

Title: The Development of the Maintenance Operations Safety Survey: challenges in transferring a predictive safety tool from flight operations to aircraft maintenance.

Authors:

Marie Langer, Cranfield University (ST6173)

Marie Langer is a final year research student pursuing a PhD degree at the Cranfield Safety and Accident Investigation Centre (CSAIC) at Cranfield University. Besides her PhD research Marie regularly delivers lectures on safety data analysis and performance assessment in aircraft maintenance, and has been working on consultancy projects for the International Air Transport Association (IATA). Prior to joining the PhD programme she spent a year working for IATA in Montreal initially as an intern and then as an assistant manager in the safety department. Marie also holds a Master's degree in Air Transport Management from Cranfield University.

Prof. Graham Braithwaite, Cranfield University (MO3644)

Graham Braithwaite is Head of the Department of Air Transport at Cranfield University, UK where he also holds the Chair in Safety and Accident Investigation. Having completed a PhD in aviation safety management at Loughborough University, Graham worked for 6 years at the University of New South Wales in Sydney, Australia before joining Cranfield in 2003. His research and teaching covers human factors, accident investigation and the role of culture in safety. In 2011, Graham led Cranfield's successful bid for a Queen's Anniversary Prize for its contribution to aviation safety through teaching and research in accident investigation. He was presented the award by Her Majesty The Queen and HRH The Duke of Edinburgh at Buckingham Palace.

Introduction

Predicting human behaviour and managing human error is arguably the greatest challenge facing the aviation industry today (Shappell and Wiegmann, 2009). In order to gain a better understanding of human behaviour and overall organizational safety performance, the industry is moving towards monitoring of normal operations (Helmreich et al., 2003). One of the key advantages is the learning opportunity without the negative consequences and associated costs of an incident or accident. Predictive tools such as the Line Operations Safety Audit (LOSA) provide objective information of routine operational performance, complement existing safety data collection programs and are endorsed by the International Civil Aviation Organization (ICAO).

However, while aviation safety is heavily dependent on the management of human error in all parts of the complex system, predictive tools are still largely limited to flight operations. Despite its safety critical role, aircraft maintenance in particular arguably remains over-reliant on safety information from traditional sources such as accident and incident investigations or quality audits. Perhaps one of the reasons is the very nature of the work where engineers seldom receive feedback on their actions, a key difference between flight and maintenance operations. While pilots can usually see the results of their actions almost immediately, the consequences of a maintenance error may be uncovered only when it leads to an incident or an accident. Aircraft maintenance is also associated with extensive human

involvement which, when coupled with a complex system, makes it highly susceptible to error. Yet, the knowledge of maintenance error often does not extend to the organizational context in which it occurs (Hobbs, 2008). A new approach is needed to gain information about the nature of maintenance errors, frequencies and successful recoveries present in the maintenance environment on day-to-day basis. Rather than focusing on negative events, a LOSA-like tool designed for the maintenance environment can enrich the current knowledge of maintenance error by uncovering organizational roots and human responses to everyday threats and errors.

This paper is based on a PhD research project testing the application of LOSA principles in line maintenance and focuses on the lessons learned during the development and practical application of the tool. In cooperation with Thomas Cook, a major UK airline and maintenance organization, the existing LOSA concept was adapted to suit the maintenance environment and introduced in real operations. Through structured observations conducted by trained observers during routine operations, narrative data were collected describing how engineers detect, manage or mismanage threats and errors. The tool captured the complex maintenance environment, identified the causes of human failure, human responses to failures and successful recoveries.

A quick look at LOSA

LOSA is a proven concept which has been successfully running in flight operations since late 1990's. It is defined by ten operating characteristics which address pilot trust, observer reactivity, data reliability and validity concerns:

- Jump seat observations of routine flight operations
- Anonymous, confidential and non-punitive data collection
- Voluntary flight crew participation
- Joint management/pilot union sponsorship
- Systematic/safety targeted observation instrument
- Trusted and trained observers
- Trusted data collection repository
- Data verification roundtables
- Data derived targets for enhancements
- Results feedback to line pilots

Since LOSA assures non-punitive action it has the ability to capture flight crew performance much closer to operational reality than ever before. Through structured observations during routine operations trained observers record how flight crew detect, manage or indeed mismanage threats and errors. The collected information is in form of a written narrative where the observers are encouraged to write a story of the flight describing a sequence of events for each flight phase. Although the process may seem like a traditional line check, in this case the observer is not an examiner. Anonymity and confidentiality are also guaranteed so flight crew cannot be identified and punished for any errors unveiled by LOSA. Importantly, before any observation can take place the flight crew are asked for a consent. Their

participation is voluntary so if anyone declines there are no questions asked and the observer selects another flight. Thus, trust is the core of a successful LOSA that allows observers to see normal flight crew performance and collect safety data.

At the heart of LOSA is a theoretical model, Threat and Error Management (TEM), which reflects the complex aviation environment. The model takes into account not only active failures (errors) but also latent conditions (threats), which are always present within the system. It is based on a principle that flight crews have to manage threats and errors on daily basis to maintain safe operations. LOSA can highlight not only weaknesses but also strengths of the system by focusing on successful human performance (Klinect et al., 2003; Klinect, 2005).

Applying LOSA principles in the maintenance environment

LOSA has been recently successfully transferred to air traffic control, despatch and even the medical industry resulting in similar benefits to those achieved in flight operations (see Patterson, 2007; FSF, 2005; Helmreich and Sexton, 2004). However, applying the same LOSA principles in aircraft maintenance is a complex task. Not only is the nature of the work and the physical environment different, but also the distinct working culture had to be considered.

The first and perhaps the most important step was to move beyond the blame culture, overcome the existing engineer skepticism in new initiatives and gain their trust - the cornerstone of the LOSA methodology. The attitude of many engineers towards audits meant that the first obstacle was the actual title of the project. Line operations safety audit was clearly not acceptable term for this cultural setting: Not only because of the misconception that it was only relevant to line maintenance, but particularly because of the associated implications of audits (focused on individuals and punitive). As a result, the project was introduced as the Maintenance Operations Safety Survey (MOSS). It covers all aspects of maintenance operations and the word "survey" is associated with the fundamental attributes of this tool: anonymity, confidentiality and non-punitive action.

The physical working environment coupled with the nature of the work posed a great challenge particularly in terms of identifying a suitable unit of observation. While pilots work in a highly ordered, relatively task-limited, confined space, engineers perform a much greater variety of tasks in diverse settings. Aircraft maintenance has different forms ranging from line maintenance associated with an open airport environment, changing weather conditions and high operational pressure, to hangar maintenance defined by a large number of complex tasks, often carried out at heights or in confined spaces. Considering all the complexities it was vital to identify a common unit or event which could be observed from the beginning to the end. In flight operations this common unit is a flight, with a clear starting and finishing point, and always following the same flight phases. This breakdown allows comparison of information from each flight across operators irrelevant of other attributes such as the departure or arrival points or the length of the flight. Hence, line maintenance seemed to be the appropriate environment to start with especially as all activities including defect rectification, inspections and routine checks are incorporated into different types of checks (e.g. transit, daily, weekly).

For the equivalent to the flight phases used in LOSA, Liston (2005) proposed to break an observation into three stages: job set-up, job procedure and job close-up. He noted that Airbus and Boeing split all maintenance tasks into those three stages and the breakdown is used throughout all their published documentation. Importantly, this natural decomposition of tasks is familiar to the maintenance personnel and provides a clear structure for data collection and subsequent analysis. As a result, an observation form was designed incorporating the three stages. As in flight operations the observation form would contain a written narrative of events observed during the course of each stage (phase) and would be filled only after the observation ended.

In terms of executing observations, the observer on the flight deck is like a “fly on the wall”: sits on the jump seat and will not interrupt the flight crew unless safety is compromised. Because of the close proximity it is possible to observe all actions and communications performed by the flight crew. In the open line maintenance environment, on the other hand, the observer becomes perhaps more visible to the observed engineer which may feel more intrusive than in flight operations. Engineers move around the aircraft, collect spares from stores and frequently return to the office to access documentation as part of a check. So rather than observing an engineer performing maintenance tasks, the observer role is to follow and observe which can make it awkward and impact on the engineer’s behaviour. Although it is possible to keep a distance when observing a walk-around, the observer has to get much closer when accessing the cabin and the flight deck. Also, practically speaking, the observer can only follow one person at a time taking into account the size of the aircraft. Hence, if two or three engineers are allocated to carry out a check the observer has to select one person for an observation. Although this represents a limitation, the practical tests noticed that comprehensive contextual information can still be captured. Since one engineer is responsible for certification of the check, all involved engineers would communicate about their individual tasks or issues which would be observed and reflected in the narrative.

Dialogue with line personnel throughout the entire process was vital for MOSS success. Engineers need to understand not only the purpose and the process, but importantly how the collected data are going to be stored and used. While in flight operations, LOSA is an established process where flight crews can contact other operators to gain information about the tool, MOSS in maintenance is yet to prove itself. Because of inherent distrust and the fact that MOSS is in its infancy, the greatest concern was over a possible misuse of information, identification of individuals and punitive action for errors committed. Even though it was communicated at the beginning of the process and addressed by a letter from the management to all engineers, such concerns required frequent reassurance of anonymity, confidentiality, non-punitive action and the independence of Cranfield University as the data repository. It is also important to note that like LOSA, MOSS application is a lengthy process. To ensure successful results, LOSA in flight operations may take up to eight months from the initial planning stage to the final report of findings. Approximately 70% of the time is attributed to analysis which is not visible to the line personnel. Since MOSS was run for the first time, essentially testing the concept, the elapsed time from the planning stage to report delivery was twelve months. The perceived lack of activity could trigger doubts about the tool and damage the established trust. Recognizing the possible issue early in the

process, it was decided to provide regular updates of progress which coincided with company organized forums for line engineers and in turn, received positive feedback.

TEM in maintenance operations

Threat and error management provides the framework not only for the data collection but importantly for the data analysis (Klinec, 2005). Firstly, for the purpose of MOSS the definitions of threats, errors and undesired aircraft states needed slight alteration. As noted earlier, whilst pilots can usually see the results of their actions almost immediately, the consequences of a maintenance error may not be immediately apparent or observable even when safety margins are reduced. Therefore in terms of defining errors and undesired states the potential to reduce safety margins would be taken into account as reflected in the TEM definitions adapted for MOSS and outlined in table 1.

TEM definitions	LOSA Flight operations (Klinec 2005; Merritt and Klinec, 2006)	MOSS Aircraft maintenance
Threats	Events or errors that occur outside the influence of the flight crew (i.e., not caused by the crew); increase the operational complexity of a flight; and require crew attention and management if safety margins are to be maintained.	Observable events or errors which occur outside the influence of the engineer; increase the operational complexity of a task; and require engineer’s attention and management if safety margins are to be maintained (e.g. ground equipment obstructions, despatcher interruptions to engineer duties, operational pressure, aircraft defect).
Errors	Flight crew actions or inactions that lead to a deviation from crew or organizational intentions or expectations; reduce safety margins; and increase the probability of adverse events on the ground or during flight.	Observable engineer actions or inactions that lead to a deviation from engineer intentions or organisational expectations; <u>have the potential to reduce safety margins</u> ; and increase the probability of adverse events on the ground or during flight (e.g. check performed from memory, failure to make an entry into technical log, servicing error).
Undesired aircraft states	Flight crew error-induced aircraft deviations or incorrect configurations associated with a clear decline in safety margins.	Undesired states ¹ – observable engineer error-induced states or situations which reduce or <u>have the potential to reduce safety margins</u> (e.g. APU left running unattended, areas of aircraft not checked for damage).

¹ In MOSS “undesired aircraft states” will be referred to as “undesired states”

In order to derive clear and quantifiable results from the collected narratives, which are understandable to the management, threat and error categories reflecting maintenance operations had to be developed. Essentially, the top level categories (threats, errors and undesired states) come from the conceptual framework and define mutually exclusive categories. The TEM model further takes into account responses and outcomes of threats, errors and undesired states. Threats are focusing on information about the context or environmental settings, while the responses and outcomes focus on individual strategies to achieve or manage situations.

The specific (lower level) TEM categories are based on the collected narrative data. A list of TEM categories emerged progressively during data collection and it is expected to expand with further MOSS applications. Using principles from content analysis (see Krippendorff, 2004; Smith, 2000) and applying the adapted TEM definitions to each narrative specific types of threats, errors, undesired states, their responses and outcomes were identified. Developing categorization from the actual data rather than using pre-defined categories means that the results truly reflect the operations at the time of data collection. While the current TEM list will serve as a basis for the next MOSS application the flexibility of the content analysis method will help in recognizing new types of threats, errors and undesired states when MOSS is conducted again in the same organization or adopted by a different operator.

MOSS data collection

Once the methodology was adapted and tested in the operational environment, the full implementation of the MOSS process was agreed with the operator. Considering the project time constraints, available resources and analytical capability, it was decided to limit MOSS application to four line stations. Collecting information at four different line stations is a means of building a verification process into the data collection and increasing data reliability. As a first step, a letter explaining the MOSS project and the management commitment was sent out to all personnel at the line stations. Then, seven volunteers, all line engineers including a union representative, underwent a two-day training to become observers.

Overall, a sample of 56 randomly selected observations was collected. While the original plan was to complete the data collection within just 4 weeks, unforeseen operational issues meant that it actually took 15 weeks. Only by testing the entire process, did it become apparent that data collection is not possible in the same way as in flight operations. Although it requires extensive planning, an airline usually has the capability to de-roster flight crew (observers) and allocate their duty time for LOSA data collection for the required period of time. This is not feasible in aircraft maintenance. Taking observers off duty for data collection was heavily influenced by shift patterns, number of people on shift, workload, license cover and even holiday arrangements. Rather than having a planned schedule for data collection like in flight operations observations had to take place when it was operationally possible. As a result, it is estimated that approximately 50% of observations required the observers to conduct observations on an over-time basis with associated additional costs. Essentially, the key lesson was that the data collection must be a compromise between time and resources availability. To collect data in a shorter period of time, more resources in terms of number of observers and over-time pay would be necessary.

Discussion of MOSS findings

Arguably the first and the most important indication of MOSS success is the degree of acceptance from the line engineers. In LOSA, this is measured by the number of flight crew refusals to be observed. A high denial rate could indicate a lack of organizational trust in the methodology and compromise the quality of results. During the MOSS data collection only two refusals were noted and both occurred at the beginning of the process. This suggests that MOSS was well embraced and achieved a high level of confidence with observers and participants.

As the methodology generates a large amount of data, the TEM model provides the direction for analyses. Drawing upon many years of LOSA experience, the MOSS data analytical process began with TEM based performance indicators in the form of calculated frequencies of threats, errors and undesired states the results provide a general overview of systemic and engineer performance. These initial results point towards organizational strengths and areas of weaknesses which need further investigation.

Highlighting the operational complexity, the preliminary MOSS results indicate that threats were present in all collected observations at an average of 7.8 threats per observation. A large proportion, 68% of threats were organizational in nature and mainly ground/ramp related, followed by 27% of aircraft related threats. While the majority of threats were detected and effectively managed 16% of threats were linked to engineer error. This means that these threats were either not detected or were not managed effectively.

Errors were noted in 86% of observations, at an average of 2.5 errors per observation. The vast majority of errors were procedural, nearly 80%, and mainly associated with non-compliance. This corresponds with findings from internal quality audits and results from the UK Confidential Human Factors Incident Reporting Programme (CHIRP). For illustration, in flight operations LOSA archive also noted non-compliance errors as the most common error type frequently associated with non-standard checklist use. Interestingly, many years of LOSA experience suggest that increased non-compliance reflects the safety culture and leads to decreasing safety margins as the operator usually experiences higher rates in mismanaged threats and errors, undesired aircraft states and mismanaged undesired aircraft states (Merritt, 2007). This link, however, is yet to be established with MOSS as the results are based on data collection in one maintenance organization to date.

Ineffective error management can potentially lead to additional errors or undesired states, events with a high likelihood of decreasing safety margins. Even though the majority of observed errors were inconsequential, a proportion of errors resulted in undesired states and were noted in 34% of observations. The 36 observed undesired states were mainly associated with aircraft areas not checked for damage at any point during the check, APU left running unattended or failure to complete all checklist items before certification. Importantly, it was noted that 20 of these events originated in ground/ramp related threats disrupting walk-arounds and through inadequate threat and error management, resulted in undesired states. Even though these specific undesired states did not contribute to any incident or accident at the time these are areas of concern that need to be addressed.

Inadequate walk-arounds failing to notice missing or not correctly closed panels, cowling or door latches have previously led to serious incidents, damage and considerable costs. For illustration, in one particular incident investigated by the UK AAIB, a Boeing 777 suffered substantial damage to cabin windows and minor damage to the fuselage and fin after a large access door detached from the aircraft shortly after take-off. A deviation from standard operating procedure during a routine maintenance check in the hangar resulted in inadequate fastening of the Air Driven Unit (ADU) door catches. The aircraft was subject to multiple walk-around inspections by hangar and line personnel before and after the handover to flight operations. The investigation revealed that the open catches were not detected during 11 subsequent walk-around inspections. MOSS results, therefore, can identify specific threats contributing to the failed systemic defence (e.g. walk-around inspection) and opportunities for errors with the potential to result in similar incidents so these can be addressed and reoccurrence prevented.

While threat, error and undesired state frequencies provide valuable information, examining TEM management strategies is perhaps more diagnostic. The experience from LOSA reveals that flight crew typically notice first the undesired state rather than the error or errors which contributed to the state. Arguably, this is linked to the immediacy of feedback to flight crew actions as the error might not be as noticeable as the aircraft state. MOSS results in terms of detection of errors and undesired states were different. Perhaps because of the lack of feedback, errors and undesired states in particular were frequently undetected.

Differences in threat and error management strategies between day and night shifts were also highlighted. Interestingly, due to increased complexities such as operational pressure, threats and errors were managed more effectively during night shifts than during day shifts. This finding is quite contrary to current industry views based on results from accident and incident investigations, that night shifts provide much higher opportunity for maintenance error. However, it has to be noted that MOSS data were collected at the beginning of the winter season. Given that Thomas Cook is a seasonal operator, the line maintenance workload during day and night shifts reduces in the winter season.

Rather than focusing on detailed results specific to the operator it is essential to examine if MOSS delivered comparable benefits to LOSA in flight operations. Referring to the three-stage breakdown of an observation in maintenance, an equivalent to flight phases, the results suggest that similar to flight operations, each stage is linked with specific types of threats and errors. For instance, procedural errors frequently associated with the failure to use checklists were noted during the set-up stage.

Importantly, the TEM framework allows a comprehensive link between threats, errors and undesired states to be established. It is well known that given the nature of the job engineers have to frequently deal with obstructions and interruptions. However, the frequency or implications are not well understood. As illustrated earlier, the information would not be accumulated unless it transpires in a serious incident or accident that is subject to investigation. The MOSS project revealed that although engineers generally deal well with interruptions there are still instances when additional threats are present and interruptions then lead to errors and potential safety issues. For example, MOSS results identified that ground/ramp related obstructions frequently stop engineers from inspecting parts of the aircraft. This represents an early link in a chain of events. As complexity increases the engineer fails to

manage this threat as other threats coupled with operational pressure to depart on time are present. Frequently, by the end of the check the engineer commits an error, most likely unintentionally, and fails to return to inspect the areas even when all cargo doors are closed and the aircraft is clear of all ground equipment. The result is an undesired state as the engineer certifies incomplete check and aircraft departs even though the areas most prone to damage were not examined.

Overall, MOSS has the capability to highlight not only the weaknesses but also the strengths of the system in terms of successful threat and error recoveries. Similar to LOSA results in flight operations, the majority of threats and errors identified by MOSS were inconsequential. This not only shows the error tolerance of the aviation system but also the merit of the LOSA methodology. Given the absence of negative outcomes these events would not be picked up by any other safety tool limiting the learning opportunity.

Conclusion

Despite the challenges associated with a distinct working culture, physical environment and nature of work it was possible to adapt and implement the LOSA methodology to line maintenance operations. Factors such as appropriate introduction and effective communication throughout contributed to the MOSS success. Operational issues in terms of availability of observers and scheduling observations were discovered during the practical implementation of the process to provide important lessons for future MOSS applications. In summary, the findings of this research highlighted that the quality of MOSS results is dependent not only on the methodology but also on proper execution taking into account the characteristics of maintenance operations.

Acknowledgement

This PhD research would not have been possible without the support and close cooperation of the industry partner, Thomas Cook Aircraft Engineering UK, part of Thomas Cook Group PLC. Personal thanks go to everyone who contributed to the MOSS project and made it a success.

References

- FSF (2005), "Line operations safety audit (LOSA) provides data on threats and errors", *Flight safety digest*, Vol. 24, No. 2, pp. 1-11.
- Helmreich, R. L., Klinect, J. R. and Wilhelm, J. A. (2003), "Managing threat and error: data from line operations", in Edkins, G. and Pfister, P. (editors), *Innovation and consolidation in aviation*, Ashgate, Aldershot, pp. 15-22.
- Helmreich, R.L. (2004), "Managing threat and error to increase safety in medicine", in Dietrich R. and Jochum, K. (editors), *Teaming up: components of safety under high risk*, Ashgate, Aldershot, pp.117-132.
- Hobbs, A. (2008), *An overview of human factors in aviation maintenance*, ATSB, Australia.
- Klinect, J. R. (2005), *Line operations safety audit: a cockpit observation methodology for monitoring commercial airline safety performance*, (PhD dissertation), University of Texas, Austin.

Klinect, J. R., Murray, P., Merrit, A. and Heilmreich, R. (2003), "Line operation safety audits (LOSA): definition and operating characteristics", in proceedings of the 12th *International symposium on aviation psychology*, Dayton, Ohio, pp. 663-668.

Krippendorff, K. (2004), *Content analysis: an introduction to its methodology*, (2nd eds), Sage, California.

Liston, P. (2005), *Human factors competence in aircraft maintenance*, (PhD thesis), University of Dublin, Dublin.

Merritt, A. (2007), "The LOSA archive: the data and how it can be used", presented at the 2nd *ICAO global symposium on TEM and NOSS in ATC*, 7-8th February 2007, Washington DC, USA.

Merritt, A. and Klinect, J. R. (2006), "Defensive flying for pilots: an introduction to threat and error management", available at:

<http://homepage.psy.utexas.edu/homepage/group/HelmreichLAB/Publications/pubfiles/TE M.Paper.12.6.06.pdf>.

Patterson (2007), "Finavia NOSS trial report", available at:

http://www.eurocontrol.int/humanfactors/public/standard_page/NOSS.html

Shappell, S. and Wiegmann, D. (2009), "A methodology for assessing safety programs targeting human error in aviation", *International journal of aviation psychology*, Volume 19, Issue 3, pp. 252-269.

Smith, C. P. (2000), "Content analysis and narrative analysis", in Reis, T. and Judd, C. M. (editors), *Handbook of research methods in social and personality psychology*, Cambridge University press, Cambridge, pp. 313-335.