UNRELIABLE AIRSPEED – MAKING VALID ASSUMPTIONS ON FLIGHT CREW PERFORMANCE

L. Fucke*, R. Nicholson, R. Mumaw****
***Boeing R&T Europe, Madrid, Spain **Boeing Commercial Airplanes, Seattle, WA**
Lars.Fucke@boeing.com; Roger.Nicholson@boeing.com; Randall.L.Mumaw@boeing.com

Keywords: *human performance, flight deck, safety*

Abstract

As part of airplane certification the type certificate holder, i.e. manufacturer is required to monitor in-service events to ensure that events with a catastrophic outcome occur with extremely improbable likelihood. The reliability of system components can typically be obtained based on failure statistics. In situations where flight crew intervention is required however, a reliable statement on the flight crew performance is harder to make. For this purpose a Boeing team has developed a tool named the Cockpit-operations Reliability Evaluation Worksheet (CREW). CREW collects all factors impacting flight crew performance and supports a decision on whether and which changes to interface, procedures or training are advisable to ensure safe operations. This paper discusses some considerations leading to the development of CREW as well as its application to the unreliable airspeed scenario. The analysis of the unreliable airspeed scenario was one of the first larger CREW campaigns.

1 Introduction

The design of complex, safety-critical systems, such as airplanes, must ensure that non-normal situations can be managed by human operators to avoid undesired outcomes. That is, when airplane system failures occur, there may be a need for the flight crew to intervene in order to ensure continued safe flight and landing.

When airplanes are designed and certified, safety analysis is used to identify key failure scenarios where flight crew intervention will be required. Then, alerting, indications, system controls and flight controls are designed to reduce risk associated with flight crew performance. Thus, an appropriate level of flight crew reliability is "designed in." However, unanticipated failure scenarios can emerge as the airplane fleet operates around the world, and safety analysis needs to determine whether the appropriate level of reliability exists for these unanticipated failures. These scenarios may involve unexpected system failure sequences or may arise because system reliability is less than was originally anticipated. Additionally, the fleet may be subject to more extreme operational and environmental factors than anticipated. For this analysis, which is referred to as continued airworthiness, it is important to estimate the actual level of reliability in the system including the human element.

This in-service safety process is monitoring events such as accidents, incidents, events reportable to the regulator (in the U.S. according to FAR 21.3) as well as in-service reports from operators. As shown in Figure 1 the manufacturer is reviewing a large number of events every year to identify potential safety issues which have to be addressed by appropriate design action, procedural changes or changes to training. In the decision process of whether an in-service scenario is to be considered a potential safety issue and will be addressed by manufacturer action, a risk-based approach is being followed.

Analysis of the unreliable airspeed scenario was part of a larger industry effort. On the following pages the scenario will be detailed, the mechanics of the CREW tool will be discussed and some conclusions on flight crew handling of unreliable airspeed will be offered.

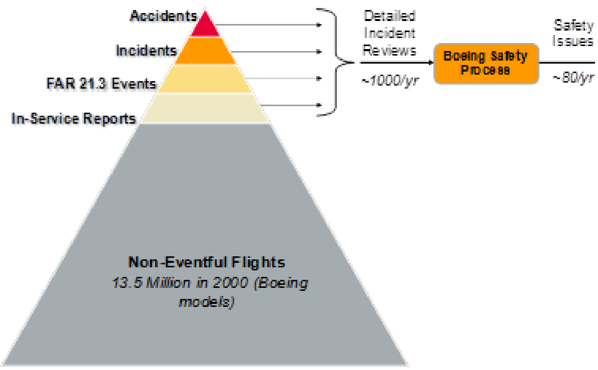

Figure 1: In-service safety assurance [1]

2 The Scenario

In recent years industry began a set of analyses to better understand flight crew handling of unreliable airspeed events. First step of the company response was to query internal and external aviation safety data sources in order to scope the problem. This search highlighted well-known accidents and significant incidents (Figure 2), and also a relatively large number of in-flight events, yielding an overall rate of about 10-7/flight hour. Analysis of this dataset showed that flight crews responded appropriately to erroneous airspeed during the takeoff phase, attributable to the well-trained 80 knot speed check. However, crew response to erroneous airspeed in flight cast doubt on an expectation that at least 99 in a hundred crews would perform reliably.

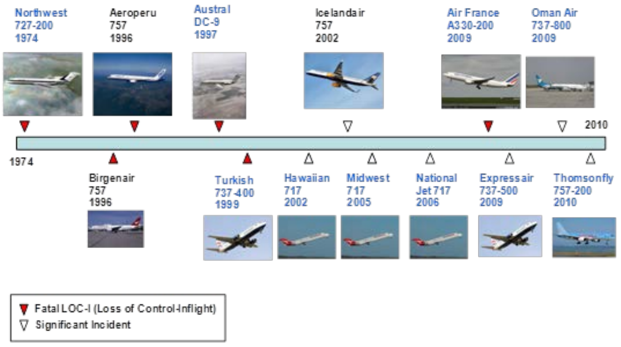

Figure 2: Significant Pitot-Static events leading to unreliable airspeed

The next step was an extensive simulator campaign reproducing a large set of unreliable airspeed conditions as well as flight deck effects across the various aircraft models. Further, the team worked with technical pilots in order to establish the desired flight crew actions for the scenarios. Both~~,~~ the flight deck effects as well the desired flight crew actions are essential inputs to the CREW analysis, as the method is based on scrutinizing which guidance the flight receives from procedures, indications and training. Before some of the results will be discussed the mechanics underlying the CREW tool will be explained.

3 The CREW-Tool

The CREW methodology, especially its development and theoretical basis was discussed extensively in [3]. The tool determines whether based on the guidance provided by interface, procedures and training sufficient flight crew performance can be expected in a given non-normal situation. That implies that rather than asking which reliability can be achieved, the CREW-method up-front calculates the flight crew reliability required based on the reliability of the other system components and the reliability requirement for the total system (e.g. the maximum catastrophic event probability of 10-9 per flight hour as set forth in [6], shown in Figure 2). Subsequently the method runs through a step-by-step analysis of the factors impacting the crew's ability to reach the target flight crew reliability.

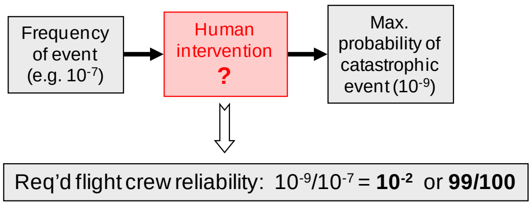

Figure 3: CREW determines required flight crew reliability

The theoretical basement of the method is Rassmussen's step ladder model (SLM), a symbolic processing model described in [2]. It was chosen due to its proven validity and high level of robustness.

3.1 The CREW Framework

The method is based on the "Cognitive Demands Checklist" described in [4] which was conceived for use in the nuclear industry. It is built around the Rasmussen SLM with only slight modifications. The tool systematically analyzes flight crew support from interface and procedures in performing the necessary recovery actions. To achieve this, CREW breaks the process of crew decision making and taking action into six phases which are derived from the Rasmussen approach.

Figure 4 gives a schematic view of these phases.

The six phases are:

1) Detect: The flight crew has to detect that a non-normal situation is occurring.
2) Understand: The flight crew has to interpret the indications and alerts associated with the scenario in order to come to a coherent explanation of the event. This may include interpreting flight deck mediated effects as well as external manifestations of the event.
3) Prioritize: In this step the flight crew prioritizes between several concurrent goals. This step is vital in events where the flight crew is confronted with multiple failure scenarios.
4) Select action: Based on the interpretation of the event and the goal prioritization the flight crew has to determine which the correct recovery actions to be taken are.
5) Intentional deviation: Based on alternative explanations of the indications or a wrong mental model of the situation that the flight crew might develop for a variety of reasons the crew might choose to deviate from the standard recovery path, may it result in successful recovery or not.
6) Execute: The flight crew has to physically execute recovery actions.

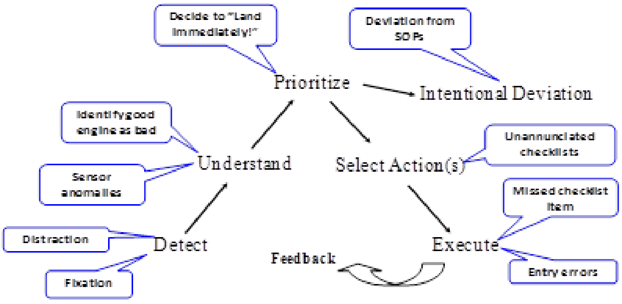

Figure 4: CREW implementation of the Rasmussen model and typical sources of flight crew errors

Service history shows that the chain of action can break down in any of the above phases leading to unsuccessful recovery.

Figure 4 also provides some typical sources of error. For successful recovery from a non-normal situation the crew has to pass all phases which can either be facilitated or hindered by interface and procedural factors. These factors to be discussed in the following section are the 'helps and hinders' of the CREW tool.

3.2 Helps and Hinders

In a similar fashion as proposed for nuclear applications by Roth and Mumaw in [4], the CREW development team set out to compile a comprehensive, flight-deck relevant checklist of factors that support or adversely affect flight crew performance in the respective phases described above. The set of these influential factors was produced with extensive help from safety, technical, research and training pilots. Naturally it is here where the CREW method is most flight deck specific, a result of a manufacturers design philosophy and architectures. However, based on the experience from adapting the tool to the flight deck environment, it can be said that part of the help and hinder factors are universally applicable to operators of complex systems, as is the general CREW framework.

The goal was to make explicit all the implicit reasoning performed by expert pilots whenever evaluating flight crew performance. It was cast it into a checklist format consisting of binary questions at a level of abstraction just high enough to be applicable to virtually all flight deck scenarios. For additional benefit of the analyst when applying the checklist, every help/hinder question is exemplified by several operational examples. The total checklist covers approximately 100 help and hinder factors across the performance phases discussed in the previous section extended by the two phases: detection of errors and recovery from errors. All information pertinent to the help/hinder analysis or the detailed description of the non-normal scenario, such as checklists to be executed by the flight crew, required actions outside of checklists, controls to be actuated, etc., is captured by the CREW tool.

Some examples of hinder factors are:

Detection phase

- An indication is not readily distinguishable as a non-normal indication; it is normal by interface presentation, but non-normal in the operating context.
- Indication is not in forward field of view or at central location.

Selecting an action

- The checklist to be executed is an unannunciated checklist.
- Criterion for selection of correct checklist requires sustained or repetitive monitoring to judge.

Execution

- Execution of the necessary actions is physically demanding.
- Execution requires closed loop monitoring of the action’s effects.

3.3 Integration of CREW Assessment Components

Key for the CREW assessment is how all the information is integrated to inform an appropriate safety decision, i.e. answering the question can be reasonably assumed that the support provided by interface and procedures guides the flight crew to successfully and reliably take appropriate recovery actions or does the situation merit design, procedure or training changes. In order to do this, the analyst has to determine what the required performance and urgency of pilot action is in this situation. How required flight crew reliability is determined is shown in Figure 2. Urgency of the flight crew response is determined based on 3 simple questions and is being categorized as either

- *Time Critical Warning level:* Typically flight path related issues, that require immediate flight crew action,
- *Warning level:* Typically system failures that require remedial crew action,
- *Caution level:* Typically failures that require crew awareness, or
- *Routine attention level.*

To automatically evaluate the flight deck effects for a given non-normal situation against the urgency requirements stated above comprehensive, airplane model specific flight deck effects reporting modules have been developed. Figure 5 is providing a graphical picture with erroneous flight deck effects (misleading the flight crew) shown in amber, accurate indications in green and system consequences in gray.

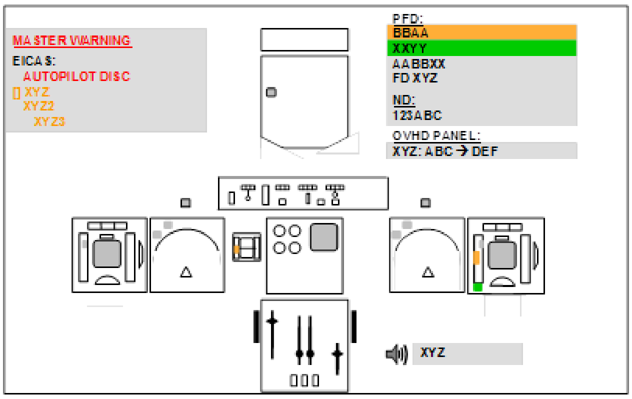

Figure 5: Generic graphical visualization of flight deck effects

The CREW tool then evaluates the flight deck effects against the alerting requirements defined by the Boeing flight deck philosophy, i.e. certain modes (voice, tone, visual, tactile) are required for certain alerting levels. If those requirements are not met CREW identifies a potential issue in the detection phase as shown in Figure 6.

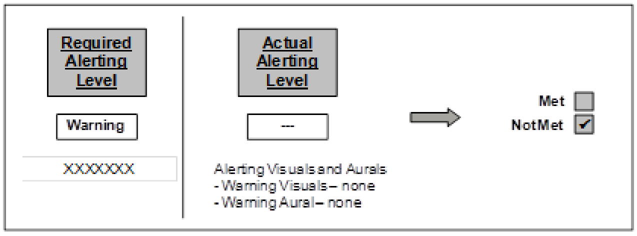

Figure 6: Automatic evaluation of urgency, i.e. required alerting level

Naturally, evaluation of the adequacy of the alerting level allows a statement typically limited to the detection phase described above. To get the complete picture of the scenario, this part of the analysis has to be fused with the helps and hinders the assessment has identified as applicable to the scenario. For this purpose the help and hinder factors are visualized by Rasmussen performance phase in context with the required flight crew reliability, e.g. 99/100 pilots are required to perform adequate recovery from the non-normal situation in order to reach the maximum allowable fleet risk.

Figure 7 gives a template CREW decision support screen.

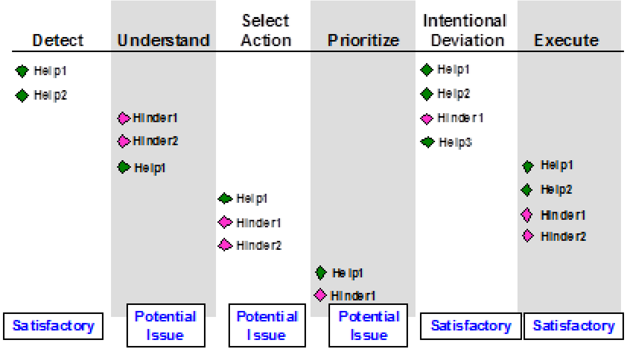

Figure 7: CREW decision support screen with help and hinder factors broken out by Rasmussen performance phase

Part of the decision process, the relating of helps/hinders to the required crew performance, is still up to subjective judgment. However, this decision is made phase by phase, with a comprehensive set of impacting factors in view. The challenging question of "How many crews will do this right?" is decomposed into more manageable portions, i.e. yes/no questions when performing the actual help/hinder analysis. The factors applicable to the situation are identified and a final assessment on whether adequate flight crew performance can be reasonably assumed is made. Thus, the decision process becomes auditable and consistent. In order to pass the assessment, support to the flight crew should be rated satisfactory for all phases.

3.4 CREW Application

The CREW assessment is typically performed by a team consisting of the safety engineer, human factors personnel, technical/safety pilots and, if required, systems experts. Especially for the definition of the scenario, including flight deck effects in the situation to be studied, systems experts might be required. The pilots are extremely important during the help/hinder assessment as well as for the final evaluation. The safety engineer takes ownership of the assessment process and ensures adequate expert input at the right points of the process.

4 Results

As of today, CREW has been applied to a number of in-service scenarios and has proven fit for purpose. The unreliable airspeed scenario was one of the most challenging scenarios studied so far. As was found in particular multi-channel unreliable airspeed events may be challenging to handle as they can produce interrelated effects in multiple systems which can be partially contradicting and bear the risk of misleading the flight crew. As is the case with other failure scenarios in complex, highly integrated systems, the ability of the flight crew to correctly interpret flight deck effects and integrate those into a coherent explanation of the problem might become very important. Further, multiple failure effects across systems may trigger several warnings and cautions from the airplane's centralized alerting system which may make it harder for the flight crew to get to the right procedure to execute. Figure 8 is showing in which of the Rasmussen phases CREW analysis has uncovered potential issues to flight crew performance.

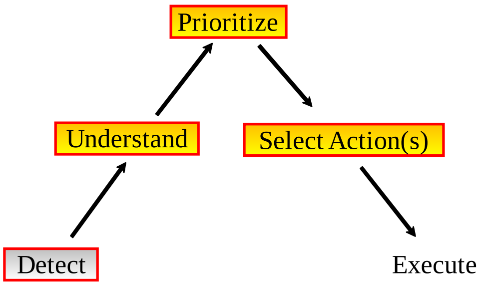

Figure 8: Rasmussen phases with potential hinders to appropriate flight crew action

Figure 8 also highlights potential issues in detection phase which do not directly relate to the detection of the non-normal situation itself. Rather they are related to automatic evaluation of the alerting level as shown in Figure 6 applied to the failure of flight envelope alerts due to the lack of valid airspeed information. Statistical analysis of historical unreliable airspeed events presented by Silva and Nicholson in [5], seems to support that for unreliable airspeed related accidents the majority of issues occur in action selection and understanding, while for incidents the majority of breakdowns show in understanding of the non-normal condition.

5 Conclusions

CREW analysis has unveiled certain potential issues in handling of unreliable airspeed events especially between understanding the non-normal condition and selecting the appropriate procedure/flight crew action to ensure safe continued flight and landing. This correlates with a statistical review of historic unreliable airspeed accidents and incidents performed by another team. Further, detection of certain unsafe airplane states regarding safe operating speeds might be hampered in case a total lack of valid airspeed information arises. The CREW tool, informed by an extensive simulator campaign to develop a comprehensive picture of the flight deck effects, has proven to be a valuable method for analysis of non-normal in-service events and providing effective decision support to the manufacturer’s safety review board. The findings have been reviewed and will be considered in future designs.

References

[1] R. Kennedy, BCA Aviation Safety - The Role of the Manufacturer in Enhancing Safety in the Aerospace Industry, RNLI Safety Conference, 2010.

[2] J. Rasmussen, Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering, Elsevier Science Publishing Co., Inc., New York, 1986.

[3] L. Fucke, R. Mumaw, R. Kennedy, R. Nicholson, Assessment of Risk from Human Performance on the Flight Deck – Driving Sound Design Decisions. IET System Safety Conference, Birmingham, September 2011.

[4] E. Roth, R. Mumaw, Preliminary Draft of a Cognitive Demands Checklist, Internal Research Report, Westinghouse Electric Corporation, 1992.

[5] S. Silva, R. Nicholson, Categorization of Unreliable Airspeed Events Using Rasmussen’s Human Performance Model, ICAS Conference, Brisbane, September 2012.

[6] Federal Aviation Administration, AC 25.1309-1A, System Design and Analysis, 1988.

Copyright Statement

The authors confirm that they, and/or their company or organization, hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, or have obtained permission from the copyright holder of this paper, for the publication and distribution of this paper as part of the ICAS2012 proceedings or as individual off-prints from the proceedings.