the failure before the clacker, while more than half (56%) of the RCO crews and two-thirds (67%) of the two pilot crews did so. One two-crew noticed the failure before the IAS DISAGEE annunciation.

Flight Path Control and Failure Handling

The average time between IAS DISAGREE and overspeed was 8-10 seconds so even for those that detected the IAS DISAGREE light, all crews/pilot(s) experienced an overspeed warning clacker. The clacker is extremely loud and distracting and continues until the overspeed condition is cleared. The clacker sounded for an average of 1.5 minutes with a range of 30 seconds to 15 minutes.

Because of some high profile accidents, this failure has been extensively trained for the past few years. Prior to referring to the appropriate checklists, nearly all crews immediately disconnected the autopilot and autothrottle from memory due to this training and to ensure that the automation was not causing the problem.

A few pilots found the clacker so distracting that they attempted to locate the circuit breaker before attempting to troubleshoot. As this clacker sounded immediately following an ATC command to climb, the majority of pilots sought to reverse the most recent action and requested a descent to the previous altitude. When the aircraft returned to that altitude the majority of the non-normal indications cleared and the aircraft behaved normally until the crew initiated the descent.

In the midst of the failures, all pilots alerted ATC to an airspeed problem but only 14 crews declared an emergency. All but two crews requested a descent or block altitude clearance from ATC with an average time of approximately 45 seconds. If this occurred during the two minute delay in an RCO configuration, the failure effects were no longer apparent when the resting pilot re-engaged.

Some pilots were erroneous in conceptualization of the flight control warning system and indicated a concern that the failed pitot would trigger the stick shaker as they descended. The stick shaker system that warns of aircraft stall conditions is based on an angle of attack sensor and not just airspeed.

Typically, the autopilot was re-engaged to the non-failed side within a minute or less.

Even for this short period of time, there were many control difficulties. At high altitude, there is a small airspeed range between stall warning and over speed which requires only a small pitch excursion to go from over speed to stall warning. Eight of the 18 crews experienced one or more stick shaker events that precede a stall. Five of the 18 crews experienced an actual overspeed because of inappropriate pitch control during the event. At least one crew received a bank angle warning. Three crews experienced both stick shaker and overspeed during recovery. Although all crew configurations had at least one event, 67% of single pilot crews experienced a stick shaker or actual overspeed. Approximately 50% of the total stick shaker events occurring during SPO and the majority of the stick shaker events during the RCO conditions occurred before the resting pilot was re-engaged. Therefore, approximately 90% of the total stick shaker events occurred when the pilot flying was essentially performing SPO. The crew resource management when the other pilot indicates “watch your airspeed” occurs long before the aircraft warnings.

Checklist Usage

Time to first correct checklist is an indicator of crew understanding of the problem. This data is shown in Figure 2. Time to complete the checklist is another indicator, as well as how closely the crew follows the checklist and if they complete additional checklists that apply to the failure.

Checklist use for this failure was complicated by the fact that the first annunciation, IAS DISAGREE, points to a checklist with the only action “Refer to the unreliable airspeed checklist”, which is the required checklist for this failure. The checklist has recently been re-designed to handle multiple failures as well as other failures. The checklist requires qualitative decisions and some of the indications disappear before completing the checklist.

The time to start the checklist was significantly faster in the two-crew condition. The time for the resting pilot to re-engage in the flight was fixed at two minutes and that is the approximate difference in times between two-crew and RCO condition. On average, SPO pilots took 50% longer to start the checklist than Two-Crew pilots did.

This time is essentially a reflection of not being able to delegate any tasks like talking to ATC, gathering weather information, talking to dispatch, and maintaining aircraft control.

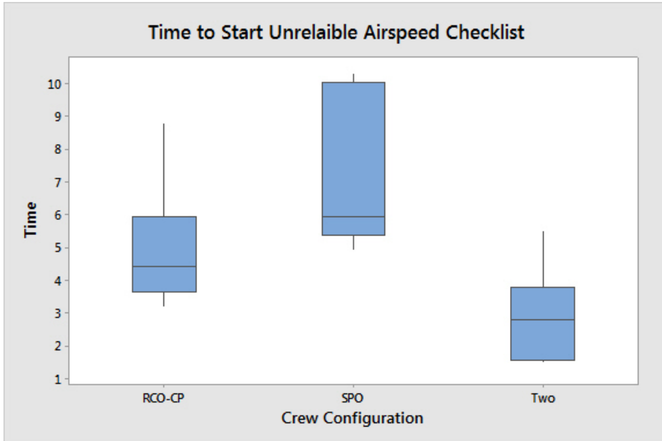

Figure 2. Time to first starting the Unreliable Airspeed Checklist.

Diversion Decision

The diversion decision is an indicator of how well the non-normal was handled and includes many factors in the decision making process. Airspeed failures, especially while in the clouds, can become critical. The diversion decision includes factors like icing potential and location of the nearest suitable airport with good weather.

The weather at Denver (DEN) and Colorado Springs (COS) was configured to be similar and relatively bad, with DEN being especially problematic. Weather at Albuquerque (ABQ) was okay and weather at Grand Junction (GJT) was good. Crews, in general, knew what the weather was like at Denver and Albuquerque but had to ask specific questions of ATC or dispatch to get other weather. The flight management system was already configured for a landing at ABQ.

Figure 3 shows diversions by airport and crew configuration. For RCO where the first officer is initially flying (RCO-CP) and doesn't make any diversion decisions until the captain is back active on the flight deck, only one of six crews diverted to other than the destination. When the Captain was flying the SPO condition (SPO-CP), they always diverted and the majority found the good weather at Grand Junction, but for First Officers flying an SPO flight (SPO-FO), only one in three diverted. For two-crew condition, half diverted to Denver. These data suggest a correlation between the perceived criticality and crew experience in the diversion decision. Not all crews considered the airspeed failure a critical problem and when flying two crew decided the risk of weather less important than the expediency of the closer airport, Denver. For single pilot Captains, the failure was critical enough that they all diverted and they felt that getting to better weather was a priority.

Workload

Overall workload was measured using the NASA Task Load Index (TLX) presented to the pilots immediately after completion of each run (Figure 4). There was an increase in workload for airspeed failure compared to the nominal runs but this difference was not found to be significant. Overall workload increased more for the first officer. Analysis of the TLX components found a significant difference in the temporal subscale ($F=3.24$, $p= 0.035$) likely due to time pressure of the first officer while completing the checklist items.

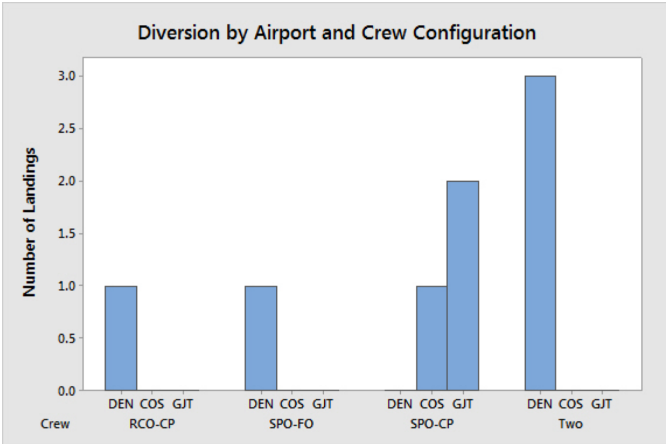

Figure 3. Diversions by Airport and Crew Configuration.

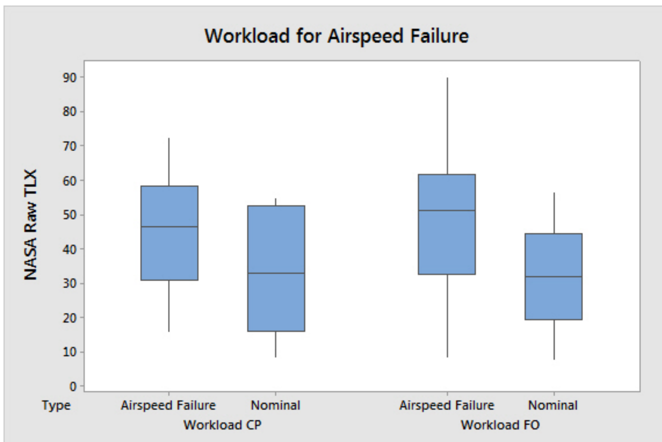

Figure 4. NASA Task Load Index Workload

Perceived Level of Safety

Post-run the crews used a Likert scale to self-assess their perceived level of safety for the airspeed failure by crew configuration and crew member, captain or first officer. The first three data columns are crew configuration as perceived by the Captain (CP) and the last three columns are crew configuration as perceived by the first officer (FO). A safety level of 1 is completely acceptable, 4 is neutral and 7 is completely unacceptable. Although a blocked pitot tube is a simple failure, RCO and SPO crew configurations rate this as unacceptable.

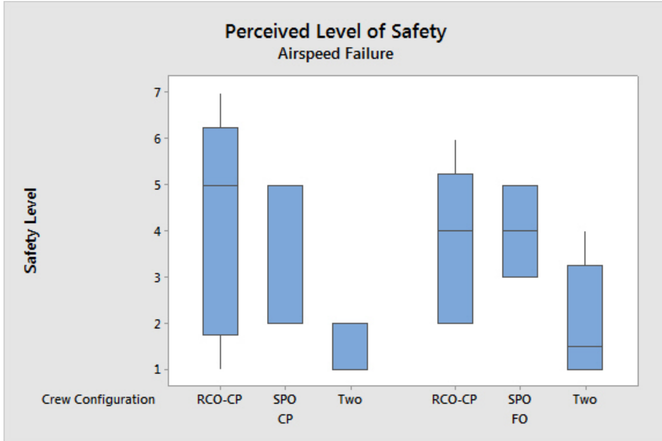

Figure 5. Perceived Level of Safety for the Airspeed Failure by Crew Configuration and Crew Member.

Conclusion

Although relatively benign, the Pitot tube failure presented some challenges that were especially problematic as the crew complement was reduced from the nominal two-crew condition. Unreliable airspeed is a well-trained event. Many crews had the initial procedures memorized; however, crews still had difficulty with aircraft control as stick shaker, overspeed, and overbank warning were common. Time to identify and begin to work the correct checklist was double for SPO compared to two-crew. Data analysis for this failure (and the other five, including nominal runs) is being used to establish quantitative baseline levels of performance and safety during nominal crew configuration. These data are being used to assess the performance and safety decrement in reduced crew and single pilot crew configurations using current-day flight deck design and certification. From this baseline, technology requirements will be identified that may inform future normal two crew operations and may eventually help enable reduced crew or possibly even commercial single pilot operations.

References

Air Line Pilots Association, "Air Line Pilots Association White Paper on Unmanned Aircraft Systems," April 2011.

Li, G., Baker, S.P., Grabowski, J.G., and Rebok, G.W., "Factors Associated with Pilot Error in Aviation Crashes," Aviation, Space, and Environmental Medicine, Vol. 72, No. 1, January 2001.

Emergency and Abnormal Situations in Aviation Symposium, LOSA Data and Emergency and Abnormal Situations http://human-factors.arc.nasa.gov/flightcognition/download/EAS_Symposium_Presentations/LOSA.pdf , June 10-11, 2003

Etherington, T.J., Kramer, L.J., Kennedy, K.D., Bailey, R.E., and Stephens, C.L., *Quantifying Pilot Contribution To Flight Safety For Normal And Non-Normal Airline Operations*. Paper presented at AIAA Digital Avionics Systems Conference.

Acknowledgements

This work was sponsored by NASA's Airspace Operations and Safety Program, Safe Autonomous Systems Operations project, led by Parimal Kopardekar. It was conducted under an interagency agreement between NASA and the FAA. This work could not have been possible without the hard work, diligence, assistance and support of Mr. Harry Hodges (AFS-440), Mr. Doug Rodzon (AFS-440), the entire AFS-440 Flight Simulation Branch, and Mr. Bruce DeCleene (AFS-400). The technical assistance of Dr. Robert "Buck" Joslin and Mr. Bryan Burks, Alaska Airlines, was greatly appreciated.