Applied Ergonomics, Volume 82, January 2020, Article number 102953
DOI:10.1016/j.apergo.2019.102953

Modelling and supporting flight crew decision-making during aircraft engine malfunctions: developing design recommendations from cognitive work analysis

Saryani Asmayawati, Jim Nixon
Safety and Accident Investigation Centre, Cranfield University, Cranfield, MK43 0TR, UK

Abstract

In this article, we analyse flight crew response to an in-flight powerplant system malfunction (PSM) using control task analysis. We demonstrate the application of the decision ladder template and the skills, rules, and knowledge (SRK) framework to this new area of inquiry. Despite the high reliability of turbofan engines, accidents and incidents involving PSM still occur. During these unusual events, flight crew have not always responded appropriately, leading to a reduction in safety margins or disruption of operations. This article proposes recommendations for technological and information system that can support flight crew in responding safely and appropriately to a PSM. These recommendations focus on new ways in which information from engine health monitoring system and other sources of data can be utilised and displayed. Firstly, we conducted knowledge elicitation using Critical Decision Method (CDM) interviews with airline pilots who have experienced real or simulated PSM events. We then developed generic decision ladders using the interview data, operations manual, training manual, and other guideline documents. The generic decision ladders characterise the different stages of responding to PSM identified as part of the research. These stages include: regaining and maintaining control of aircraft, identifying PSM and selecting appropriate checklists to secure the engine, and modifying the flight plan. Using the decision ladders and insights from the CDM interviews, we were able to identify cognitive processes and states that are more prone to error and therefore more likely to generate an inappropriate response. Using the SRK framework, we propose design recommendations for technological and information systems to minimise the likelihood of such inappropriate response. We conclude that this combination of methods provides a structured and reliable approach to identifying system improvements in complex and dynamic work situations. Our specific contributions are the application of these techniques in the unrepresented area of flight operations, and the development of evidence-based design recommendations to improve flight crew response to in-flight powerplant system malfunctions.

Keywords: cognitive work analysis; flight crew decision-making; aircraft engine malfunction

1. Introduction

Modern turbofan engines powering most large transport aircraft have unparalleled reliability and rarely demand constant attention from the flight crew. The Australian Transport Safety Bureau (ATSB) (2014) reviewed engine related occurrences reported to them between 2008 and 2012 and found that there is approximately one powerplant-related occurrence affecting turbofan aircraft every 20,000 flight hours. This low occurrence rate is evidence of the high reliability of modern turbofan engines; however, they do fail and therefore the flight crew must be able to respond appropriately to ensure continued safe flight.

Inappropriate flight crew response to powerplant system malfunctions (PSM) is a threat to safe and efficient operation. The high reliability of turbofan engines, their failure characteristics, and flight crew experience and training were identified as causal factors to these inappropriate responses (FAA, 2009). Roelen and Wever (2002) evaluated 476 events involving commercial western-built turboprop and jet aircraft between 1990 and 2000 and found that flight crew response to system failure was inappropriate in 19% of the sample, and that incorrect diagnosis or decision was a factor in 38% of those events. In Roelen and Wever's study, almost half of the events they evaluated were PSMs. From these data, it is clear that understanding decision-making during PSM events is an important line of inquiry. In addition, Aerospace Industry Association/European Association of Aerospace Industries (AIA/AECMA) (Sallee and Gibbons, 1998), report that there has been an increase in the rate of inappropriate crew response to PSM despite the significant improvements in engine systems since the 1980s. These improvements include engine control and alerting systems that enable flight crew to identify and manage the more common types of PSM appropriately. However, despite such systems, the studies indicate that flight crew still have difficulty identifying certain PSM and thus responding appropriately (Sallee and Gibbons, 1998; FAA, 2009), potentially reducing safety margins and/or disrupting operations. This indicates the significant potential for improvement in how aircraft and engine systems support flight crew decision-making during PSM events.

Additional information streams that are currently being researched from advanced engine health management (EHM) systems may help in minimising the likelihood of inappropriate flight crew responses leading to unnecessary in-flight shut down (IFSD) and diversions (Karnofski, 2015; Ulizar et al., 2016). Modern EHM systems are capable of acquiring hundreds of parameters from measurements on engine gas path, oil-fuel system, vibration, structural assessment sensors, and sensor conditions and fidelity (Ulizar et al., 2015). Combined with engine models and maintenance history, comprehensive diagnostic and prognostic knowledge regarding the health of the engine and failure conditions will play a significant role in a not-too-distant future. This has the potential to improve the range of decisions made in respect of powerplant system malfunctions (PSM) beyond IFSD or diversions, whilst ensuring safety. However, improvements need to be incorporated properly into the flight deck, taking into account how decisions are made in the operational context.

In this article we model decision-making in response to engine malfunction. Using this analysis we specify design recommendations for technologies that can improve decision-making and decision outcomes in response to PSM. Using a combination of the decision ladder method and the skills, rules and knowledge (SRK) framework we develop design recommendations for new powerplant information systems. The contribution of this article is two-fold. Firstly, applied study has been offered which focuses on this important but under-researched part of flight operations, especially in light of the new health monitoring technologies available. Secondly, we contribute to the methodological approaches in this area through the combination of decision-making models and SRK model to elicit technology requirements.

Figure 1 shows that many PSM have similar symptoms and therefore precise determination of PSM type from the engine instrument indications and other cues may not always be practicable. A loud 'bang' heard in the flight deck followed by the aircraft yawing could relate to a number of different PSM or indeed not be related to an engine at all. In any case, it is important to note that diagnosis of the exact cause of the PSM is neither necessary nor safe if it diverts resources from flying the aeroplane (Sallee and Gibbons, 1998; FAA, 2009; Airbus, 2006). Therefore, any additional technology and its interface should support decision-making with minimal demand on those resources.

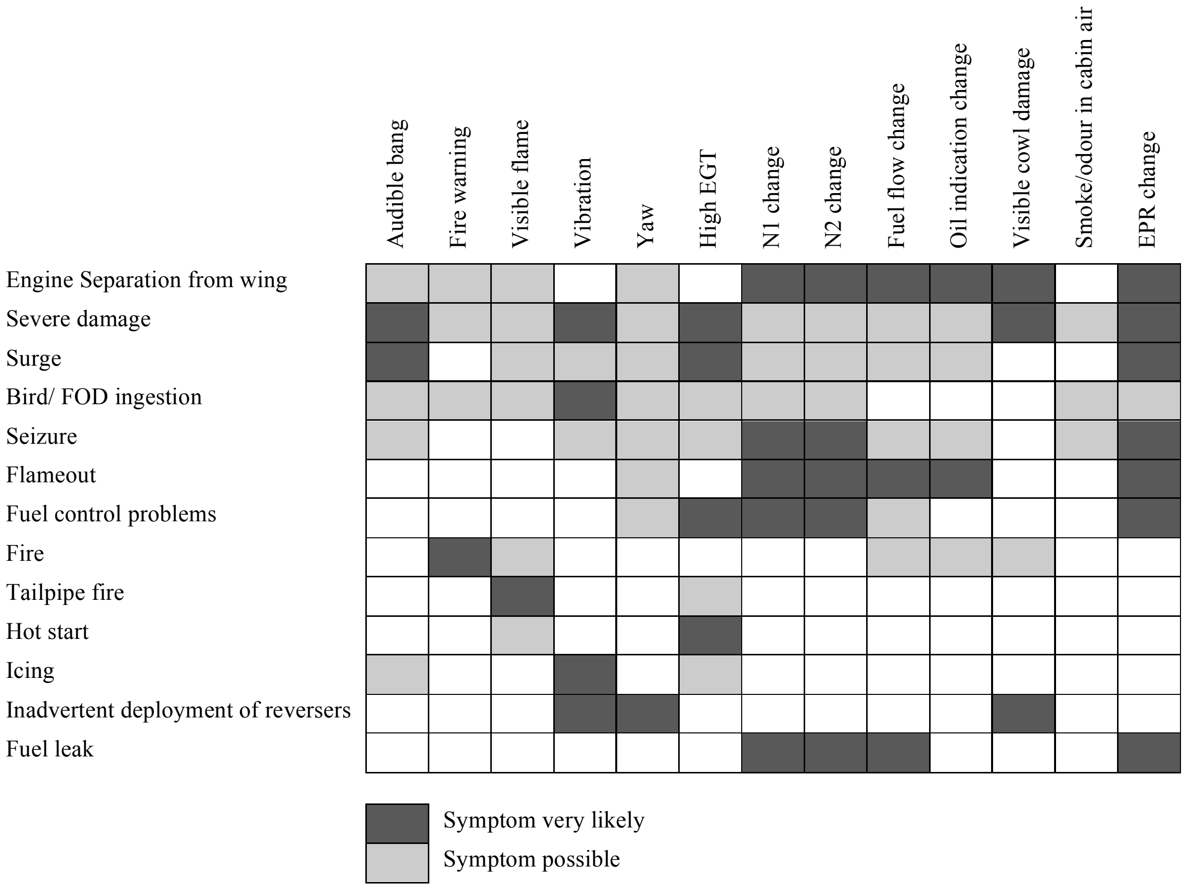

Figure 1 Summary of engine conditions and malfunction symptoms that can be used by flight crew to diagnose type of non-normal condition or malfunction during flight (FAA, 2009).

To begin to understand flight crew decision-making in response to PSM we employed the aeronautical decision-making (ADM) model developed by Orasanu and Fischer (1997), as shown in Figure 2. Two key categories of events are specified by the model as *well-defined* and *ill-defined* events. We have used these two categories to structure and understand the different decision-making processes taken in response to various PSM. Generally, well-defined PSM events are specified in the airline's standard operating procedures and training manuals (Orasanu & Connolly, 1993), whilst an ill-defined PSM event will require further information gathering and can also lead to a creative 'solution' since there is no specified procedure to address the problem. Alternatively, when there are not enough resources to acquire further information, flight crew will resort to 'satisficing' decision-making by applying a procedure such as in-flight engine shut down (IFSD). This could explain ATSB's findings where 34 of the 55 events with abnormal engine indication resulted in IFSD of the affected engine, even though many abnormal engine indications can be spurious or insignificant (ATSB, 2014).

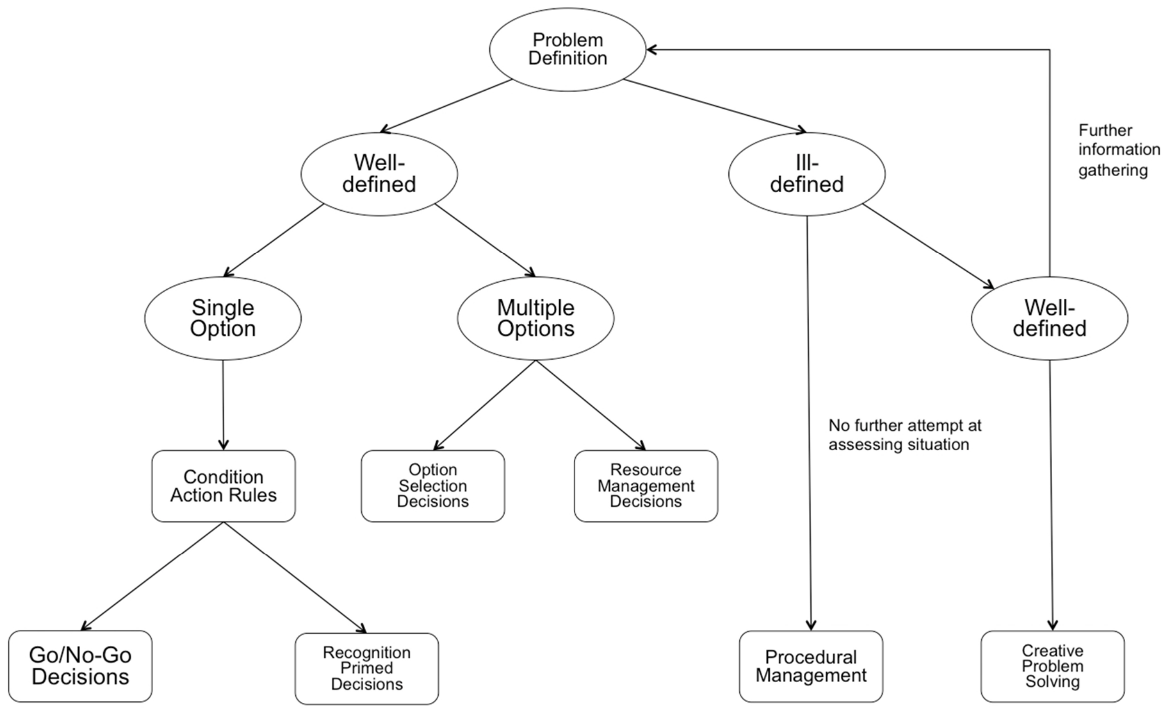

Figure 2 Taxonomy of decision types in aeronautical decision-making (Orasanu & Fischer, 1997, as adapted by Harris, 2011) used in this study to provide structure in analysing decision-making processes in various powerplant system malfunction events.

In this article we examine flight crew decision-making strategies employed during the two categories of PSM events in order to identify how technology can support them. The work has two aims. First, to analyse flight crew decision-making during PSM in a structured way. Second, to derive recommendations for the technology and information system to support flight crew decision-making from this structured analysis.

The analysis method used to meet these aims is control task analysis using the decision ladder template. For this study, the method has to be able to generate system requirements associated with expert actions not only when responding to familiar scenarios such as well-defined problems but also when they are confronted with ill-defined ones. Rasmussen's decision ladder template (1974; Vicente, 1999) is a tool that was designed to do exactly this (Naikar, 2010). The original decision ladder template was proposed by Rasmussen based on his field observations on power plant operators. The tool has since been applied in safety critical sectors such as finance (Li, Burns, & Hu, 2016), rail (Mulvihill et al., 2016), traffic safety (Cattermole, et al., 2015), automotive (McIlroy & Stanton, 2014), manufacturing industry (Hassall & Sanderson, 2014), military (Jenkins et al., 2010; Jenkins et al., 2011a), policing (Jenkins et al., 2011b), and unmanned aerial vehicle design (Elix & Naikar, 2008; Jenkins, 2012), among others. For this study, we used the decision ladder template modified by Naikar, Moylan and Pearce (2006) as shown in Figure 3, which was based on their review of a range of papers by Rasmussen and Vicente. In this modified template, shading is used to highlight cognitive tasks and processes involved in the activity under investigation.

The decision ladder template also incorporates skill-based, rule-based, and knowledge-based behaviour (SRK) framework (Rasmussen, 1983, 1986). According to Rasmussen, the three levels of operator performance correspond to the decreasing levels of familiarity with task or environment, or increasing level of cognitive complexity associated with those tasks. At skill-based level, performance is essentially subconscious, based on the automated activation of highly learned set of behaviours (Rasmussen & Lind, 1982; Rasmussen, 1983). Errors at this level are typically associated with the variability of human performance (Reason, 1990). At rule-based level, performance is based on recognition of situations in relation to the rules for actions derived from previous experience, instructions, or procedures (Rasmussen & Lind, 1982; Rasmussen, 1983). Errors at this level are generally related to the misclassifications of situations or incorrect recall of procedures (Reason, 1990). Knowledge-based performance typically occurs during unfamiliar situations, where the operator has to use conscious analytical processes and existing knowledge to solve novel problems. Thus, errors at this level are mainly due to incomplete or incorrect knowledge and resource limitations (Reason, 1990). This framework is used here to identify where and how inappropriate response might occur and determine technology requirements to minimise its incidence. In the aviation sector, this is the first use of this combination to understand flight crew response in this new applied domain.

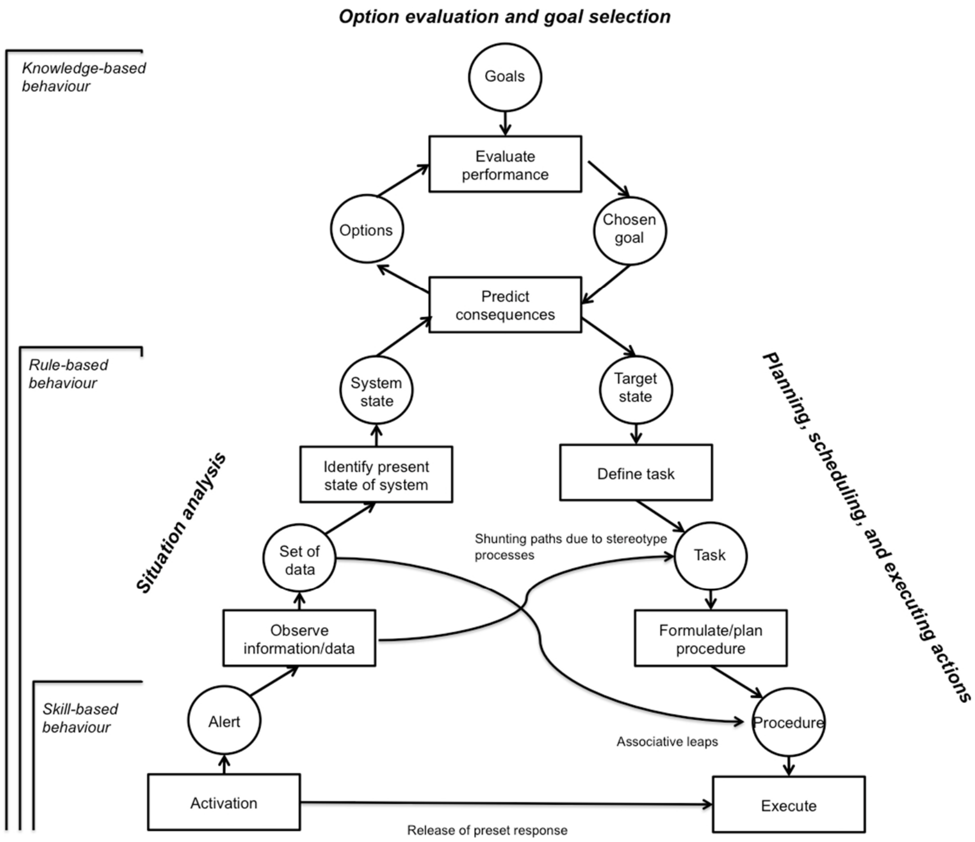

Figure 3 Decision ladder template as modified by Naikar, Moylan and Pearce (2006) incorporating Rasmussen's skills, rules, knowledge framework as adapted by McIlroy and Stanton (2015).

The decision ladder template was developed to conduct control task analysis, which is the second phase of cognitive work analysis (CWA). With the template, we are able to investigate (1) task goals and the constraints on those goals, and (2) the information that is relevant for particular situations (Vicente, 1999). The design requirements could take form in constraint-based procedures, automation, or context-sensitive interface that presents the right information at the right time (Mitchell & Saisi, 1987; Woods, 1991). This aspect is critical for work presented here since we generate design recommendations for new systems through a structured, naturalistic analysis of flight crew response to both well- and ill-defined PSM event.

2. Methodology

2.1 Design

We employed critical decision method (CDM) to interview subject matter experts (Klein et al., 1989; Hoffman et al., 1998), which has been shown to be effective in constructing decision ladders (Cattermole et al., 2015; Naikar, Moylan & Pearce, 2006; Naikar & Saunders, 2003). The CDM interviews also assisted in understanding the information and cues that flight crew use when making decisions associated with PSM events, and any difficulties they encounter at any stage during the event. Since decision ladders should be independent of system or event (Vicente and Tanabe, 1993), we developed generic decision ladders that cater for any PSM event regardless of the aircraft type/level of automation.

The decision ladders were developed by the first author (SA) using the CDM interview tables, complemented by applicable operating and training manuals as well as guidelines documents. The decision ladders were then validated by the second author (JN) and an airline pilot with fourteen years of commercial flying experience and over 6,000 total flight hours on two types of twin-turbofan large transport aeroplane that represent older and newer generation.

The validated decision ladders were then evaluated using SRK framework to identify corresponding improvements to system design that can support flight crew decision-making process and minimise potential for errors or inappropriate responses.

2.2 Participants

According to Weitzenfeld et al. (1990), CDM interviews for knowledge elicitations are typically conducted with 7-8 SMEs. In this study, seven pilots operating on commercial twin-turbofan-powered large transport aircraft from three different airlines were recruited as SMEs directly and via referral sampling. All participants volunteered and were briefed of the objectives and the nature of the study and gave written consent prior to the interviews. Table 2 shows the details of the participants, including current and past type ratings, flight hours, rank (Captain and Senior First Officer), age group, and previous operational background (civil or military). The seven pilots covered 12 different current and past type ratings, with total flight hours on commercial twin turbofan-powered large transport aircraft ranging between 5,000 to 20,000 hours.

Ten incidents were recounted by the participants. Three of the airline pilots were interviewed based on training/check events, and four airline pilots used real experience. Three of these four pilots experienced more than one real in-flight PSM event. The ten incidents covered both ill-defined and well-defined PSM more than once with insignificant variations on the main decision points, which gave us confidence in the number of participants interviewed. Summary of the incident information can be found in Table 3.

2.3 Procedure

Table 1 shows the CDM process and probes adapted from Hoffman, Crandall & Shadbolt (1998), which were used to explore task goals, planning, decision-making, and strategy formation. Line pilots were recruited and interviewed as subject matter experts (SMEs) and asked to develop a timeline of the cues and actions that were taken in response to a real or simulated PSM event. Interviews were conducted individually by the first author. All but two interviews were conducted face-to-face; remote interviews were conducted using video conferencing. A timeline was created by the interviewer as instructed and checked visually by the interviewees. Interviews lasted between 55 to 180 minutes, as some participants recalled more than one incident.

Table 1 Stages of the CDM process and the probes used during the interviews.

Stage	Description and probes used
(1) Incident recall and timeline verification	The participant was asked to recount an event where they functioned in a key decision making role. The interviewer drew the events out on paper, with approximate elapsed time, which were then confirmed by the participant as the timeline. This timeline was used as the basis for further cognitive probes and questions.
(2) Decision and task goals	The participant was asked to describe the objectives of decisions and actions taken in each meaningful event in the timeline.
(3) Cues/information used	The participant was asked about the information and cues used in each meaningful event in the timeline.
(4) Progressive deepening	The participant was asked to go into further details in each meaningful event in the timeline. This includes any difficulties that they encounter, and additional information or system changes/improvements that could have assisted in the tasks or decision-making processes that they had identified in stages 2 and 3.
(5) What-if queries	When appropriate, the participant was asked how they would respond when given different factors or scenarios. Participant was also asked what could have gone wrong or what likely errors they could have made in each meaningful event.

Table 2 Background information of the participants interviewed in this study.

Participant number	Current Type Rating	Engine make/model installed on the aircraft currently flying	Approximate flight hours on current type	Past type ratings	Approximate total flight hours*	Rank	Length of service in current airline	Age category	Background	Operator
A1	B787	RR Trent 1000	500 hours	B737-300/400/500	5,000 hours	Senior First Officer	> 3 – 5 years	31 – 40	Civil	Airline A
A2	B777	RR Trent 800, GE 90	4,000 hours	B757, B767	12,000 hours	Senior First Officer	> 10 years	41 – 50	Military	Airline A
A3	A330, A340	RR Trent 700, RR Trent 500	10,000 hours	A319, A320, A321	13,000 hours	Senior First Officer	> 5 – 10 years	41 – 50	Military	Airline B
A4	A320	IAE 2500	460 hours	B737, B747-400	8,000 hours	Captain	> 10 years	31 – 40	Civil	Airline A
A6	B777	RR Trent 800, GE 90	15,000 hours	L1011, DC10	20,000 hours	Captain	> 10 years	51 – 60	Civil	Airline A
A7	A330	GE CF6	2,800 hours	A320, B737 EFIS, BAe 146	18,500 hours	Captain	> 10 years	41 – 50	Civil	Airline C
A8	A330	GE CF6	3,000 hours	A320, B737, BAe 146	17,000 hours	Captain	> 10 years	51 – 60	Civil	Airline C

* On commercial large transport aeroplane (multi-engine turbofan) only

Table 3 Summary of powerplant system malfunction events recounted by participants in CDM interviews.

Event number	Type of event	Aircraft type flown during event	Role during event	Event	Phase of flight	Relevant actions and outcomes
A1	Simulated (LOE)	B737-400	PF	Single engine surge	Cruise FL340	Affected engine did not recover after completion of ENGINE SURGE checklist, so SEVERE ENGINE DAMAGE checklist was selected and affected engine was shut down. Diverted.
A2	Real	B757	PM	Single engine surge	Cruise FL250, top of descent	Affected engine was set to idle and continued to destination. Captain's previous experience had influenced the decision making process during the event. Participant stated that in hindsight, they might have considered shutting down the engine.
A3	Simulated (LOE)	A330	PF then PM	Dual engine failure and fire due to bird strike, instrument lost	Climb FL270	Restart attempted on both engines. Did not follow ENGINE FIRE checklist which would have shut down the engines as unsure if it could be restarted afterwards.
A4	Simulated (LOE)	A320	PF	Single engine fire	After take off	Affected engine was shut down (followed procedure ENGINE FIRE). Turned back.
A6_1	Real	B777	PM	Dual engine failure	Approach	The remainder of the flight was focused on flying the approach and landing, mainly controlling speed and height. Continued approach and conducted emergency landing in destination airport.
A6_2	Real	B777	PM	Vibration due to mechanical damage	Top of climb	Vibration was not reduced when engine was set to idle, so it was shut down (followed SEVERE ENGINE DAMAGE checklist). Turned back.
A7_1	Real	A330	PF	Gradual decrease of oil quantity	From top of climb until landing	Closely monitored oil quantity and calculated rate of decrease. Continued to destination.
A7_2	Real	B737	PF	High oil temperature	Cruise	Shut down affected engine (followed procedure ENGINE FAILURE/SHUTDOWN). Diverted.
A7_3	Real	B737	PF	Compressor stall	Cruise	Nothing done to engines as no abnormal indications of engine parameters, just bang noises. Decided to divert just in case.
A8	Real	A330	PM	Bird strike followed by vibration	After take off	Affected engine set to idle. Diverted.

LOE: Line Operational Evaluation PF: Pilot Flying PM: Pilot Monitoring

3. Results and Analysis

We describe the activities involved in responding to an in-flight PSM with control task analysis. The interview data was translated into generic decision ladders that are then used to identify potential sources of error or variable response. Based on this analysis we develop design recommendations which are directly and transparently linked to the decision ladders created.

In accordance with the principles of control task analysis (Vicente, 1999), the decision ladders are created with the flight crew as a unit regardless to their specific duties and responsibilities as pilot flying and pilot monitoring. Additionally, some tasks can be done by the flight crew or the aircraft systems, and the decision ladder does not discriminate between them. We also re-emphasise that the scope of this study is limited to in-flight PSM that occurs during climb and cruise phases.

It should be noted that whilst there are many aspects of flight operations that are highly procedural, flight crew still have some discretions in certain aspects of their tasks. In this study, we focus on using control task analysis to describe how flight crew *can* handle PSM events, taking into account how they *should* carry out certain tasks (based on procedures and training) and how those tasks are actually done (based on CDM interviews). Therefore, whilst certain patterns/sequence of activities in the decision ladders are defined by the constraints that are present during PSM (for example gaining/maintaining control of the aircraft before any other activities), other patterns will depend on skills/knowledge of the pilots and circumstances surrounding the PSM. The aim of developing these decision ladders was therefore not to specify the actual sequence of activities; rather, it was to identify the nature of the key control tasks or decisions during PSM, particularly with reference to the SRK framework.

Based on the interviews and reviews of Flight Crew Operating Manuals and Flight Crew Training Manuals, we found that flight crew response to PSM is based on a common philosophy, regardless of manufacturer, aircraft type, or operator. Diagnosis of the PSM is not started until certain conditions are met. These conditions are: (1) flight path and configuration of the aircraft are established to ensure a continued safe flight and landing; (2) the aircraft is not at a critical stage of flight; (3) the aircraft is at least a certain altitude, or obstacle clearance altitude, whichever is higher (Sallee and Gibbons, 1998; FAA, 2009; Airbus, 2006). Once the aircraft is controlled and the flight path is safe, the flight crew can focus on addressing the PSM. This normally involves diagnosis of the indications and 'symptoms' to enable selection of the most appropriate checklists and their execution. The diagnosis and outcome of the checklist actions will inform flight planning options, which could be to return to departure airport, divert, or continue to destination. This is inline with the overarching principle of flight operations, which is to "aviate/fly, navigate, communicate". With this information, we define three work situations for analysis using decision ladders, which are segmented according to time or stages (Rasmussen et al., 1994). These work situations are: (1) flying the aircraft, (2) addressing the PSM, and (3) modifying flight plan. Each work situation has specific goal and primary activities, as shown in Table 4.

Table 4 Work situations and their corresponding goals and primary activities represented by the generic decision ladders.

Work situation	Goal	Primary activities
WS1 – Flying the aircraft	To ensure that the aircraft and its path are safe and under control following the event.	(1) Regain and maintain control of the aircraft immediately following the PSM. (2) Confirm a non-normal situation that will need to be resolved. (3) Ensure safe flight path.
WS2 – Addressing the PSM	To identify the nature of the problem in order to select the appropriate checklist whilst maintaining safe flight.	(1) Diagnose of PSM type and affected engine. (2) Select and confirm applicable checklist.

		(3) Apply checklist.
WS3 – Modifying flight plan	To ensure a continued safe flight and landing at a suitable airport.	(1) Review and assess aircraft and engine system capability. (2) Identify alternatives to flight plan. (3) Review revised flight plan. (4) Carry out and communicate revised flight plan. (5) Manage flight (and occupants).

The first generic decision ladder addresses the initial stage of flight crew's response to a non-normal or emergency situation, which is regaining and maintaining control of the aircraft. In the second stage ('addressing the PSM'), we created two generic decision ladders to differentiate the activities involved in responding to a *well-defined* and *ill-defined* PSM. In this study, a *well-defined* PSM event is an event where the type of PSM and affected engine is annunciated [1], or, if unannunciated, the indications/symptoms are coherent and unambiguous which enables definitive identification of PSM type. In this case, it is implied that there is a suitable checklist/procedure to address the PSM. An *ill-defined* PSM event is an event where the PSM is accompanied with contradictory, ambiguous, or unfamiliar abnormal indications and symptoms, and/or there is no suitable checklist/procedure to address it. The last generic decision ladder addressed activities related to modifying the flight plan and managing the rest of the flight.

3.1 Generic decision ladder for Work Situation 1 (Flying the aircraft)

The first generic decision ladder (Figure 4) addresses flight crew's cognitive processes and states in regaining and maintaining control of the aircraft following a PSM event. As shown in the decision ladder, all activities here consist of skill-based and rule-based behaviours, which are as expected in a well-trained element of flight operations. Based on the CDM interviews, difficulties that could lead to ineffective or inappropriate response at this stage could occur on steps (1), (2), (3), and (4), with further improvements suggested for steps (6) to (10).

[1] Annunciated failures or malfunctions are those that can be automatically identified and announced by the aircraft and engine systems, usually through interfaces such as EICAS (Engine-Indicating and Crew-Alerting System) or ECAM (Electronic Centralised Aircraft Monitor). Such systems usually also provide associated checklist(s) to the flight crew.

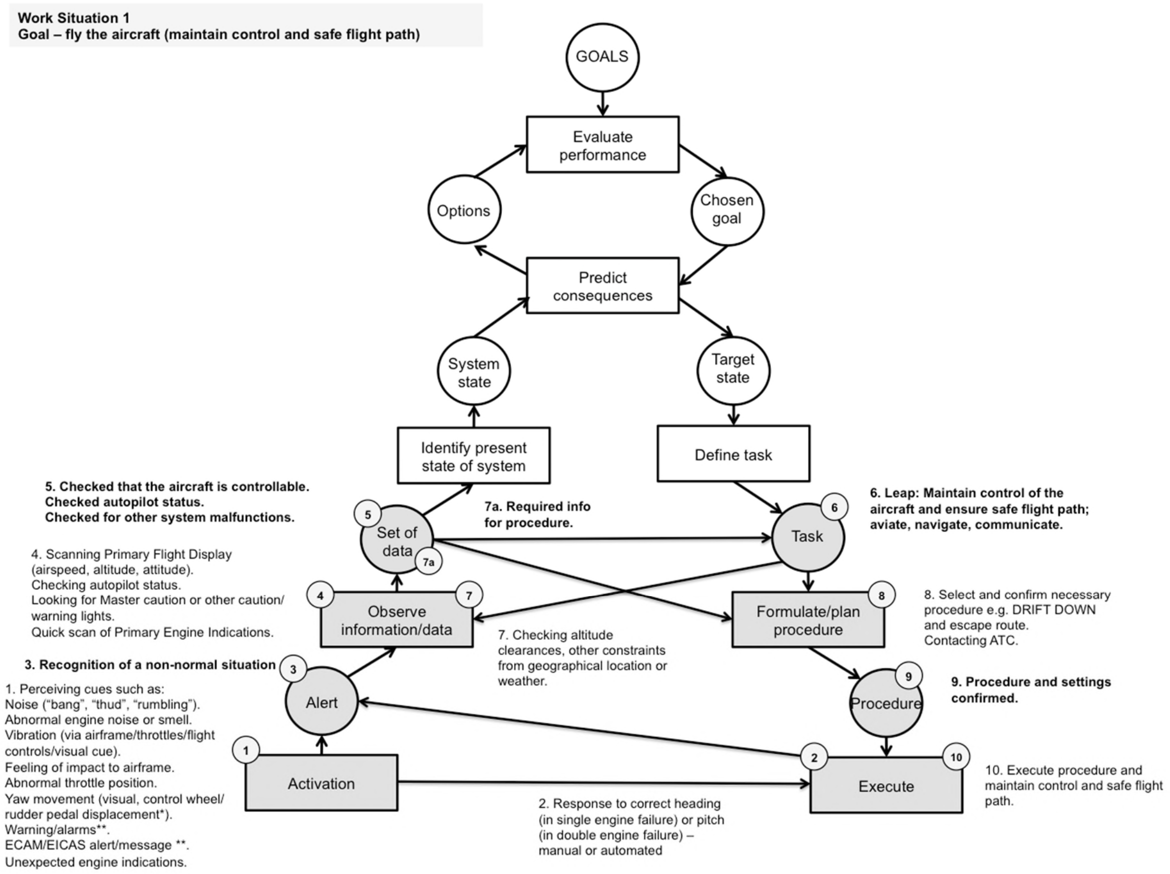

Figure 4 Generic decision ladder for the first stage of responding to a powerplant system malfunction, representing cognitive states and processes involved in regaining and maintaining control of the aircraft and securing a safe flight path.

For (1), (2) and (3), inappropriate responses could be a) the pilot reacting incorrectly to regain or maintain control of aircraft; b) not being aware when the aircraft system automatically corrected changes in flight performance or attitude due to PSM; c) not being immediately aware that there is a non-normal condition; d) not being immediately aware that the non-normal condition is related to PSM. According to Rasmussen and Vicente (1989), these skill-based activities could be influenced adversely by *lack of resources* (lack of speed, precision, or force), and *stochastic variability* (variability of attention or motor performance). These types of errors could be managed by automation and information/alerting system as discussed in Recommendations 1 and 2.

As rule-based activities, potential for error or inappropriate response in steps (6) to (10) could also be caused by *lack of resources* (inadequate memory for rules) and *stochastic variability* (erroneous recall of data or parameters related to rules). Recommendation 3 addresses these sources of potential error or inappropriate response.

Design Recommendation 1: The technology should automate aircraft control recovery action where possible and notify flight crew of this action promptly.

Lack of resources and stochastic variability in executing a pre-set response (step (1) to (2)), such as the application of counter rudder for asymmetric thrust, could be managed with automated functions or action supports as proposed in Lintern (2010). For example, if the pilot's focal attention interrogates which foot should press the rudder pedal rather than letting their pre-set reaction responds to the yaw (*overattention* – see Reason (1990)), they could press the wrong rudder pedal to counteract the yaw as a consequence. Additionally, pilots need to apply the right amount of rudder to avoid over-correction.

An automated aircraft control recovery system would prevent an incorrect response, and such systems are already installed on more modern aircraft. It should be noted that in such aircraft, the initial cue of a PSM event perceived by the flight crew will be the rudder pedal displacement instead of the yaw movement. Depending on the strength of the cues, flight crew may or may not be immediately aware of the non-normal condition, so it is important that flight crew is notified of any automated response actions executed by the aircraft system to regain control and maintain the flight path.

Design Recommendation 2: The technology should automatically detect and alert non-normal engine conditions promptly.

Some activation cues, particularly external ones such as noises, are not always immediately associated with an engine malfunction by the flight crew. This is especially true if the PSM did not instantaneously affect flight performance or trigger a specific alert/warning messages. Some alert/warnings are only triggered after a certain threshold, for example low oil quantity or vibration level, and that may occur when the engine condition is already deteriorated at the more critical phase of the flight. Perceiving the cues and recognition of non-normal situation in step (1) to (3) is a rule-based behaviour that could be supported by automated detection and alerting of non-normal engine condition. Immediate recognition of the situation could provide additional time for flight crew to address the condition before it deteriorates. The detection and alerting of non-normal engine parameter values in the context of the flight should encompass *abrupt changes (transient and permanent)* as well as *abnormal rate of change (increasing and decreasing)* of engine parameter values on the affected engine.

Design Recommendation 3: The technology should display synthesised information necessary to secure and control flight path.

Since the primary goal in Work Situation 1 is to gain/maintain control and ensure a safe flight path, the information sought and processed in step (4) is therefore not used to diagnose the engine problem. Rather, it is used to assess the overall aircraft conditions and quickly identify critical problems or faults that can affect safety, such as aircraft controllability issues or issues with other safety-critical aircraft systems such as cabin pressurisation. A summary or synthesis of information on aircraft controllability and system status can support the flight crew in achieving the goal of this work situation quickly and with minimum cognitive resource.

Once the status of the aircraft status and thrust availability is confirmed in step (5), steps (6) to (9) are reflective of the application of a well-trained procedure. During these steps, the flight crew are trained to recognise the situation as requiring the application of a specific procedure; for example, in most single engine or reduced thrust cases, flight crew will have to apply DRIFT DOWN[2] procedure with a previously discussed escape routes and necessary communications. A display of summary of required information, such as aircraft performance limitations coupled with traffic, weather, as well as geographical and topographical constraints, would reduce flight crew workload in planning the procedure whilst maintaining situation awareness. In the future, an integration with automated air traffic management system is another possible improvement that should be explored further.

3.2 Generic decision ladder for Work Situation 2 (Addressing PSM)

[2] In the event of an engine failure during latter stages of climb or during cruise, the aircraft will not be able to maintain its current altitude and hence it will have to descend (drift down). The DRIFT DOWN procedure involves setting maximum continuous thrust on the operating engine whilst countering any resulting yaw with rudder and trim, and disconnect autothrottle/autothrust if applicable. During this procedure, the flight crew needs to know local traffic to ensure separation and terrain clearance.

The next generic decision ladders address the second stage of flight crew's response to a non-normal or emergency situation, which is identifying the PSM to select the appropriate checklist to secure the engine. As shown in the decision ladder for Work Situation 2a for a *well-defined* PSM (see Figure 5), the activities consist of skill-based and rule-based behaviours, which are as expected in a flight crew response to a non-normal situation that are covered in training and/or procedures. Figure 6 describes the decision ladder for Work Situation 2b for *ill-defined* PSM, where the activities consist of all three types of behaviour – skill-based, rule-based and knowledge-based behaviours. Based on the CDM interviews, difficulties that could lead to inappropriate response during ill-defined events are mainly on the activities relating to interpreting cues and engine parameters to diagnose PSM, with the possibility of selecting inappropriate procedure or executing the procedure incorrectly. Whilst this is less likely to occur in a well-defined PSM event on aircraft with higher level of automation, further areas of improvements have been identified and addressed by the following design recommendations.

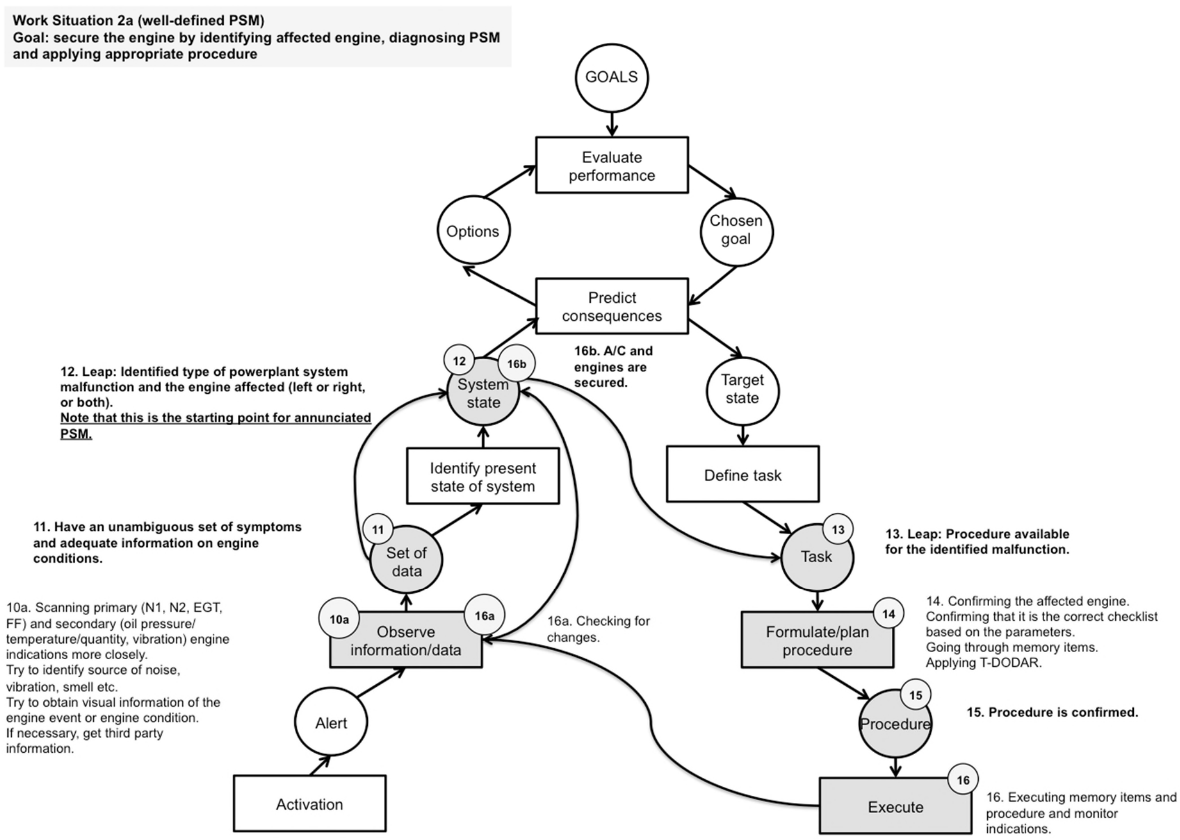

Figure 5 Generic decision ladder for the second stage of responding to a well-defined powerplant system malfunction event, representing cognitive states and processes involved in identifying the affected engine(s), diagnosis of malfunction, and selecting and applying appropriate procedure.

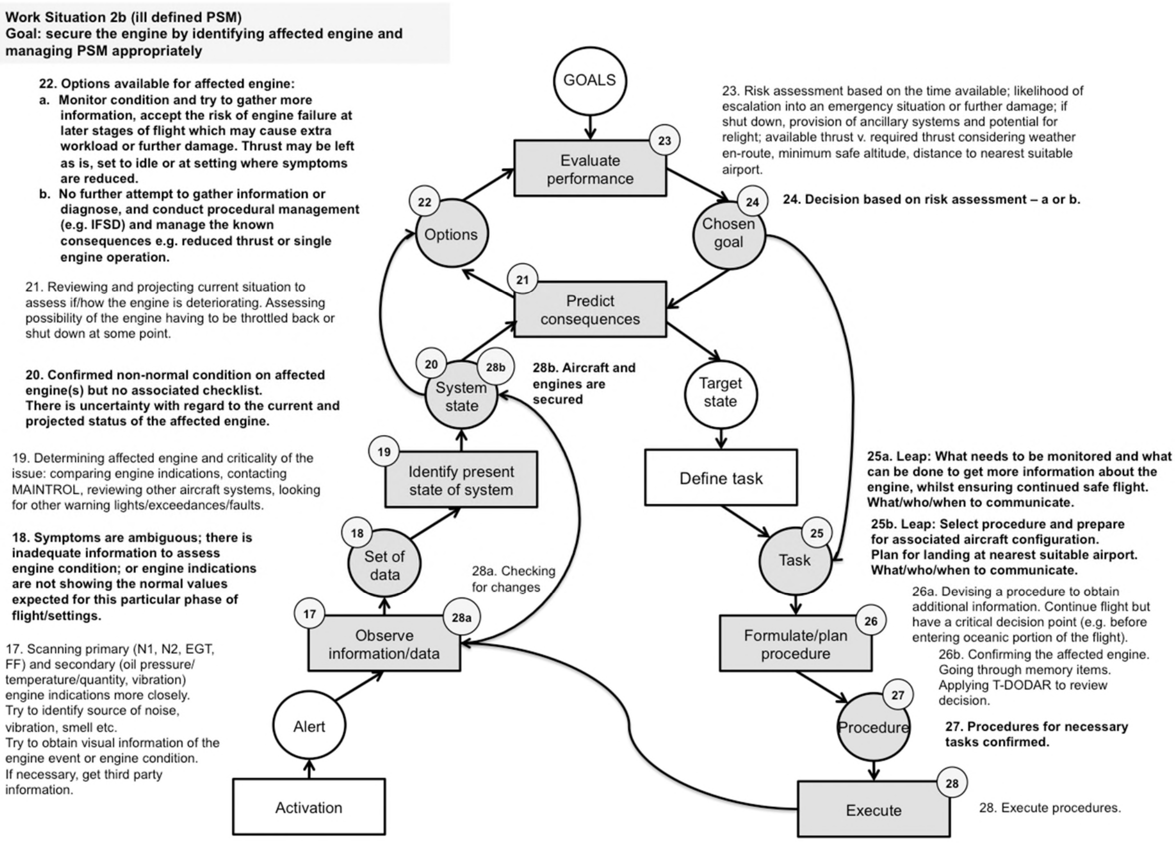

Figure 6 Generic decision ladder for the second stage of responding to an ill-defined powerplant system malfunction event, representing cognitive states and processes involved in identifying the affected engine(s), diagnosis of malfunction, and selecting and applying appropriate procedure.

Design Recommendation 4: The technology should exploit engine health monitoring data to improve the range of automated diagnosis of PSM.

This recommendation addresses activities relating to interpreting cues and engine parameters to diagnose PSM within steps (10a) to (12) and steps (17) to (20). For annunciated PSM, the decision ladder for WS2a (well-defined) starts at step (12), which means the resource-heavy and error-prone activities of *Observe information/data* and *Identify present state of the system* are no longer carried out by the flight crew. In an annunciated PSM event on newer generation aircraft, the flight crew need only to confirm and carry out the suggested checklist. Currently however, annunciated PSM only covers the more common types such as engine fire, engine failure/separation, engine surge/stall, engine overheat, and some component failures. Flight crew error and inappropriate response in PSM events are likely to be reduced if the automated diagnosis and alerting system is capable of a wider range of PSM types. It should be noted that engine health data must be captured and processed accurately and reliably, and alerts must be prioritised properly. Prioritisation of alerts will be required particularly in multiple or cascading system failures such as in the QF32 accident (Australian Transport Safety Bureau, 2013). Without prioritisation, multiple system messages/warnings may alter a *well-defined* problem into an *ill-defined* one. Nevertheless, in most cases, it is more likely that flight crew will respond appropriately to an annunciated PSM.

Design Recommendation 5: The technology should promptly notify flight crew of engine recovery actions and engine status updates.

Some engine systems are capable of employing automated self-recovery actions following an automated diagnosis of certain PSM types, such as a recoverable engine surge (e.g. due to icing). This means that steps (10) to (16b) are conducted entirely automatically in the background. It is important that flight crew are alerted of such automated self-recovery action on the affected engine and status with as little lag as possible, as this would minimise the risk of inappropriate actions such as an unnecessary IFSD, which occurred during the Singapore Airlines 9V-SSF dual engine surge incident on 15 May 2015 (TSIB, 2018).

Design Recommendation 6: The technology should display engine parameters in the way that supports accurate and reliable diagnosis, decision-making, and monitoring of engine health and performance.

This recommendation also addresses activities relating to interpreting cues and engine parameters to diagnose PSM within steps (10a) to (12) and steps (17) to (20), and obtaining feedback from execution of actions on steps (16a) to (16b) and steps (28a) to (28b). The designs of engine parameter indicators have not changed significantly since the 1960s, even as the instruments changed from electromechanical to glass cockpit with LCD screens. Ideally, the designs should be such that the diagnosis activities in steps (10a) – (12) and (17) – (20) are not prone to rule-based errors such as *erroneous recall of data/parameters related to rules*, *adherence to familiar rules* related to sequential observation of information (Rasmussen & Vicente, 1989), and *partial matching of rule* due to informational overload or rule strength (Reason, 1990).

Diagnosis of unannunciated PSM starts at step (10a) or (17), where flight crew assess the state of the powerplant system by scanning primary and secondary engine indications more carefully, and obtain more information from other sources if necessary. CDM interviews revealed that engine parameter indications (see Figure 7 as a typical example) are mainly scanned for differences between engine (asymmetry) and abnormal values (very high or very low), usually in the order that they are displayed from top to bottom. With standard engine instruments, a systematic scan is demanded of the flight crew, especially when there is no specific indication for asymmetry or abnormal values. For some engine parameter indications, other than very low or very high values, the definition of abnormal values may depend on the context of the flight such as phase of flight, or even the aircraft's specific condition. Additionally, relying on comparison of the indications between engines could be ineffective in identifying problems if both engines are malfunctioning in the same way, which is a rare, but possible, event. Since flight crew do not continuously monitor engine parameter indications, it is also possible to miss momentary abnormal changes, for example the fluctuations of N1[3] gauge needle during a surge, and a gradual increase or decrease of a parameter.

[3] N1 value indicates fan speed of a turbofan engine, often used as an indication of the amount of thrust produced by the engine.

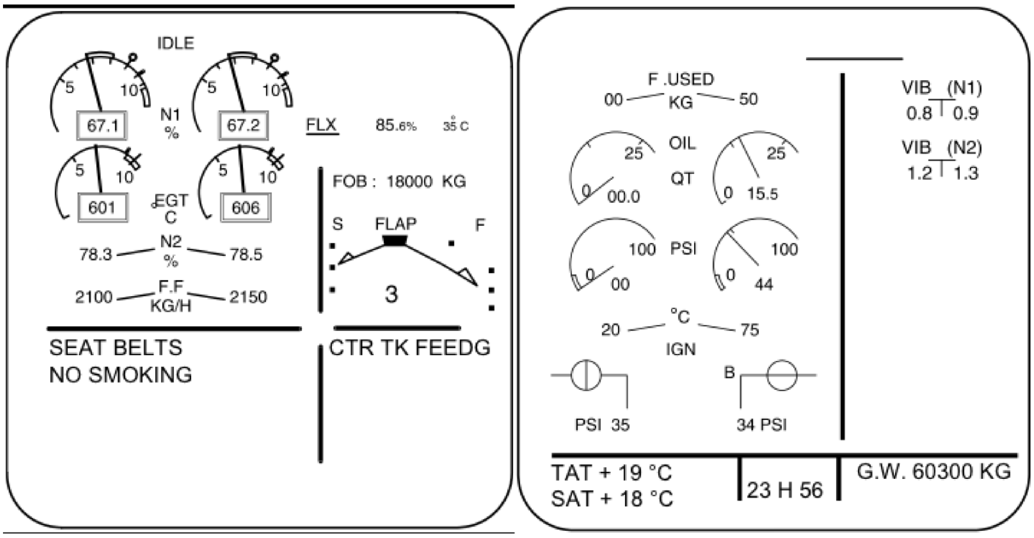

Figure 7 Example of primary engine display (left) and secondary engine display (right) on the Airbus A319/A320/A321 (with CFM engines) on ECAM (Airbus, 1998)

Design recommendations emerging from how flight crew observe information from engine parameter indications include highlighting values and rate of change of those values when they are out of normal range for the particular flight phase or settings. Asymmetrical engine parameter values between engines should also be highlighted when the difference is out of normal range. Conversely, flight crew will also benefit from notification when previously abnormal engine parameters have returned to normal.

The main difference between the two decision ladders in this work situation is the knowledge-based activities involved during an ill-defined PSM event (Figure 6), as denoted by steps (21) to (24). It is important to reiterate that flight crew will only undertake an extensive diagnosis effort or additional information gathering if time and workload level allow. In order to select the best option to achieve the goal at this stage, flight crew will benefit from information on projected engine and aircraft condition and operational status for the remainder of the flight. This will also allow flight crew to make a more accurate risk assessment and therefore potentially better decision into the next stage. There is a huge potential in utilising current engine health monitoring data to provide decision-making support such as this.

3.3 Generic decision ladder for Work Situation 3 (Modifying flight plan)

The fourth generic decision ladder (Figure 8) addresses the third stage of flight crew's response to a non-normal or emergency situation, which is modifying flight plan. A formal diversion plan is usually created before flight, and could be activated by myriad factors including weather conditions, passenger issues, or engineering issues. Changing flight plan is also a well-trained aspect of flight operations, hence the decision ladder contains mainly skill-based and rule-based activities. The recommendation arising from this decision ladder is aimed at synthesising information to assist decision-making and minimise workload.

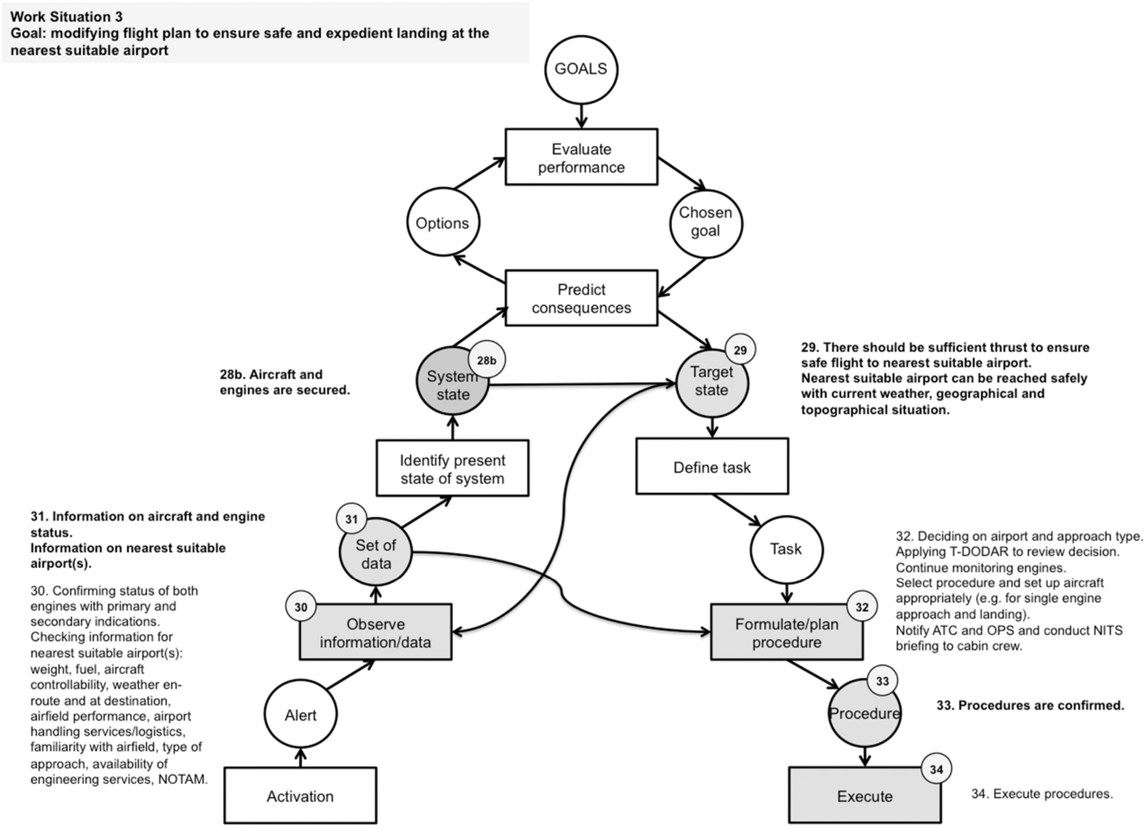

Figure 8 Generic decision ladder for the third stage of responding to a powerplant system malfunctions, representing cognitive states and processes involved in modifying the flight plan and ensuring safe and expedient landing.

Design Recommendation 7: The technology should provide synthesised information required for modifying a flight plan.

In step (30) of the decision ladder, there is a great deal of information that the flight crew need to obtain and process from various sources to plan for the tasks demanded in step (32). Integrated display of this information would reduce the workload and support decision-making and planning tasks at this stage. In addition, multiple alternatives could be shown with corresponding airfield performance data, for example. In this way, flight crew could plan intelligently to minimise disruption.

4. General discussions

We aimed to characterise flight crew decision-making process during PSM using control task analysis with the decision ladder template. This approach is an effective method to characterise the process and then to transparently generate the design recommendations for future systems. Our analysis shows that flight crew response to PSM in the current operating environment can be inappropriate and variable due to various factors. We suggest that a system encompassing our new design recommendations from the decision ladders would address these factors and improve response, minimising variability whilst maintaining a degree of flexibility.

In contrast to Lintern (2010), our study shows that decision ladder template works well to model decision-making in the operational environment. The method is capable of characterising the role of procedures and automation in a complex system. The combination of models and methodologies used in this study provided a structured and reliable methodology in analysing flight crew decision-making in a PSM event. Orasanu and Fischer's (1997) aeronautical decision-making model was useful in classifying PSMs in terms of how the decision making process may differ for each category, and determining the number of generic decision ladders that needed to be developed. In line with findings from Naikar and Saunders (2003), data from the CDM interviews were instrumental in constructing and validating the decision ladders. The CDM interview data was also used to identify areas of difficulties and how they could be improved according to the flight crew. Combining decision ladders with Rasmussen's (1983) SRK error taxonomy assisted further in determining potential errors or variable flight crew responses. According to Rasmussen et al. (1994) and Vicente (1999), shortcuts in the form of shunts and leaps are created by heuristics decision processes, but in this application we show that shunts and leaps are primarily signposted by procedures and in some cases already created by automation itself. Where heuristics, procedures or automation are not available, knowledge-based activities take place. The decision ladders generated in this study identify where shortcuts could be created to reduce the number of cognitive processes and potential for errors or variable responses. It is these shortcuts to which some of our recommendations refer.

Secondly, we aimed to develop design recommendations for new systems, exploiting engine heath monitoring data and other sources of data. In support of this second aim we propose seven new design recommendations based on the decision ladder analysis. A system that meets these design recommendations would support flight crew during both ill-defined and well-defined PSM events. A new, intelligent system using synthesised information that can be generated by engine health monitoring data, aircraft systems and flight performance data, and third-party data such as traffic and weather could be specified to minimise disruption in the event of a PSM, give the flight crew more options or constrain others. Greater focus on interoperability is needed to take these individual systems to a new environment to deliver reliable information that can be exploited by the flight deck. These types of systems may become even more important in the future when considering new operational concepts such as single-pilot operations or autonomous, centrally co-ordinated flights.

One important element of our design recommendations is automation of different functions, evolving the current engine management system. With recent automation-related aviation accidents, the role of automation in aviation is under renewed scrutiny. The selection of automation types in the areas addressed by the proposed design requirements will need careful consideration, particularly because automation can result in new activities and new coordination demands from the flight crew (Pritchett, Kim, & Feigh, 2014). While cognitive work analysis aims to provide flexibility in adapting to disturbances in complex sociotechnical systems (Vicente, 1999), consequences of incorrect flight crew response in this case can potentially be grave and therefore designers need to allow some variability in flight crew response whilst creating a more defined constraint boundary, unless variable response is completely undesirable. For example, automation type in Design Recommendation 1 is *action implementation*, which involves automation in all other functions including *information acquisition*; *information analysis*, and *decision and action selection* (Parasuraman et al., 2000). Airframe manufacturers implement this type of automation in cases where a different response may never be required – for example, a counter rudder to correct the yaw resulting from a single engine failure. However, variability of responses within a defined constraint boundary might still be useful in other areas addressed in the other design recommendations. For example, in the case of engine fire, the aircraft/engine systems provide *information acquisition* and *information analysis* automation in the form of engine fire warning, rather than *action implementation* to automatically extinguish the fire. This is because currently the procedure for engine fire involves shutting down the engine, which may require consideration of other information by the flight crew. The determination of type of automation in the areas addressed by our design recommendations should be investigated further in future research, in which specific contexts such as single pilot operations should also be considered.

Another aspect of our design recommendations is the interface and display of information. Existing automation in aircraft and engine systems have assisted flight crew. However, current interfaces have created problems, as evidenced in the QF32 accident (ATSB, 2013) and SQ836 (TSIB, 2018) incident. With new information generated by the intelligent systems, we focus on displaying the synthesis of information rather than adding pieces of information that require more cognitive processing from the flight crew. This is where the goal for each decision ladder is key to our analysis. The goal should be the driver of how the interface and display should be designed for the particular work function.

Our findings need to put in the context of the study limitations. Firstly, we only consider twin-engine operations in this article. Extension of findings to four, or increasingly rarely, three engine types is not always warranted. In a four engine aircraft, loss of thrust in one engine causes a reduction in total thrust by 25 percent. Although a twin-engine aircraft is capable of flying with one engine, the loss of fifty percent total thrust can create a more immediate need to re-plan effectively especially since the margin of safety is reduced more significantly.

We do not consider phases of flight other than climb and cruise in this article. In other phases of flights such as final approach and landing, the priority is to fly the aircraft and land safely as soon as possible. However, we do not exclude the possibility that new engine health monitoring data might assist these phases of flight in the future.

The application of this method to a new area in commercial air transport operation has resulted in new concepts and user-centred design recommendations. The models of decision-making have characterised and defined this hidden area of aircraft operations. Our analysis has shown ways in which technology could make flight crew action in the event of engine malfunctions even safer, further minimising disruption to crew and passengers alike.

Acknowledgement

This work was supported by Cranfield University Aerospace Integration Research Centre, in collaboration with Rolls-Royce.

References

Australian Transport Safety Bureau (2013). *Investigation Report AO-2010-089: In-flight uncontained engine failure Airbus A380-842, VH-OQA, overhead Batam Island, Indonesia, 4 November 2010.* Canberra: ATSB.

Australian Transport Safety Bureau (2014). *Aviation Research Report AR-2013-002: Powerplant failures in turbofan-powered aircraft.* Canberra: ATSB.

Airbus (1998). *319/A320/A321 Flight deck and systems briefing for pilots, STL 945.7136/97, issue September 1998.*

Airbus (2006). *Flight Operations Briefing Notes – Supplementary Techniques: Handling Engine Malfunctions (FLT_OPS – SUPP_TECH – SEQ 07 – REV 01 – DEC. 2006).*

Cattermole, V., Horberry, T., Burgess-Limerick, R., Wallis, G., and Cloete, S. (2015), 'Using the Critical Decision Method and Decision Ladders to analyse traffic incident management system issues, in *Proceedings of the 2015 Australasian Road Safety Conference 14-16 October, Gold Coast, Australia.*

Elix, B., and Naikar, N. (2008). Designing safe and effective future systems: A new approach for modelling decisions in future systems with Cognitive Work Analysis. *Proceedings of the 8th International Symposium of the Australian Aviation Psychology Association*.

Federal Aviation Administration (2009). *Turbofan Engine Malfunction Recognition and Response Final Report*. Burlington, MA: FAA Engine & Propeller Directorate.

Harris, D. (2011). *Human Performance on the Flight Deck.* CRC Press, Boca Raton.

Hassall, M.E., and Sanderson, P.M. (2014). 'Can the decision ladder framework help inform industry risk assessment processes?', *Ergonomics Australia, 10:3.*

Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998). *Use of the Critical Decision Method to Elicit Expert Knowledge: A Case Study in the Methodology of Cognitive Task Analysis.* Human Factors, 40(2), 254–276. https://doi.org/10.1518/001872098779480442

Jenkins, D. P., Salmon, P. M. Stanton, N. A. Walker, G. H., and Rafferty. L. (2011a). "What Could They Have Been Thinking? How Sociotechnical System Design Influences Cognition: A Case Study of the Stockwell Shooting." *Ergonomics 54 (2)*, pp. 103–119.

Jenkins, D. P., Stanton, N. A., Salmon, P. M., and Walker, G. (2011b). A formative approach to developing synthetic environment fidelity requirements for decision-making training. *Applied Ergonomics. 42 (5)*, pp. 757-769

Jenkins, D. P., Stanton, N. A. Salmon, P. M., Walker, G. H. and Rafferty. L. (2010). "Using the Decision Ladder to Add a Formative Element to naturalistic Decision-Making Research." *International Journal of Human-Computer Interaction 26 (2–3)*, pp. 132–146.

Jenkins, D. P. (2012). Using Cognitive Work Analysis to describe the role of UAVs in Military Operations. Theoretical Issues in Ergonomics Science. 13(3)335-357

Karnofski, K. (2015). 'Cost/Benefit of Advanced Engine Health Monitoring', in *5th GRC Propulsion Controls and Diagnostics Workshop. September 2015*. Seattle, WA: The Boeing Company.

Klein, GA, Calderwood,R and MacGregor, D. 1989. Critical decision method of eliciting knowledge, in IEEE Transactions on Systems, Man and Cybernetics, 19: 462–472.

Li, Y., Burns, C., and Hu, R. (2016). 'Representing stages and levels of automation on a decision ladder: The case of automated financial trading', *Proceedings of the Human Factors and Ergonomics Society*, pp. 328–332.

Lintern, G. (2010). 'A Comparison of the Decision Ladder and the Recognition-Primed Decision Model', *Journal of Cognitive Engineering and Decision Making, 4(4)*, pp. 304–327.

McIlroy, R.C., & Stanton, N.A. (2015) A decision ladder analysis of eco-driving: the first step towards fuel-efficient driving behaviour, *Ergonomics, 58:6,* 866-882, DOI: 10.1080/00140139.2014.997807

Mitchell, M., and Saisi, D., (1987). *Use of Model-Based Qualitative Icons and Adaptive Windows in Workstations for Supervisory Control* IEEE Transactions on Systems, Man and Cybernetics. 17. 573 - 593.

Mulvihill, C.M., Salmon, P.M., Beanland, V., Lenne, M.G., Read, G.J.M, Walker, G.H., and Stanton, N.A. (2016). 'Using the decision ladder to understand road user decision making at actively controlled rail level crossings'. *Applied Ergonomics, 56,* pp. 1-10.

Naikar, N (2010). DSTO-TR-2397: *A Comparison of Decision Ladder Template and the Recognition-Primed Decision Model.* Fishermans Bend, VA: Defence Science and Technology Organisation.

Naikar, N., Moylan, A., and Pearce, B. (2006). Analysing activity in complex systems with cognitive work analysis: Concpets, guidelines, and cases for control task analysis. *Theoretical Issues in Ergonomics Science, 7*(4), pp. 371-394.

Naikar, N. and Saunders, A. (2003). Crossing the boundaries of safety operation: An approach for training technical skills in error management. *Cognition, Technology & Work, 5,* pp. 171-180.

Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 3-20). Westport, CT, US: Ablex Publishing.

Orasanu, J., and Fischer, U. (1997) Finding Decisions in Natural Environments: The View from the Cockpit. In Zsambok, C.E., and Klein. G.A. (eds). *Naturalistic Decision Making*. Mahwah, NJ: Lawrence Erlbaum Associates.
Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000). 'A model for types and levels of human interaction with automation', *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 30*, 286-297.
Pritchett, A. R, Kim S. Y., & Feigh, K. M. (2014) Modeling Human–Automation Function Allocation. *Journal of Cognitive Engineering and Decision Making, 8 (1),* pp. 33-51.
Rasmussen, J. (1974). *The Human Data Processor as a System Component Bits and Pieces of a Model*. Riso: Library of the Danish Atomic Energy Commission.
Rasmussen, J., & Lind, M. (1982). *A Model of Human Decision Making in Complex Systems and its Use for Design of System Control Strategies*. Roskilde: Risø National Laboratory. Risø-M, No. 2349
Rasmussen, J. (1983). 'Skills, Rules and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models', *IEEE Transactions on Systems, Man and Cybernetics, Vol.SMC-13, No. 3*, 257-266.
Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. New York, NJ: North-Holland.
Rasmussen, J., & Vicente, K. (1989). Coping with human errors through system design: implications for ecological interface design, in *International Journal of Man-Machine Studies, 31,* pp. 517-534.
Rasmussen, J., Pejtersen, A. M., Goodstein, L. P. (1994). *Cognitive Systems Engineering*. New York: Wiley-Interscience.
Reason, J. (1990). *Human Error.* Cambridge: Cambridge University Press.
Roelen, A.L.C. and Wever, R. (2002). *NLR-CR-2002-547-PT-1 An Analysis of Flight Crew Response to System Failures, PT-1 Flight Crew Intervention Credit in System Safety Assessment, Phase 1 Report*. Amsterdam: National Aerospace Laboratory NLR.
Sallee, G.P. and Gibbons, D.M. (1998). *AIA/AECMA Project Report on Propulsion System Malfunction Plus Inappropriate Crew Response (PSM + ICR)*.
Transport Safety Investigation Bureau. (2018). *Report No. AIB/AAI/CAS.112 - Final Report A330-300, Registration 9V-SSF Dual Engine Surge 23 May 2015.* Singapore: Ministry of Transport Singapore.
Ulizar, I., Azkoitia, I., and Minguez, B. (2016). Aircraft Engine Advanced Health Management: The Power of the Foresee, in *8th European Workshop on Structural Health Monitoring (EWSHM 2016), 5-8 July 2016, Spain.*
Vicente, K.J. (1999). *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Vicente, K. J., & Tanabe, F. (1993). *Event-independent assessment of operator information requirements: Providing support for unanticipated events.* In Proceedings of the American Nuclear Society Topical Meeting on Nuclear Plant Instrumentation, Control and ManMachine Interface Technologies (pp. 389-393). La Grange Park, IL: ANS.
Weitzenfeld, J.S., Freeman, J.T., Riedl, T.R., and Klein, G.A., (1990). The critical decision method (CDM): A knowledge-mapping technique. In *Proceedings of the '90 Annual AT&T Human Factors & Behavioural Science Symposium.* AT&T Bell Labs, Holmdel, NJ.
Woods, D. D. (1991). The cognitive engineering of problem representations. In G. R. S. Weir and J. L. Alty (Eds.), Human-computer interaction and complex systems (pp. 169-188). London: Academic Press.

Modelling and supporting flight crew decision-making during aircraft engine malfunctions: developing design recommendations from cognitive work analysis

Asmayawati, Saryani