Cogn Tech Work (2003) 5: 171–180
DOI 10.1007/s10111-003-0125-z

EAM-2002 AWARD WINNERS

Neelam Naikar · Alyson Saunders

Crossing the boundaries of safe operation: An approach for training technical skills in error management

Received: 17 October 2002 / Accepted: 15 April 2003 /

Abstract Robust error management within the cockpit is crucial to aviation safety. Crew resource management (CRM) focuses on *non-technical* skills for error management but the training of *technical skills* for error detection and error recovery is also a potentially valuable strategy. We propose a theoretical basis for training technical skills in error management as well as a cognitively oriented technique for analysing accidents and incidents to identify specific training requirements. To evaluate the strengths and limitations of this new approach, we present a case study of its application to the F-111, a strike aircraft in the Royal Australian Air Force. This case study demonstrates that the new training approach is both feasible and useful, although an empirical validation of the approach is still necessary. In addition, the case study highlights the limitations of the current F-111 simulator for training technical skills for error detection and error recovery.

Keywords Aviation · Safety · Training · Error management · Cognitive work analysis

1 Introduction

Initial efforts to improve aviation safety were based on the view that humans are unreliable and that safety interventions should focus on preventing human error. Over the last decade, there has been an increasing recognition that errors are not only difficult to eliminate completely but that they are a consequence of the same cognitive mechanisms that allow humans to operate flexibly under demanding conditions. Moreover, although humans often make errors, they also provide a critical line of defence in averting the adverse consequences of errors because of their ability to adapt to dynamic situations. The focus of safety efforts has therefore shifted from error prevention to error management; that is, creating systems that are better able to tolerate the occurrence of errors and contain their damaging effects (Hollnagel 1993; Maurino 2001; Paries and Amalberti 2000; Rasmussen and Vicente 1989; Reason 2000 and 2001; Sarter and Alexander 2000; Shappell and Wiegmann 2000).

In this paper, we propose a new approach for training *technical* skills for error management. This approach was motivated by three main observations. Firstly, the training of technical skills in aviation emphasises error prevention. Yet, errors still occur. Second, crew resource management (CRM) programs provide aircrew with *non-technical* skills for error management. Yet, accidents still occur. Third, although poor CRM is a causal factor in many aircraft accidents, poor technical skills in operating the aircraft after it has been placed in an unsafe or undesired state by human error are also often a factor (Naikar, Saunders and Hopkins 2002). Before we present the new approach for training technical skills for error management, we elaborate the concept of error management and explain existing approaches to error management in light of this concept.

1.1 Error management

The concept of error management can be illustrated by characterising work systems as having boundaries of safe operation (Flach and Rasmussen 2000; Rasmussen, Pejtersen and Goodstein 1994). Using this characterisation, a system is in a safe or desired state when it is operating inside the boundaries of safe operation (Fig. 1). When a system crosses the boundaries of safe operation, it is in an unsafe or undesired state and an accident or incident can result. This is inevitable in competitive work environments where the pressure to achieve system objectives within tight resources and operating constraints leads to systematic migration

N. Naikar (✉) · A. Saunders
Defence Science and Technology Organisation,
P.O. Box 4331, VIC 3001 Melbourne, Australia
E-mail: neelam.naikar@dsto.defence.gov.au

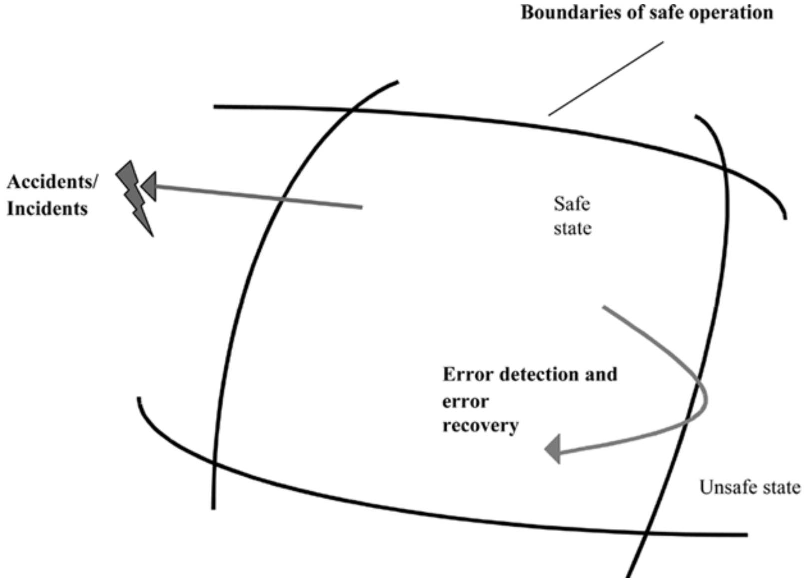


Fig. 1 The characterisation of work systems as having boundaries of safe operation

towards the boundaries. Safety interventions must therefore focus not only on preventing errors to keep a system within the boundaries of safe operation but also on managing errors by making it possible for a system to return from crossing the boundaries of safe operation. Therefore, in terms of aviation, aircrew must be able to detect that an aircraft is in an unsafe state and to recover the aircraft to a safe state.

Rather than there being a tangible boundary between a safe and an unsafe state, an alternative view may be to think of a dimension from safe to unsafe states with migration along this dimension making an accident more likely. This characterisation avoids the complication of pinpointing where the boundaries are. However, the original characterisation by Rasmussen is adequate for the purposes of this paper.

1.2 Existing approaches to error management

An existing approach to error management is error-tolerant design (Billings 1997; Kontogiannis 1999; Noyes 1998; Rasmussen and Vicente 1989). A key strategy of this approach is to design interfaces that make errors and their effects visible and reversible. Error-tolerant interfaces, therefore, highlight the boundaries of safe operation, present feedback when the system crosses the boundaries and provide opportunities for operators to return the system to a desired state.

Another existing approach to error management is CRM. The latest generation CRM relies on observations of normal line flights to pinpoint areas for organisational improvement and to develop training strategies for error management (Gunther 2002; Helmreich 2001; Helmreich, Klinect and Wilhelm 1999; Helmreich, Merritt and Wilhelm 1999). Forms of organisational improvement include: modifying standard operating procedures, changing the nature and scope of technical training, altering scheduling practices and establishing or enhancing safety departments. These interventions are aimed at preventing error or minimising the likelihood of error. CRM strategies for error management involve making the flight crew aware of sources of threats to flight safety and training them in core behaviours for managing errors effectively, such as maintaining vigilance and cross-checking team members, reviewing and modifying plans and leadership and communication skills. CRM therefore trains aircrew in non-technical skills for detecting unsafe states and recovering the aircraft to a safe state.

2 Training technical skills for error management

While CRM programs provide training in non-technical skills for error management, the training of technical skills in aviation is still focused on error prevention. A technical training approach that is based solely on error prevention gives aircrew the opportunity to practise error-free performances and therefore to operate the aircraft within the boundaries of safe operation. Consequently, aircrew have considerable experience in operating the aircraft when it is in a safe state. However, aircrew have little opportunity to operate the aircraft when it has crossed the boundaries of safe operation into an unsafe state. As a result, they get little practice in detecting cues that the aircraft has been placed in an unsafe state by human error and recovering the aircraft to a safe state. The typical response when an aircraft enters an unsafe state is often confusion as aircrew seek to understand what is happening (Naikar, Saunders and Hopkins 2002). Recovery action is usually straightforward if the aircrew understands the situation.

To develop systems that are more prone to error, an approach for training aircrew in technical skills for error management may be necessary. The training approach we propose would give aircrew the opportunity to cross the boundaries of safe operation so that they can practise detecting the cues that the aircraft is in an unsafe state and strategies for recovering the aircraft to a safe state. Then, if aircrew cross the boundaries of safe operation during real missions, they are more likely to detect and recover from the error, and consequently avert an accident or incident. This training approach generally requires a simulator because it would be dangerous to cross the boundaries of safe operation in a real aircraft.

The technical training approach to error management that we propose can lead to novel ways of training. Consider, for example, the training of procedures in both commercial and military aviation. The most common approach is to drill aircrew in executing the steps of a procedure to minimise the likelihood of error during real operations. However, a slip or lapse in executing procedures is inevitable, as any incident database will show. Adopting a technical training approach to error management would require that aircrew are also given practice in dealing with the evolving situation if they make an error in executing a procedure. Therefore, aircrew should be given the opportunity to not follow a procedure or parts of a procedure in a training simulator and to practise detecting and recovering from the error.

Some empirical support for this approach is available from studies in the area of human-computer interaction (Dormann and Frese 1994; Frese and Altmann 1989; Frese, Brodbeck, Heinbokel, Mooser, Schleiffenbaum and Thiemann 1991). These studies showed that trainees who were encouraged to make errors while learning computer programs perform better on test tasks than do trainees who were required to follow procedures or instructions. Two of the explanations offered for the superior performance of trainees who were encouraged to make errors were that they learn strategies for recognising errors and dealing with the resulting situations, and that they develop more knowledge of the system by experiencing it in phases that are not normally present during error-free performance. The differences between the two groups were more pronounced with difficult tasks than with easy tasks, presumably because greater knowledge is necessary to solve difficult tasks and errors are more likely to occur.

The technical training approach to error management that we propose is also consistent with findings about expert decision making in high-risk operations (Klein 1997; Klein, Calderwood and MacGregor 1989). Studies of fire fighters, military commanders and paramedics have shown that in time critical, high-workload situations, experts can use their prior experience to make rapid and effective decisions by matching situations to ‘templates’ of cues, diagnoses and solutions that have worked in the past. In most of the domains, between ninety and ninety-five percent of the decisions were made in this way. These studies demonstrate that prior experience provides an effective basis for recognising situations and implementing plans and actions under demanding circumstances.

Finally, certain aspects of training in the Royal Australian Air Force suggest that the technical training approach to error management that we propose may have validity in military aviation. Firstly, the approach is consistent with that taken for training aircrew to manage equipment failures. Specifically, equipment failures are presented to aircrew during simulator training so that they can practise processes for managing the failures if they occur on real missions. Secondly, in some cases, the aircrew receives training in detecting and recovering from human error, although usually not to the extent advocated by the approach that we propose. For example, the aircrew is required to place the aircraft in an unusual attitude and to practise recovering from this situation. However, to recover from an unusual attitude in real operations it is critical to detect the unusual attitude in the first place. Therefore, aircrew should also receive training in detecting unusual attitudes.

In other cases, aircrew receive limited training in error detection. For example, a training instructor acting as a navigator in a two-person strike aircraft may deliberately enter a wrong weapons delivery mode to check if the pilot detects the error. If the pilot does not detect the wrong entry, the training instructor will usually alert the pilot to the error and correct the weapons delivery mode prior to the attack on the target. However, since failing to detect a wrong entry can also happen on real missions, the training instructor should leave the error uncorrected and give the pilot more of an opportunity to learn to recognise the cues that an error has occurred and to practise recovering from the evolving situation.

3 Techniques for identifying training requirements

To train aircrew in technical skills for error management, it is necessary to identify both potential errors and strategies to detect and recover from the errors, as well as to incorporate the strategies into training programs. Therefore, in addition to a theoretical basis for training technical skills in error management, we have developed a technique to identify specific training requirements. This technique involves analysing accidents and incidents to examine the boundaries of safe operation that the aircrew has crossed and their error, error detection and error recovery processes. Subsequently, this analysis can be used to develop requirements for the unsafe states that an aircrew should experience during training and the strategies that they should practise for error detection and error recovery.

The technique that we have developed consists of three main steps. The first step is to identify the critical events in an accident or incident mission. The second step is to use Cognitive Work Analysis (CWA) (Rasmussen et al. 1994; Vicente 1999), in particular the decision ladder or step ladder model, to examine an aircrew’s error, error detection and error recovery processes during the critical events. The third step is to identify training requirements for technical skills in error management on the basis of the preceding analysis. In the following sections we illustrate the three steps of the technique with a hypothetical F-111 accident.[1]

3.1 Identifying critical events

The first step of the technique is to identify the critical events in an accident or incident mission. Critical events are points during a mission when the aircrew crossed the boundaries of safe operation, and/or the aircrew had the opportunity to detect that the aircraft was in an unsafe state, and/or the aircrew had the opportunity to return the aircraft to a safe state. Therefore, to identify critical events, analysts should search for points during the mission when:

1. The actions or decisions of aircrew, or the absence of actions or decisions, placed the aircraft in an unsafe state
2. Information was available to aircrew that the aircraft was in an unsafe state
3. Opportunities were available to the aircrew to recover the aircraft to a safe state

To illustrate these points, consider a hypothetical accident involving an F-111 aircraft in which a pilot tries to perform a manoeuvre manually without disengaging the autopilot. The pilot finds it difficult to perform the manoeuvre because the autopilot is fighting him for control of the aircraft, but the pilot perseveres with completing the manoeuvre. The autopilot has a bank limit on it, so when the pilot achieves a 45° bank angle during the manoeuvre, the autopilot produces a fail tone and then disengages from controlling the aircraft. As the autopilot disengages while the pilot is exerting high stick forces to gain control of the aircraft and complete the manoeuvre, the aircraft is thrown into a hazardous state and then hits the ground.

In this accident, the first critical event occurred when the pilot tried to perform a manoeuvre manually without disengaging the autopilot. This action placed the aircraft in an unsafe state where the pilot and autopilot were fighting for control of the aircraft. The pilot had cues that the aircraft was in an unsafe state because he experienced difficulty in executing the manoeuvre against the control inputs of the autopilot. At this point the pilot also had the opportunity to recover the aircraft to a safe state by disengaging the autopilot. The second critical event occurred when the autopilot disengaged and produced a fail tone, which was another cue that the aircraft was in an unsafe state. At this point the pilot also had the opportunity to recover the aircraft to a safe state by reducing stick forces while the autopilot disengaged, and then taking manual control of the aircraft.

3.2 Examining the aircrew's error, error detection and error recovery

The second step of the technique is to consider possible explanations for the aircrew's error, error detection and error recovery processes during each critical event using the decision ladder (Rasmussen et al. 1994; Vicente 1999). We chose the decision ladder over more traditional models of information processing (Norman 1986; Wickens 1992) since the steps in the decision ladder need not be followed in a linear sequence, the decision ladder accommodates many starting points and the decision ladder accommodates shortcuts or shunts and leaps from one part of the model to another (see Vicente 1999 for a detailed discussion of these points). This model was motivated by observations that experts rarely follow the strict linear sequence from perception to execution that characterises traditional models of information processing (Rasmussen 1974). The decision ladder has previously been used in accident analysis for classifying errors (O'Hare, Wiggins, Batt and Morrison 1994; Rasmussen 1982).

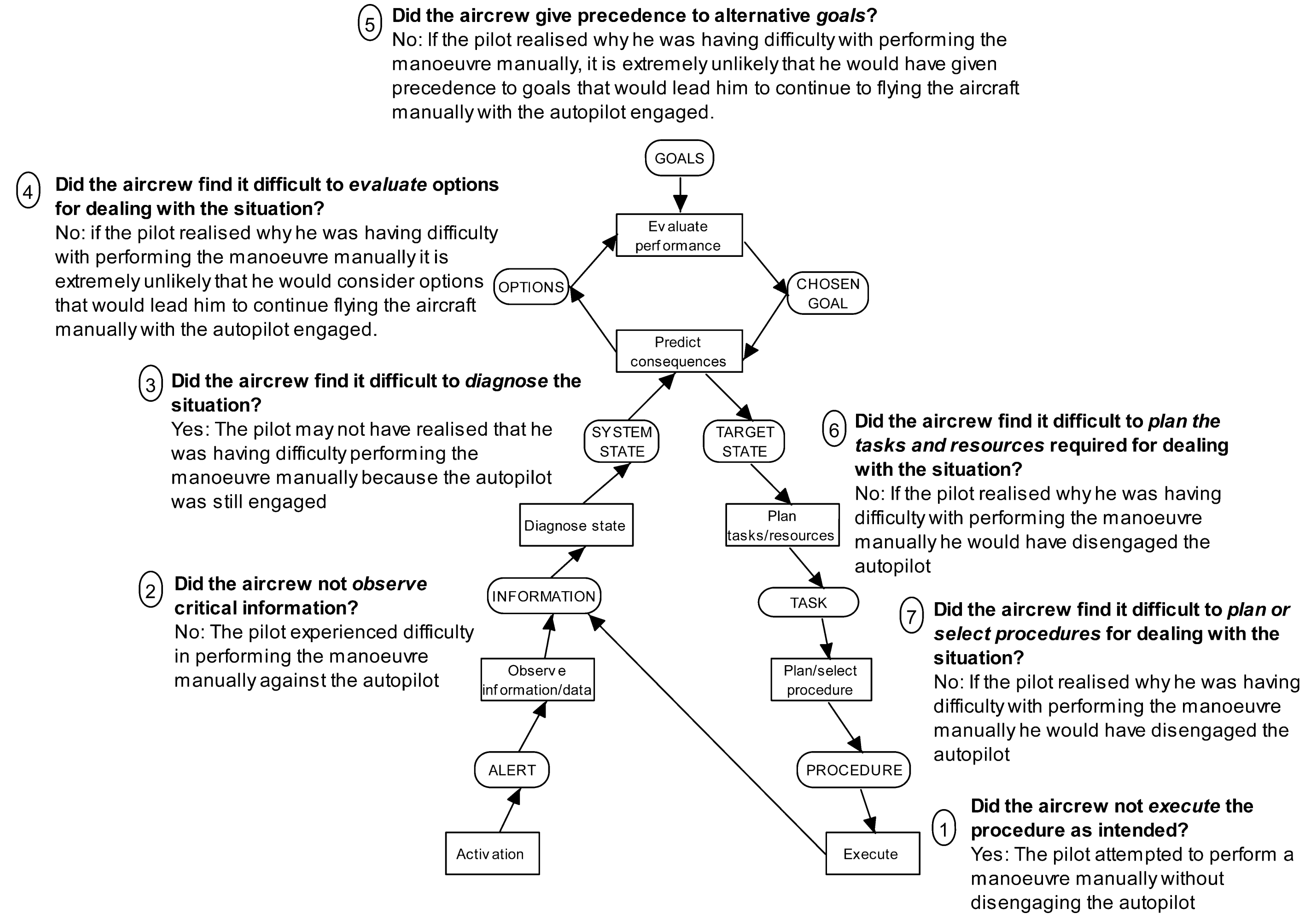
Using the decision ladder (Fig. 2), the aircrew's error, error detection and error recovery processes during each critical event are modelled in terms of the following: observation, diagnosis, option evaluation, goal prioritisation, planning and execution. Table 1 shows a sample of prompts relating to these different parts of the decision ladder. The cells that are filled in the last two columns indicate that error detection processes are related to the left side of the decision ladder, whereas error recovery processes are related to the top and right side of the decision ladder. By using these prompts to review accident or incident data, analysts can analyse which aircrew errors placed the aircraft in an unsafe state, why the aircrew did not detect or recover from the error or how the aircrew detected and recovered from the error.

Specifically, at each critical event, analysts should first use the prompts in the second column of Table 1 to analyse what aircrew errors placed the aircraft in an unsafe state. Then, if the aircrew did not detect the error or recover from the error, the analysts should use the same prompts to analyse why they did not detect the error or recover from the error. If the aircrew did detect and recover from the error, the analysts should use the prompts in the third and fourth columns, respectively, to analyse how they did this. When using the prompts to model these processes, it is important to avoid the hindsight bias (Dekker 2002; Woods, Johannesen, Cook and Sarter 1994) and rely solely on information that was available to the aircrew at that time, rather than bringing in knowledge about the accident that is available to the analysts after the fact.

To illustrate, at the first critical event in the hypothetical accident, the aim is to analyse the pilot error that placed the aircraft in an unsafe state and then to examine why the pilot did not detect and recover from the error at this point. The decision ladder prompts to use to analyse the critical event are shown in the second column of Table 1. Figure 2 shows a decision ladder for

---

[1] We use a hypothetical accident because information about actual F-111 accidents is classified.



⑤ **Did the aircrew give precedence to alternative *goals*?**
No: If the pilot realised why he was having difficulty with performing the manoeuvre manually, it is extremely unlikely that he would have given precedence to goals that would lead him to continue to flying the aircraft manually with the autopilot engaged.

④ **Did the aircrew find it difficult to *evaluate* options for dealing with the situation?**
No: if the pilot realised why he was having difficulty with performing the manoeuvre manually it is extremely unlikely that he would consider options that would lead him to continue flying the aircraft manually with the autopilot engaged.

③ **Did the aircrew find it difficult to *diagnose* the situation?**
Yes: The pilot may not have realised that he was having difficulty performing the manoeuvre manually because the autopilot was still engaged

⑥ **Did the aircrew find it difficult to *plan the tasks and resources* required for dealing with the situation?**
No: If the pilot realised why he was having difficulty with performing the manoeuvre manually he would have disengaged the autopilot

② **Did the aircrew not *observe* critical information?**
No: The pilot experienced difficulty in performing the manoeuvre manually against the autopilot

⑦ **Did the aircrew find it difficult to *plan or select procedures* for dealing with the situation?**
No: If the pilot realised why he was having difficulty with performing the manoeuvre manually he would have disengaged the autopilot

① **Did the aircrew not *execute* the procedure as intended?**
Yes: The pilot attempted to perform a manoeuvre manually without disengaging the autopilot

Fig. 2 A decision ladder for a hypothetical accident

this critical event. The annotations to the decision ladder describe the responses to the prompts with the numbers indicating the order in which the decision ladder should be followed. From the decision ladder, we can see that the pilot did not execute a procedure correctly. That is, he tried to perform a manoeuvre manually without disengaging the autopilot. The pilot then had difficulty in detecting the error he had made. This was not because he did not *observe* the critical information (he was aware that it was hard to perform the manoeuvre) but rather because he was unable to *diagnose* why it was hard to perform the manoeuvre. If the pilot had made the right diagnosis, he would probably have detected the error (that is, he would have realised that he had forgotten to disengage the autopilot) and consequently, he would have recovered from the error (that is, disengaged the autopilot while reducing stick forces).

3.3 Defining training requirements

The third step of the technique is to define training requirements for technical skills in error detection and error recovery from the preceding analysis. In particular, the elements of the preceding analysis that are relevant here include: the aircrew's errors, the unsafe state that the aircraft was placed in by the aircrew's errors, the cues that were available to the aircrew that the aircraft was in an unsafe state, the problems that the aircrew had in observing the cues or diagnosing the situation, the options that were available to the aircrew to recover the aircraft to a safe state and the problems that the aircrew had in evaluating options, prioritising goals, planning and executing procedures to return the aircraft to a safe state. This information can be used to specify training requirements in terms of the unsafe states that aircrew should experience during training and the strategies that they should practise for error detection and error recovery.

The training requirement from the analysis of the hypothetical accident is to give aircrew the opportunity to fly the aircraft manually with the autopilot engaged so that they can experience the difficulty in performing a manoeuvre. In addition, aircrew should practise disengaging the autopilot while exerting appropriate stick forces so that the aircraft is not thrown into a hazardous state when the autopilot relinquishes control of the aircraft. Then, if the aircrew forgets to disengage the autopilot before performing a manoeuvre manually on a real mission, they are more likely to recognise the cues that the autopilot is still engaged and disengage the autopilot to regain control of the aircraft without throwing the aircraft into a hazardous state.

Table 1 A sample of prompts for developing decision ladder models of the aircrew's error, the error detection and the error recovery processes

Parts of the decision ladder	What was the error? Why didn't the aircrew detect the error? Why didn't the aircrew recover from the error?	How did the aircrew detect the error?	How did the aircrew recover from the error?
Observation	Did the aircrew not observe critical information?	What information or cues did the aircrew observe?	
Diagnosis	Did the aircrew find it difficult to diagnose the situation?	What was the aircrew's diagnosis of the situation?	
Option evaluation	Did the aircrew find it difficult to evaluate options?		What options did the aircrew consider? Why did the aircrew select/reject options?
Prioritisation of goals	Did the aircrew give precedence to alternative goals?		What were the aircrew's goals?
Planning of tasks and resources	Did the aircrew find it difficult to plan the tasks and resources required for dealing with the situation?		What tasks and resources did the aircrew plan to use to recover from the error?
Planning or selection of procedures	Did the aircrew find it difficult to plan or select procedures for dealing with the situation?		What procedure did the aircrew plan to recover from the error? What standard procedure did the aircrew select to recover from the error?
Execution	Did the aircrew not execute the procedure as intended?		What procedure did the aircrew execute?

4 Applying the technique to the F-111

Having proposed a theoretical basis for training technical skills in error management and a new technique for identifying specific training requirements, we now offer a case study that shows how we applied this technique to the F-111, a strike aircraft in the Royal Australian Air Force. This case study allows us to evaluate the strengths and limitations of the new training approach. We begin by providing some information about the F-111.

The F-111 is a strike aircraft that is manned by a pilot and a navigator. F-111 missions typically involve flying in formation to a series of waypoints to deliver weapons to land and maritime targets. Weapons must be delivered to the targets on time, while evading threats to the platform, and with ingress to the targets occurring at a low level—approximately 200 feet from the ground. These missions are characterised by a high aircrew workload and tight time constraints (Lintern and Naikar 2000). Over the last 32 years, there have been nine F-111 accidents, six of these fatal. In addition, 124 incidents have been entered into the Royal Australian Air Force's incident database over the last two years.

5 The F-111 accident analysis

To identify F-111 technical training requirements for error management, we used the technique that we developed to analyse the three most recent F-111 accidents which occurred in 1987, 1993 and 1999. All three accidents were fatal. To perform the analyses, we developed profiles of the accident missions based on data in the reports of the accident investigations and boards of inquiry. In these profiles we listed the various phases of the accident mission, for example, the ingress to the target, the pull up and weapons release and the turn execution. Then, for each mission phase, we described the duration of the phase, the crew communications and control actions, the system alerts, the aircraft status including the altitude, the bank angle, pitch, etc., the status of the displays including radar images, the standard operational procedure at that phase and the alternative explanations for the aircrew's behaviour or for the aircraft's departure from the standard operational procedure at that phase.

Typically, it took between three to five days to construct a profile of each accident mission. From these profiles we identified the critical events in the accident missions and developed decision ladders for each critical event. It took approximately one day to identify the critical events and two days to develop decision ladders. From this analysis it took approximately one day to identify training requirements. The analysis of the three F-111 accidents resulted in six training requirements.

5.1 The F-111 incident analysis

We also used the technique that we developed to analyse F-111 incidents that were reported in the Royal Australian Air Force's incident database between

January 2000 and June 2002. We eliminated incidents that reported equipment malfunctions and incidents that involved human error outside the aircraft while it was still on the ground. We were left with 21 incidents that reported aircrew errors inside the cockpit of the aircraft.

A major difference in analysing incidents, compared to fatal accidents, is that the incident aircrew can provide important information about how they detected and recovered from the error. We found, though, that the incident reports are typically very brief, and limited to describing the sequence of events in the incident. For instance, aircrew may report that after performing the weapons checks they suddenly realised that the aircraft had descended inadvertently. However, they generally do not describe the cues that alerted them to the situation. This is not surprising because this type of implicit knowledge is usually difficult to articulate (Klein et al.1989).

To elicit information about the aircrew's errors and their error detection and error recovery processes, we interviewed the aircrew about the incidents they had reported using the critical decision method (Hoffman, Crandall and Shadbolt 1998; Klein et al. 1989). This technique allows interviewers to gradually shift operators from an operational description of the incident to a description of the cognitive processes behind the incident. This is typically done in four 'sweeps'. In the first sweep, the operator recounts the whole incident in their own terms with minimal interruption from the interviewer. In the second sweep, the operator and the interviewer develop a more detailed account of the incident including the sequence of events. In the third sweep, the operator and the interviewer establish a timeline for the incident and identify the decision points. In the final sweep, the interviewer uses a number of probes to elicit information about the operator's decision-making processes during the decision points.

For the purposes of this study, some tailoring of this interview procedure occurred at sweeps 3 and 4 (Naikar and Saunders 2002). Specifically, for each incident, the interviewer targeted at least three decision points: (1) the error(s) (2) error detection (3) error recovery. The interviewer used general probes to prompt a free recall of the aircrew's experiences at each decision point, followed by specific probes where necessary for constructing decision ladders. So, for example, Table 2 shows that for error detection, a general probe is, "How did you know that something was wrong?", whereas a specific probe is, "What did you see, hear or smell that alerted you that something was wrong?" Some of the specific probes are similar to those described in Klein et al. (1989) and Hoffman et al. (1998). Finally, after probing all of the critical events with the general and specific probes, the 'in hindsight' probes in Table 2 may be useful for uncovering further information to define training requirements for error management.

It took approximately one hour to conduct each interview and approximately one day to transcribe each interview. Analysing the interview data with the technique that we developed took approximately two days. The analysis of 21 F-111 incidents resulted in 8 training requirements. Several of the incidents involved the same type of error, for example, setting a particular autopilot parameter incorrectly.

5.2 An evaluation of the training requirements

To evaluate the training requirements, we interviewed seven F-111 aircrew and seven F-111 instructors about the training requirements from the accident analysis, and six F-111 aircrew and six F-111 instructors about the training requirements from the incident analysis.[2]

For each training requirement, we asked the interviewees the following questions:

1. whether they had already conducted the training
2. whether they thought that the training would be useful for helping them deal with errors on real missions
3. whether they had been in an unsafe situation similar to the accident or incident that had motivated the training requirement

Table 3 shows the number of "yes" responses and the total number of responses for each training requirement. The first six training requirements were from the F-111 accidents, whereas the remaining eight training requirements were from the F-111 incidents. The total number of responses for the training requirements from the F-111 accidents varies between ten and fourteen because in a few cases it was difficult to assign a categorical "yes" or "no" to the responses in the interview transcripts. During the interviews for the training requirements from the F-111 incidents, which were held on a later occasion, we asked interviewees for a categorical "yes" or "no" answer if we weren't sure about their responses.

The table shows that for each training requirement the majority of interviewees responded that they did not conduct the training already. For those responses that are coded as "yes", the interviewees generally reported that they had been exposed to the training once before when the opportunity arose through aircrew error. None of the training requirements were conducted systematically or were documented in the F-111 training syllabus.

For each training requirement, the majority of interviewees responded that they thought the training would be useful for helping the aircrew deal with errors on real missions. A few of the interviewees judged that a training requirement was not useful because they believed that they were unlikely to make the error that was the basis of the training requirement. However, for each training requirement, several interviewees reported that they had been in an unsafe situation similar to the

---

[2] We are not able to describe the F-111 training requirements in this paper due to the classified nature of this material.

Table 2 A sample of general, specific and 'in hindsight' probes for interviewing aircrew about error(s), error detection, and error recovery. The specific probes are organised according to the parts of the decision ladder

		Error(s)	Error detection	Error recovery
General probes		What went wrong?	How did you know that something was wrong?	How did you respond to the situation?
Specific probes	Observation	What information did you have at this point? Did you miss anything important?	What did you see, hear, or smell that alerted you that something was wrong?	
	Diagnosis	What was your assessment of the situation at this point? Did you find it difficult to understand what was going on?	What did you think had happened at this point? Did you know what had gone wrong?	
	Evaluation of options	What options, if any, did you consider? Did you find it difficult to select/reject options?		Did you consider several options? Why did you select/reject options?
	Prioritisation of goals	What were your specific goals at this point? Did you find it difficult to prioritise your goals?		What were your goals at this point? How did you prioritise your goals?
	Planning of tasks and resources	Did you have a plan at this point? What was your plan? Did you have any difficulty developing a plan for dealing with the situation?		Did you develop a plan for dealing with the situation? What was your plan?
	Planning and selection of procedures	Were there standard procedures for dealing with the situation? If not, how did you work out what steps to take? Did you have any difficulty formulating this procedure?		Did you use established procedures to recover from the situation? If not, how did you work out what steps to take?
	Execution	What actions did you carry out? Did you have any difficulty carrying out these actions?		What actions did you carry out?
In hindsight probes		In hindsight, is there anything you could have done to prevent the error from occurring and, if so, what? In hindsight, why do you think the error occurred?	In hindsight, could you have detected the error earlier and, if so, how?	In hindsight, could you have recovered more effectively from the error and, if so, how?

Table 3 The number of interviewees that responded 'yes' to the interview questions against the total number of responses for each training requirement

Training requirements	Already done in training? Number responding 'yes' (Total responses)	Useful? Number responding 'yes' (Total responses)	Similar situation? Number responding 'yes' (Total responses)
1	0 (10)	10 (10)	8 (10)
2	1 (14)	13 (14)	7 (14)
3	0 (13)	14 (14)	8 (13)
4	1 (13)	12 (13)	4 (13)
5	4 (14)	14 (14)	13 (13)
6	5 (12)	11 (11)	9 (12)
7	1 (12)	12 (12)	6 (12)
8	1 (12)	11 (12)	8 (12)
9	0 (12)	12 (12)	11 (12)
10	0 (12)	12 (12)	5 (12)
11	2 (12)	11 (12)	10 (12)
12	1 (12)	12 (12)	4 (12)
13	0 (12)	12 (12)	9 (12)
14	0 (12)	12 (12)	5 (12)

associated accident or incident, suggesting that the errors may not be so rare.

The interviewees had some general concerns about the implementation of the training requirements. Some of the interviewees were concerned that if an aircrew deliberately make errors during training they are more likely to make errors on real missions. However, the interviewees were receptive to the idea of instructors, rather than aircrew, deliberately making errors during training so that aircrew could experience the unsafe states. Moreover, for more than half of the training requirements, it is not necessary that either the aircrew or the instructors make errors during mission training to enter the unsafe states. For example, the training requirement from the hypothetical accident was for the aircrew to fly the aircraft manually with the autopilot engaged so that they can experience the difficulty in performing a manoeuvre. Rather than 'forgetting' to disengage the autopilot during mission training for a particular weapons profile, aircrew could participate in a simulator demonstration where they 'jump in' and fly the aircraft with the autopilot engaged and practise disengaging the autopilot without throwing the aircraft into a hazardous state. The interviewees responded favourably to the suggestion of focused simulator demonstrations for implementing the training requirements.

The interviewees were also concerned about practical issues such as the potentially infinite number of errors to train for and the time and resource implications of this. In addition, they raised the issue that if the errors were very rare then it may not be worth the resources to train for those errors. These concerns highlight the need for prioritising the training requirements, perhaps based on the frequency of the errors and the seriousness of the errors in terms of potential consequences. Assessing the frequency of errors may not be as simple as counting incident reports, however, because we obtained many more accounts of similar errors during our interviews than there were incident reports.

The biggest challenge that we identified from the interviews was that the F-111 simulator does not have the capability to enter or simulate some of the unsafe states that were identified by the training requirements. To illustrate, the training requirement from the hypothetical accident was to fly the aircraft with the autopilot engaged. However, the simulator does not replicate the aircraft's performance in this state. This is perhaps not surprising, given that technical training to date has focused on avoiding errors and keeping the aircraft within the boundaries of safe operation. A potential long-term solution is to document the training requirements for future simulator acquisitions.

Meanwhile, another option may be to explore alternatives to simulator-based training. Reason (2001), for example, found that the ability of surgical teams to deal with adverse events depended in part on the extent to which they had mentally rehearsed the detection and recovery of their errors. Therefore, it may be possible to use mental rehearsal techniques, perhaps with the aid of PC-based visualisation tools, to give an aircrew 'experience' in crossing the boundaries and in detecting and recovering from their errors. Another option is case-based learning. O'Hare and Wiggins (2002) reported that written materials and desktop simulations were important sources of remembered cases in pilot decision making. The training requirements could form the basis for cases that illustrate the boundaries that an aircrew may cross and the processes for error detection and error recovery.

6 Conclusions

In this paper we have proposed a theoretical basis for training an aircrew in technical skills for error detection and error recovery and a technique for analysing accidents and incidents to identify specific training requirements. By applying the approach to the F-111 we have demonstrated that the approach appears to be both feasible and useful. The case study also highlighted a number of concerns regarding the implementation of training requirements including the lack of simulation capability for training outside the boundaries of safe operation. To develop systems that are more prone to error, the design of training simulators and the training of technical skills must move beyond error prevention.

Finally, we acknowledge the need for an empirical validation of the technical training approach to error management. As this approach has been well received by the F-111 community, as well as other parts of the Royal Australian Air Force, some of the F-111 training requirements will be trialled in the near future. Practical issues such as time and resource limitations and the availability of F-111 aircrew will dictate whether it is possible to conduct a field experiment. The findings that we have obtained so far with F-111 are encouraging and provide motivation for pursuing further research in this area.

Acknowledgements We would like to thank the following people: the Air Combat Group of the Royal Australian Air Force for sponsoring this work, the Directorate of Flying Safety and 82 Wing of the Royal Australian Air Force for their support, the F-111 aircrew and the training instructors of 1 and 6 Squadrons for giving up their valuable time for interviews, Lee Horsington, Dominic Drumm, Julia Clancy and Robyn Hopcroft from the Defence Science and Technology Organisation for their assistance on this project, Gary Klein and Laura Militello of Klein Associates for their advice on the critical decision method and Gavan Lintern from Aptima, Inc. and Russell Martin from the Defence Science and Technology Organisation for their comments on this paper.

References

Billings CE (1997) Aviation automation: the search for a human-centred approach. Lawrence Earlbaum Associates, Mahwah, NJ

Dekker S (2002) Reconstructing human contributions to accidents: the new view on error and performance. J Safe Res 33(3):371–385

Dormann T, Frese M (1994) Error training: replications and the function of exploratory behavior. Int J Hum Comput Inter 6(4):365–372

Flach JM, Rasmussen J (2000) Cognitive engineering: designing for situation awareness. In: Sarter NB, Amalberti R (eds) Cognitive engineering in the aviation domain, Lawrence Erlbaum Associates, Mahwah, NJ

Frese M, Altmann A (1989) The treatment of errors in learning and training. In: Bainbridge L, Quintanilla SAR (eds) Developing skills with information technology, Wiley, Chichester, UK, pp 65–86

Frese M, Brodbeck F, Heinbokel T, Mooser C, Schleiffenbaum E and Thiemann P (1991) Errors in training computer skills: on the positive function of errors. Hum Comput Interact 6:77–93

Gunther D (2002) Threat and error management training counters complacency, fosters operational excellence. ICAO J 57(4):12–13

Helmreich RL (2001) A closer inspection: what really happens in the cockpit. Fl Safe Mag 1/2:32–35

Helmreich RL, Klinect JR, Wilhelm JA (1999) Models of threat, error, and CRM in flight operations. In: Proceedings of the Tenth International Symposium on Aviation Psychology, Ohio State University, Columbus, OH, 3–6 May, 1999, pp 677–682

Helmreich RL, Merritt AC, Wilhelm JA (1999) The evolution of crew resource management in commercial aviation. Int J Aviat Psychol 9:19–32

Hoffman RR, 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. Hum Fact 40(2):254–276

Hollnagel E (1993) The phenotype of erroneous actions. Int J Man-Mach Stud 39:1–32

Klein GA (1997) Developing expertise in decision making. Think Reason 3(4):337–352

Klein GA, Calderwood R, MacGregor D (1989) Critical decision method of eliciting knowledge. IEEE Trans Sys Man Cybern 19:462–472

Kontogiannis T (1999) User strategies in recovering from errors in man-machine systems. Safe Sci 32:49–68

Lintern G, Naikar N (2000) Analysis of crew coordination in the F-111 mission. DSTO Client Report (DSTO-CR-0184), Aeronautical and Maritime Research Laboratory, Melbourne, Australia

Maurino D (2001) At the end of the parade. Fl Safe Mag 1/2:36–39

Naikar N, Saunders A (2002) Crossing the boundaries of safe operation: training for error detection and error recovery. In: *P*roceedings of the 21st European Conference on Human Decision Making and Control, *G*lasgow, Scotland, 15–16 July, 2002

Naikar N, Saunders A, Hopkins A (2002) Profile of human error in the F-111 system. DSTO Client Report (DSTO-CR-0260), DSTO Systems Sciences Laboratory, Adelaide, Australia

Norman DA (1986) Cognitive engineering. In: Norman DA, Draper SW (eds) User centered system design: new perspectives on human-computer Interaction, Lawrence Erlbaum Associates, Mahwah, NJ, pp 31–61

Noyes JM (1998) Managing errors. In: Proceedings of the UKACC International Conference on Control, Swansea, UK, 1–4 September, 1998, pp 578–583

O'Hare D, Wiggins M (2002) Remembrance of cases past: who remembers what, when confronting critical flight events. (in press)

O'Hare D, Wiggins M, Batt R, Morrison D (1994) Cognitive failure analysis for aircraft accident investigation. Ergonomics 37(1):1855–1869

Paries J, Amalberti R (2000) Aviation safety paradigms and training implication. In: Sarter NB, Amalberti R (eds) Cognitive engineering in the aviation domain, Lawrence Erlbaum Associates, Mahwah, NJ, pp 253–286

Rasmussen J (1974) The human data processor as a system component: bits and pieces of a model. Report No. Risø-M-1722, Danish Atomic Energy Commission, Rosskilde, Denmark

Rasmussen J (1982) Human errors: a taxonomy for describing human malfunction in industrial installations. J Occup Accid 4:311–333

Rasmussen J, Pejtersen AM and Goodstein LP (1994) Cognitive systems engineering. Wiley, New York

Rasmussen J, Vicente KJ (1989) Coping with human errors through system design: implications for ecological interface design. Int J Man-Mach Stud 31:517–534

Reason J (2000) Human error: models and management. Brit Med J 320:768–770

Reason J (2001) The benign face of the human factor. Fl Safe Mag 1/2:28–31

Sarter NB, Alexander HM (2000) Error types and related error detection mechanisms in the aviation domain: an analysis of aviation safety reporting system incident reports. Int J Aviat Psychol 10:189–206

Shappell SA, Wiegmann DA (2000) The human factors analysis and classification system—HFACS. Report DOT/FAA/AM-00/7, National Technical Information Service, Springfield, UK

Vicente KJ (1999) Cognitive work analysis. Lawrence Erlbaum Associates, Mahwah, NJ

Wickens CD (1992) Engineering psychology and human performance (2nd ed). Harper-Collins, New York

Woods DD, Johannesen LJ, Cook RI and Sarter NB (1994) Behind human error: cognitive systems, computers and hindsight. Report CSERIAC SOAR 94–01, Crew Systems Ergonomics Information Analysis Center, Ohio