2013

Extending Mission Operations Safty Audits (MOSA) Research to an Indian Sub Continent Island Airline

Sue Burdekin

EXTENDING MISSION OPERATIONS SAFTY AUDITS (MOSA) RESEARCH TO AN INDIAN SUB CONTINENT ISLAND AIRLINE

Sue Burdekin
University of New South Wales
Australian Defence Force Academy
Canberra, Australia

The aim of the Mission Operations Safety Audit (MOSA) research is to validate behavioural self-reported data from professional pilots, so that management can have confidence in this safety-critical debriefing information, and feed it back into the training continuum. In doing so, a safety loop can be established in a cost effective, operationally specific and timely program of data collection. The first study was conducted in a military F/A-18 Hornet simulator. Pilots were asked to self-report on their own operational performance across a predetermined selection of behavioural categories designed in conjunction with subject matter experts. To further test the MOSA methodology, this time in-flight, a second study was carried out with the cooperation of a civil airline in Europe. Both the military and the civil airline studies found that professional pilots were able to effectively self-report on their own performance. However, the multi-crewed European airline pilots' results revealed that first officers were more critical of their own performance. In order to determine whether national or organisational culture influenced these results, the MOSA methodology was recently tested in a regional airline operating Dash 8 aircraft between island destinations in the Indian Ocean. The results indicated that neither national culture nor aircraft operating culture appear to influence the accuracy of pilot self reports. However, the self reports from first officers do appear to be linked to their seniority and experience in that role.

The use of trained observers to rate professional pilots' non-technical skills performance during normal flight operations has become an accepted individual assessment and system evaluation method in civil aviation. There are two well know approaches utilizing in-flight observers: NOTECHS (non-technical skills) assessment and Line Operations Safety Audits (LOSA). NOTECHS was developed from a Joint Aviation Authorities (JAA) sponsored research project in the UK and Europe to provide a tool to assess pilots' CRM skills in the cockpit. These skills were defined as cognitive and social skills not directly concerned with flight control, systems management and standard operating procedures (SOPs). As an assessment tool supported by the regulator, failure to meet the predetermined standard can result in license suspension and remedial training (Flin, O'Connor & Crichton, 2008). On the other hand, LOSA was developed by The University of Texas Research Project and airline partners in North America initially to audit pilots CRM performance and then expanded to identify threats to the conduct of the flight and how errors are managed. It provides a report of an airlines strengths and weaknesses determined by non jeopardy observations of line crew performance on the flight deck (Klinect, Murray, Merritt & Helmreich, 2003). In-flight observations for both NOTECHS and LOSA are conducted under 'normal' conditions. That is, during regular, revenue raising operations.

However, the definition of 'normal operations' in a military context translates into a completely different paradigm. Normal military operations can involve high speed, rapid manoeuvre, or terrain following radar flight in a range of multi-crewed and single pilot aircraft. Many of these aircraft types can only accommodate the flight crew therefore, depending upon the particular platform, in-flight observation may not be physically possible. Because of this, the Royal Australian Air Force (RAAF) supported research into testing the accuracy of pilot behavioural self-reports, known as Mission Operations Safety Audits (MOSA). The objective was to establish a valid means of collecting targeted safety critical debriefing information which could be acted upon in a timely manner. Additionally, it was intended that MOSA could provide a cost effective evaluation of aircrew performance which would highlight issues that could feedback into the training continuum, thereby establishing a behavioural safety loop in line with the defence aviation safety management system.

The first study was conducted using a single pilot F/A-18 Hornet flight simulator. Participants were asked to rate their own behaviour on a 5 point Likert scale after 'flying' a medium or high workload mission, according to pre-determined behavioural categories, which had been designed with input from military subject matter experts (SMEs). These categories included: workload management, communication, evaluation of plans, situational awareness, monitoring and cross checking, inquiry, assertiveness and automation management. The results across both conditions of workload were highly correlated with the ratings of expert observers (Burdekin, 2012).

Following the military simulator research, Airbus supported a second MOSA study to test self-report methodology on the flight deck during normal operations in a European multi-crewed civil airline environment. SMEs from Airbus and the 'low cost model' airline Easyjet joined the researcher to further develop the methodology to incorporate a crossed design which recorded self reports from each pilot, along with the captains ratings for the first officer (F/O) and the F/Os ratings for the captain. These ratings were then compared to the ratings from an observer during the same flight sector. Sixty flight sectors were observed and the ratings by the observers were found to be significantly correlated with the self-report ratings from both the captain and the F/O. However, deeper analysis revealed that F/Os were more highly critical of, not only their own performance, but also the performance of their captain (Burdekin, 2012). It was noted that 83% of the F/Os who volunteered for the study were relatively new recruits of less than one to three years with the company and an average of 1950 hours total flight time. Given the current debate concerning the minimum flight hours required by pilots to fly for an airline (US Congress, 2010; Rural Affairs and Transport References Committee, 2011), the MOSA research has highlighted that junior F/Os expect higher standards of performance from themselves and their senior crew members. In order to determine if the level of experience and/or perhaps a cultural issue influenced the European civil airline study results a further test of the MOSA methodology was conducted.

The present study was carried out in an Indian sub continent airline operating Dash 8 turbo-prop aircraft to island holiday destinations in the Maldives. The aim of the study was to determine if the MOSA methodology would be suitable for collecting accurate aircrew self-reported information in a different cultural and operational environment.

Method

Forty one flight sectors were observed from the jump seat during normal revenue raising operations by an experienced subject matter expert (SME) who is also a commercial pilot. All crew members were volunteers and their participation was anonymous. As such, it was not possible to determine if a pilot was observed more than once. The design of the study was developed by the researcher with input from the safety manager, senior management and check and training pilots. A crossed experimental design was developed where each volunteer crew member rated his/her own performance, the performance of the other crew member and how, collectively, they performed as a crew during each sector, across a predetermined set of behavioural categories. The observer rated the performance of the captain, the first officer and also how they performed as a crew during each sector using the same predetermined behavioural categories. The protocols included eight categories of behaviour that were assessed to be a representation of the non-technical skills that the airline was keen to evaluate. Those categories were: briefing; contingency management; monitor/cross-check; workload management; situational awareness; automation management; communication; and problem solving/decision making. Each behavioural category was given comprehensive descriptors, illustrated in figure 1.

Figure 1.
Example of behavioural category and descriptors

BEHAVIOURAL CATEGORY	DESCRIPTOR	GRADING/WORD PICTURE *(1. Poor; 2. Marginal; 3. Adequate; 4. Very Good; 5. Excellent)*
Automation management	**Interaction between the operator and automated system**	**1. Incorrect crew interaction and management of aircraft automatic systems. Clear errors of competency in automation set-up, mode selection and utilization.** **2. Basic interaction with aircraft automatic systems. Appropriate mode selection and utilization barely adequate to maintain safe flight profiles.** **3. Level of automation interaction adequate to maintain prescribed SOP profiles. Mode utilization satisfactory and procedurally correct. Recovery technique from anomalies reflects limited system awareness.** **4. Automation interaction to a good standard. Effective and timely management of automatic modes. Flight path SOP profiles maintained to a proficient standard. Clear understanding of aircraft automation systems reflected in sound anomaly management.** **5. Automation management to a high standard. Clear anticipation and use of appropriate modes. All anomalies managed to a highly proficient standard reflecting a deep understanding of the automation system.**

The protocol form allowed for crew members and the observer to comment on any issue that affected the safety of that sector in the form of air traffic management, ground support, aerodrome operations, cabin crew interaction, and any other issues.

Additionally, a questionnaire was conducted during the flight that asked crews to identify any wider safety issues of concern within the airline. For example: "Can you list the top 5 safety issues currently affecting the company"; "Can you predict what the next incident/accident will be"; "What do you think would be the best way to prevent this from occurring"; "Can you nominate one CRM strategy that you have adopted that has changed the way you approach your flying"? The volunteer crew members were assured of individual confidentiality and the researcher/observer remained the 'gate keeper' of all data.

Results

The data across all categories of behaviour were collapsed and subjected to a test of correlation between the ratings of the independent observer, the captain, and the first officer (refer to Table 1). In addition to rating the performance of themselves and each other, the crew members were asked to rate how they performed together as a crew. These ratings were also correlated with the ratings of the crew's performance from the independent observer (refer to Table 2). All results were found to be statistically significant except for the crew performance ratings from the first officers. Details of crew experience can be found in Table 3.

The answers from the in-flight questionnaire were compared with routine safety data gathered from flight data monitoring, voluntary and mandatory safety reports, air safety occurrences, management safety committee meetings, and other workforce/organisational evaluation data. Analysis of this sensitive safety critical information indicated that the results were valid.

Table 1.
Results across all categories of behaviour

Rater	Mean	sd	N	r
OBS	4.12	.54	224	.669 **
CAPT	4.15	.62		
OBS	4.13	.53	224	.188 **
F/O	4.13	.58		
CAPT	4.14	.68	224	.232 **
F/O	4.13	.58		
F/O	4.31	.62	224	.112 *
CAPT	4.15	.62		

** Significant .01
* Significant .05

Table 2.
Crew performance results

Rater	Mean	sd	N	r
OBS/CREW	4.25	.52	224	.620 **
CAPT/CREW	4.26	.58		
OBS/CREW	4.25	.52	224	.186 **
FO/CREW	4.29	.59		
CAPT/CREW	4.26	.58	224	.099
FO/CREW	4.29	.59		

** Significant .01
*   Significant .05

Table 3.
Average age and experience of Crew

	Mean Age	**Average Total Flight Hours**
Captain	46	14,200
First Officer	31	4,300

Conclusion

The results from this study show that both captains and first officers were able to accurately report on their own performance, compared to the ratings from each other and an independent observer, across a range of categories of behaviour which reflected their non-technical skills. Although the observer/captain ratings were more highly correlated in both this and the European studies, in the present study, the anomaly of first officers being more critical of their own performance than the ratings issued to them by their captains and the observer was not repeated. The difference in the self-assessment of these first officers might be explained by their level of experience and the length of time that they had been employed as first officers. The majority of the F/O volunteers in this study were senior first officers who were awaiting a captaincy slot. Whereas, the first officers in the European study were relatively junior in terms total flying hours and length of time with the company.

This finding suggests that the level of pilot experience influences first officers’ ability to accurately identify their own individual performance, given a comprehensive non-technical skills behavioural scale. Therefore, to be required by the company to regularly reflect on their performance on the flight deck might help to facilitate first officer professional development.

One reason for extending the MOSA research to the Indian Subcontinent was to test the methodology in that culture. However, National culture does not appear to have influenced the results as the observed behaviours were very similar to those from the European MOSA study. The present study was conducted in a small airline operating turbo prop aircraft, flying short sectors between island landing strips, although, this type of operational culture also does not appear to have impacted the results.

This study lends support to the body of MOSA empirical research which concludes that non-technical skills self-assessment information collected from professional pilots across a predetermined range of behavioural categories is an accurate indication of performance on the flight deck. Therefore, it is suggested that aggregated and structured aircrew self-reported performance and safety information can be utilized with confidence by management to highlight developing safety issues, and indicate areas of deficiency, as well as identifying the behaviours that work well. Additionally, it has been identified that MOSA methodology can be used in conjunction with other information gathering and evaluation tools to contribute to the on-going safety feedback loop of an organisation's safety management system.

Acknowledgements

The author would like to acknowledge assistance from the following organizations in the conduct of the MOSA studies: the Royal Australian Air Force, Airbus Industries, Easyjet, and Maldivian Airlines.

References

Burdekin, S. (2012) Evaluating Professional Aircrew Safety Related Behaviour In-flight. *In proceedings of the 4th International Conference on Applied Human Factors and Ergonomics* (pp 450–461). San Francisco, USA.

Flin, R., O'Connor, P. & Crichton, M. (2008) *Safety at the Sharp End: Training non technical skills.* Ashgate, Aldershot.

Klinect, J.R., Murray, P., Merritt, A. & Helmreich, R. (2003) Line Operations Safety Audit (LOSA): Definition and operating characteristics. *In Proceedings of the 12th International Symposium on Aviation Psychology* (pp.663-669). Dayton, OH. Ohio State University.

Rural Affairs and Transport References Committee, (2011) *Pilot Training and Airline Safety*. Senate Printing Unit Commonwealth of Australia. Canberra.

United States Congress. (2010) *Airline Safety and Federal Aviation Administration Extension Act of 2010*. United States Government Printing Office. Washington DC