# A systems approach to training aeronautical decision making: from identifying training needs to verifying training solutions

**W.-C. Li**
Graduate School of Psychology
National Defence University
Taipei, South Korea

**D. Harris**
Department of Human Factors, School of Engineering
Cranfield University, Bedford

## ABSTRACT

The human factors analysis and classification system (HFACS) was developed as an analytical framework for the investigation of the role of human error in aviation accidents. A total of 523 accidents in the Republic of China (ROC) Air Force between 1978 and 2002 were analysed using this framework. The results showed that in a great many cases, poor pilot decision making was implicated. Following a survey of flight instructors' opinions, two of most promising mnemonic-based methods currently available to guide the decision making of pilots were identified (SHOR and DESIDE). These methods were developed into a short (four hour) aeronautical decision making training course. A total of 41 pilots from the Republic of China Tactical Training Wing then participated in a study to evaluate the effectiveness of this training course. Half of the participants received the short ADM training programme and half did not. Their decision making skill was evaluated in a series of emergency situations presented in a full-flight simulator. Furthermore, their decision making processes were examined in a series of pencil-and-paper based tests. The results clearly showed significant improvements in the quality of pilots' situation assessment and risk management (underpinning processes in pilot decision making) although this was usually at the expense of speed of response. Pilots used the quicker to apply SHOR mnemonic in situations that which required a fast decision and the more comprehensive but slower to perform DESIDE method when there were fewer time pressures. The results do strongly suggest that ADM is trainable and the short programme devised was effective.

## NOMENCLATURE

| | |
|---|---|
| $\chi^2$ | value in the Chi-square distribution used to test for significant differences between three (or more) independent groups when using the Kruskal-Wallis non-parametric analysis of variance where the dependent variable is measured on an ordinal scale. |
| $p$ | probability of making a type I decision error. |
| $z$ | a value in the standard normal distribution that may be related directly to a probability value to determine statistical significance when using the Wilcoxon matched pairs signed ranks test to establish if there is a statistically significant difference between two related samples where the dependent variable is measured on an ordinal scale. |

## 1.0 INTRODUCTION

To improve flight safety, the Republic of China (ROC) Air Force Headquarters routinely investigates all accidents and incidents. Accidents attributable solely to mechanical failure have decreased markedly but those attributable to human error have declined at a much slower rate and remain the primary cause of accidents. Operating a high-technology fighter aircraft is not only an issue of skilled psychomotor performance but also of real-time decision making involving situation awareness and risk assessment[1].

The interservice procedures for instructional system development (IPISD) adopts a systems perspective for the identification of needs and the development of training to achieve a goal[2]. This model, developed

**Paper No. 3086**. Manuscript received 24 April 2006, revised version received 5 February 2007, accepted 27 February 2007.

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

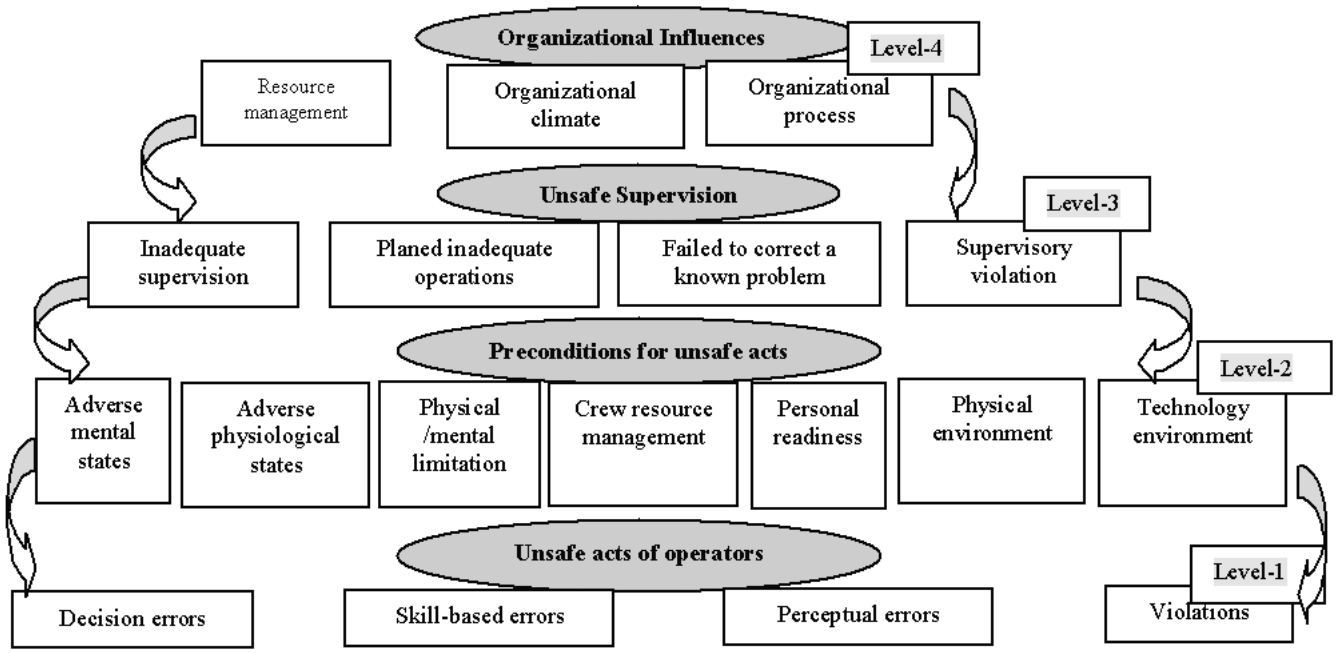



Figure 1. The HFACS framework, from Wiegmann & Shappell[4].

by the US military, provides a generalisable, context-free framework for the development of training programmes. The IPISD approach divides the development of training into five phases. These begin with identifying training needs, so that the content of training can be specified followed by a second phase specifying training objectives and mapping out the structure of training programme. Phase three develops the training contents into effective learning materials, which is followed by the fourth phase implementing training programme. The final phase evaluates the training intervention. The IPISD approach incorporates a feedback loop enable to make revisions from the evaluation during all five phases of training development. In this way the system of training development is capable of modification and improvement[3].

This paper describes three related studies. The first study is concerned with the identification of a specific training requirement to equip cadet pilots in the ROC Air Force with aeronautical decision making skills. This was achieved from the systematic human factors analysis of 523 accidents using the human factors analysis and classification system – HFACS[4]. Study two is concerned with the identification of potential training solutions, based upon mnemonic-based aeronautical decision making methods, and the subsequent design and development of a short training course. The final study is concerned with the validation of the efficacy of the training developed. After having received the training developed, cadet pilots' decision making skills were assessed in both a series of flight simulator exercises and using a pencil-and-paper based examination of their ability.

## 2.0 IDENTIFICATION OF ADM TRAINING REQUIREMENTS

### 2.1 Introduction

There are a number of perspectives for describing and analysing human errors, each based on different assumptions about the nature of error and the underlying causal factors of the human contribution in the sequence of events leading up to an accident. By examining and correlating information across a number of accidents, predictors may be identified which may then be applied to individual crews or situations in order to developing the effective prevention strategies. Many human factors accident analysis frameworks, taxonomies and analysis strategies have been devised over the years[5-9]. The human factors analysis and classification system (HFACS) is perhaps the one most commonly used and best validated, and is the one used herein as a basis for the identification of training needs.

HFACS is a generic human error analysis framework originally developed for US military aviation as a tool for the investigation of the human factors aspects of accidents. HFACS is based on Reason's[10] system-wide model of human error in which active failures are associated with the performance of front-line operators in complex systems and latent failures are characterised as inadequacies or mis-specifications which lie dormant within a system and are only triggered when combined with other factors to breach the system's defences. The development of HFACS is described in a series of books and papers[4,11-16]. However, as aviation accidents are the result of a number of causes, the challenge for accident investigators is how best to identify and mitigate the causal sequence of events leading up to an accident. HFACS examines human error at four levels. Each higher level affects the next downward level in the framework (Fig. 1).

- Level-1 'Unsafe acts of operators' is where the majority accident investigations are focused. Such causes can be classified into the two basic categories of errors and violation.
- Level-2 'Preconditions for unsafe acts' describes the latent failures within the causal sequence of events as and the context of the substandard conditions of operators and the practices they adopt.
- Level-3 'Unsafe supervision' traces the causal chain of events producing unsafe acts up to the front-line supervisors.
- Level-4 'Organisational influences' encompasses the most elusive latent failures, the fallible decisions of upper management which directly affects supervisory practices, as well as the conditions and actions of front-line operators.

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

Wiegmann and Shappell[16] suggest that the HFACS framework bridges the gap between theory and practice by providing safety professionals with a theoretically based tool for identifying and classifying human errors in aviation mishaps. The tool focuses on both latent and active failures and their inter-relationships, and it facilitates the identification of the underlying causes of human error. The objective of the first stage of the study was to identify any deficiencies which may be amenable to rectification through the development of a pilot training programme.

## 2.2 Method

The data were derived from the narrative descriptions of accidents occurring in the ROC Air Force between 1978 and 2002. The data set comprised of 523 accidents investigated during this 25-year period.

This study used the HFACS framework as described by Wiegmann and Shappell[4] to analyse the accident data. Each accident report was coded independently by two investigators. The investigators were trained on the HFACS framework together for 10 hours to ensure that they achieved a detailed and accurate understanding to the categories in the HFACS. The presence or the absence of each HFACS category was assessed in each report narrative. To avoid over-representation from any single accident, each HFACS category was counted a maximum of only once per accident. The count acted simply as an indicator of presence or absence of each of the 18 categories in a given accident.

## 2.3 Results

The 523 accidents analysed included 1,762 instances of human error recorded using the HFACS framework. Acts at the level of 'unsafe acts of operators' were involved in 725 (41·1%) of cases; the 'preconditions for unsafe acts' level was as a causal factor in 552 (31·3%) instances; the 'unsafe supervision' level was involved in 221 (12·5%) of accidents, and the 'organisational influences' level in the HFACS model was a factor in 264 (15%) cases.

### 2.3.1 Accident characteristics

Fighter aircraft were involved in 353 (67·5%) accidents, training aircraft in 113 (21·6%) accidents, and cargo aircraft were involved in 57 (10·9%) accidents; Cadet pilots were involved in 30 (5·7%) accidents, second lieutenants in ten (1·9%), first lieutenants in 92 (17·6%), captains in 144 (27·5%), majors in 148 (28·3%) and lieutenant colonel (or above) were involved in 70 (13·4%) accidents. Accidents occurred in all flight phases and missions[17].

### 2.3.2 Accident causal factors identified using the HFACS framework

At level-1, 'skill-based errors' exhibited the highest frequency of occurrence in the HFACS framework. These included actions such as inappropriate stick and rudder co-ordination, excessive use of flight controls, glide path not maintained, and adopting an improper airspeed or altitude. 'Decision errors' had the second highest rate of observations. Instances in this category included, selecting inappropriate strategies to perform a mission, improper in-flight planning, making an inappropriate decision to abort a take-off or landing, or using improper remedial actions in an emergency. The category of 'violations' included intentionally ignoring standard operating procedures (SOPs); neglecting SOPs and applying improper SOPs. The category of 'perceptual errors' exhibited the lowest frequency of occurrence. This category included spatial disorientation, experiencing visual illusions, making incorrect estimations of distance and descent rate during the approach, and vertigo during tactical manoeuvres (see Table 1).

At level-2 of the HFACS framework, instances of causal factors in the 'adverse mental states' category included issues such as over-confidence, stress, loss of situational awareness, distraction, and task saturation. 'Crew resource management' (CRM) issues, were the next most frequent category, included poor teamwork, poor communication, failures of leadership and inadequate briefing. The 'physical environment' category, included poor responses to factors such as, bad weather and terrain. The category of 'physical/mental limitations' included instances of information overload and a lack of experience to deal with a complex situation. The 'technological environment' included issues such as equipment design, cockpit display interfaces, automation and checklist layout. 'Personal readiness', encompassed issues associated with inadequate training, poor diet, and overexertion while off duty.

**Table 1**

**The frequency and percentage of accident and inter-rater reliability of HFACS categories (ranked in terms of frequency of occurrence) derived from Li and& Harris[17]**

| Categories of HFACS | HFACS level | Frequency of occurrence | Percentage agreement |
|---|---|---|---|
| Skilled-based errors | 1 | 226 | 83·4% |
| Decision errors | 1 | 223 | 81·5% |
| Adverse mental states | 2 | 184 | 86·0% |
| Resource management | 4 | 184 | 86·4% |
| Inadequate supervision | 3 | 177 | 89·7% |
| Violations | 1 | 160 | 84·9% |
| Crew resource management | 2 | 146 | 89·7% |
| Perceptual errors | 1 | 116 | 85·1% |
| Organisational process | 4 | 76 | 87·4% |
| Physical environment | 2 | 74 | 92·2% |
| Physical/mental limitation | 2 | 73 | 90·4% |
| Technology environment | 2 | 44 | 89·9% |
| Personal readiness | 2 | 29 | 72·3% |
| Planned inadequate operations | 3 | 24 | 94·6% |
| Failed correct a known problem | 3 | 12 | 95·8% |
| Supervisory violation | 3 | 8 | 96·2% |
| Organisational climate | 4 | 4 | 96·4% |
| Adverse physiological states | 2 | 2 | 96·4% |

The most frequently occurring category at level-3 was 'inadequate supervision'. Contributory factors included a failure to provide proper training, failure to track qualifications and performance, using untrained supervisors and loss of supervisory situation awareness. 'Planned inadequate operations' included issues surrounding poor crew pairings, a failure to establish if risk outweighed benefit, and failure to provide adequate time for briefing. In the category of 'failure to correct a known problem', instances included failures to remove a known safety hazard, failing to report unsafe tendencies, and failing to initiate corrective actions. 'Supervisory violations', including authorising an unqualified crew for flight, supervisors violating procedures, or a wilful disregard of authority by the supervisor, was implicated in relatively few accidents (Table 1).

At level-4, 'resource management' including the selection, staffing and training of human resources at an organisational level, excessive cost cutting and providing unsuitable equipment, was the factor most frequently involved in accidents. 'Organisational processes', which encompassed poor risk management programmes, poor mission scheduling, failing to set clearly defined objectives or failing to establish safety programmes, was the next most frequent category at this level in the framework. Issues surrounding the 'organisational climate' including inadequacies in the chain of command, poor delegation of authority, and inappropriate organisational customs, were involved in very few accidents.

The inter-rater reliability of the HFACS coding, calculated as a simple percentage rate of agreement, obtained figures of between 72·3% and 96·4%, indicating acceptable reliability (Table 1).

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

### 2.4 Discussion

'Skill-based errors' had the highest rate of occurrence (43·2%) which included actions such as inappropriate stick and rudder coordination, excessive use of flight controls, and adopting an improper airspeed or altitude. 'Decision errors' had the second highest rate of occurrence (42·6%) which included instances of improper in-flight planning, making an inappropriate decision to abort a take-off or landing, or using improper remedial actions in an emergency. The frequency of occurrence of both 'skill-based errors' (226) and 'decision errors' (223) was very similar. The initial training programmes for cadet pilots focus almost solely on factors at the skill-based level. There is no 'decision making' training programme in the ROC Air Force. Therefore, these data would suggest the need to address this issue with some urgency when training military pilots.

The importance of aeronautical decision making (ADM) has long been recognised as critical to the safe operation of aircraft[18]. In a further analysis of this data set it was observed that decision errors were often preceded by instances of poor CRM[19]. A similar trend has also been observed in commercial aircraft accidents[20]. Aviation psychologists and human factors specialists have investigated ADM for pilots and the results have strongly suggested that ADM can be improved by training[20-25]. The results from this first phase suggest that there is a need for military pilots to undergo decision making training to improve aviation safety. However, if such a training programme is to be maximally effective the most appropriate decision making training approach needs to be identified.

## 3.0 INVESTIGATING THE SUITABILITY OF ADM MNEMONICS IN TACTICAL SETTINGS

### 3.1 Introduction

There is a great variability in the nature of the tactical tasks confronting a military pilot[24]. Fighter pilots make important decisions often using ambiguous information, while under great risk and with very little time. Therefore, decision aids and training are required to provide these pilots with the necessary skills. It has been observed that the majority of fatal crashes are attributable to decision errors rather than perceptual or action errors[25]. The results of study one supported this observation. It has consistently been found that ADM can be improved with training[25-33]. However, Orasanu[32] has pointed out that there was no evidence to support the development of generic training techniques to improve all-purpose decision making skills, as there were different component skills involved in making six different basic types of decisions. These six types of decision included go/no go decisions; recognition-primed decisions; response selection decisions; resource management decisions; non-diagnostic procedural decisions; and creative problem-solving. As a result it was thought unlikely that any one single training method could improve all decision making.

There are a number of strategies (often embodied in mnemonics or acronyms) developed by aviation researchers and used by pilots as a simple guide to structure in-flight decision making. For economy of space the strategies used in this research are described in the method section. The common aim of these techniques is to form a systematic approach to decision making that is less affected by human nature[25]. Such methods could potentially produce a basis for an effective training programme for military pilots, however, there is lack of research investigating the efficacy of these ADM mnemonics. Furthermore if an ADM training programme is to be effective the most appropriate decision making strategy for each decision making situation has to be identified.

The aim of this second study is to identify the most appropriate ADM mnemonic method for each of the six basic decision types identified by Orasanu[34] as an essential step in the design and development of an effective training course.

### 3.2 Method

#### 3.2.1 Participants

A total of 60 instructor pilots with between 1,175 and 10,000 hours of flying experience were recruited. Their mean flying experience was 3,788 hours in the ROC Air Force Academy. The age of participants ranged between 32 and 54 years, with an average of 42 years.

#### 3.2.2 ADM mnemonics

Five ADM mnemonic-based methods that could potentially form the basis of the ADM training programme were selected from a review of the literature. These were:

- SHOR: The SHOR mnemonic-method[35] consists of four steps: stimuli, hypotheses, options, and response. It was originally developed for use by US Air Force tactical command and control personnel, for making decisions requiring near-real-time reactions involving threat warning, rescheduling and other types of dynamic modification.
- PASS: The PASS mnemonic-method[36] consists of four steps: problem identification, acquire information, survey strategy, select strategy. It was originally developed by a civil airline (Delta) to train pilots as part of a CRM training programme.
- SOAR: The SOAR mnemonic-based decision making method[37] comprises of four steps, situation, options, act, repeat. It was originally developed for glider pilots. It has the strength of encouraging the repeating of the evaluation of a changing situation after the initial actions have been made.
- FOR-DEC: The FOR-DEC mnemonic-based method[38], developed from the contents of a Lufthansa CRM-course, comprises of six steps: facts, options, risks & benefits, decision, execution, check. It incorporates an analysis of risk and benefits to handle situations including the effects of time pressure, continually changing conditions, distraction, and incomplete information.
- DESIDE: The DESIDE mnemonic method[39] also comprises of six steps, detect, estimate, set safety objectives, identify, do, evaluate. The DESIDE method is a practical adaptation of the conflict-theory model of Janis and Mann[40].

#### 3.2.3 Scenario selection

Six scenarios, taken from ROC Air Force accidents and incidents corresponding to the six types of decisions described by Orasanu[34], were used as stimulus material. These were:

1. Go/no go decision: When operating a Northrop F-5E, the No 2 wingman has to make a decision as the No 1 (Leader) abandons a tactical formation take-off at 145kt.
2. Recognition-primed decision: F-5E right engine fails as a result of foreign object damage just as the nose gear leaves the ground at the speed 165kt.
3. Response selection decision: No 4 in a tactical formation of F-5Es is required to make a decision when No 1 (Leader) becomes lost in cloud during formation flight (3ft distance between wing tips of the four fighters).
4. Resource management decisions: F-5E leader (No 1) of four aircraft needs to make a decision for the No 3 and No 4 aircraft when a 'no joy' call (no visual contact with No 1 and No 2) is made and No 2 calls 'one opposing target approaching at 12

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

o'clock with same altitude'. This occurs during practice of a two *versus* two engagement (air combat manoeuvre).
5. Non-diagnostic procedural decisions: Both the leader and wingman in a formation of F-5Es are unable to land at home-base in a 'bingo' (low fuel) situation during instrument flight in bad weather.
6. Creative problem-solving: When flying an F-5F both left and right generators fail at the same time during a tactical manoeuvre.

#### 3.2.4 Procedure

Each mnemonic method was evaluated using a structured rating form in terms of its suitability for situation assessment; risk management; response time; and applicability using a nine-point Likert-type scale (with a high score of 9 and a low score of 1). Supporting material was provided describing the requirements of each step in each of the mnemonic methods. The ADM rating forms were distributed to instructors in the Training Division, ROC Air Force Academy, and completed instruments were returned the next day. All instructors were fluent in the English language. None of the instructors had any previous experience in employing these structured decision making methods.

As a result of the length of the scenarios and the number of ratings required each instructor only evaluated ADM decision techniques in three scenarios selected randomly from the total of six. To eliminate order effects, the five ADM methods were presented in a randomised order in each scenario.

### 3.3 Results

In total, 180 ratings were made of the five ADM methods on the dimensions of situation assessment, risk management, response time, and applicability. A total of 29 ratings were made concerning the suitability of the ADM mnemonics in the go/no go scenario; 35 ratings in the recognition-primed decision scenario; 32 ratings in the response selection decision making scenario; 22 ratings in the resource management scenario; 32 ratings in the non-diagnostic procedural decision making scenario and 30 ratings were received concerning the suitability of the ADM methods in the creative problem-solving scenario.

The rank order of the instructors' ratings on each dimension for each ADM mnemonic-based methods in the six decision making situations used and the overall rank order of the summated scale scores are presented in Table 2.

#### 3.3.1 Treatment of data

For each participant an overall score for each mnemonic method in each scenario was created by summing the scores across these four scales giving a range of between 4 (low suitability) to 36 (high suitability). For statistical analysis using Friedman's non-parametric Analysis of Variance, the summated ratings of each ADM method in each scenario were then converted to ranking data, from 1 (highest suitability) to 5 (lowest suitability) — see Table 2.

#### 3.3.2 Statistical analysis

The highest ranking of suitability for the ADM methods in the go/no go decision scenario was received by PASS followed by FOR-DEC, DESIDE, SHOR, and SOAR. There was, however, no significant difference between these five ADM methods (Friedman's non-parametric analysis of variance: equivalent $\chi^2 = 1{\cdot}284$, $df = 4$, p = 0·864).

In the recognition-primed decision making scenario, the highest ranked mnemonic-based method was SHOR, followed by PASS, SOAR, FOR-DEC, and DESIDE. There was a significant difference between the five ADM methods (Friedman's non-parametric analysis of variance: equivalent $\chi^2 = 18{\cdot}975$, $df = 4$, $p = 0{\cdot}001$). Further examination of the relationship among those five methods by using post-hoc Wilcoxon's signed ranks test, found that five pairs of variables were differed significantly: FOR-DEC *vs* SHOR ($z = -3{\cdot}253$, $p = 0{\cdot}001$); SOAR *vs* SHOR ($z = -1{\cdot}996$, $p = 0{\cdot}046$); DESIDE *vs* SHOR ($z = -3{\cdot}515$, $p = 0{\cdot}000$); FOR-DEC *vs* PASS ($z = -2{\cdot}485$, $p = 0{\cdot}013$); DESIDE *vs* PASS ($z = -2{\cdot}505$, $p = 0{\cdot}012$).

In the response selection decision making scenario the highest ranked mnemonic method was DESIDE, followed by SHOR, PASS, SOAR, and FOR-DEC. However, there was no significant difference between these five ADM methods (Friedman's non-parametric analysis of variance: equivalent $\chi^2 = 5{\cdot}73$, $df = 4$, $p = 0{\cdot}220$).

From the analysis of the instructors' responses, the highest ranked mnemonic-based method in the resource management decision making scenario was FOR-DEC, followed by DESIDE, SHOR, PASS, and SOAR. As before, there were no significant differences between these methods (Friedman's non-parametric analysis of variance: equivalent $\chi^2 = 1{\cdot}851$, $df = 4$, $p = 0{\cdot}763$).

The highest ranked ADM method in the non-diagnostic procedural decision making scenario was DESIDE, followed by FOR-DEC, PASS, SOAR and SHOR. There was a statistically significant difference between these ADM methods (Friedman's non-parametric analysis of variance: equivalent $\chi^2 = 20{\cdot}405$, $df = 4$, $p = 0{\cdot}000$). Further examination of the relationship among these five mnemonic-based methods found that five pairs of variables were significantly different from each other: PASS *vs* SHOR ($z = -2{\cdot}085$, $p = 0{\cdot}037$); FOR-DEC *vs* SHOR ($z = -2{\cdot}973$, $p = 0{\cdot}003$); DESIDE *vs* SHOR ($z = -2{\cdot}651$, $p = 0{\cdot}008$); SOAR *vs* FOR-DEC ($z = -2{\cdot}917$, $p = 0{\cdot}004$); DESIDE *vs* SOAR ($z = -2{\cdot}519$, $p = 0{\cdot}012$).

Finally, in the creative problem-solving decision making scenario the highest ranked method was again SHOR, followed by FOR-DEC, DESIDE, SOAR and PASS. There was no significant difference between these methods (Friedman's non-parametric analysis of variance: equivalent $\chi^2 = 1{\cdot}719$, $df = 4$, $p = 0{\cdot}787$).

### 3.4 Discussion

Inspection of the results in Table 2 indicates that the SHOR mnemonic was thought to be the best ADM training mnemonic method in three of the four scenarios where fast responses were thought to be required. PASS was regarded as the best ADM training mnemonic-based method in the go/no go decision making scenario. Both SHOR and PASS were thought to be useful where a quick response is required. Both methods are simple, and were regarded by instructors as the best methods for rule-based decisions (condition-action rules). Overall, SHOR, DESIDE and FOR-DEC were regarded as good for making knowledge-based decisions (for well-defined problems) including response selection decisions and resource management decisions. However, only SHOR and DESIDE were regarded as acceptable methods for making knowledge-based decisions where the problems were ill-defined. These situations were encompassed in the non-diagnostic procedural decision scenario and creative problem-solving decision scenario. The instructor pilots' comments suggested that SHOR had the required characteristics to deal with urgent situations as it promoted quick responses. It was simple and easy to remember; it fitted the constraints inherent in time-limited and critical situations; it matched the general format of a pre-flight briefing; it was easy to put into practice; and it was thought that its logical procedures promoted safe action. As SHOR is basically an extension of the stimulus-response (S-R) paradigm of classical behaviourist psychology, it explicitly addresses the requirement to deal with two aspects of uncertainty in the decision making process; information input uncertainty (relating to hypothesis generation and evaluation) and consequence-of-action uncertainty[35] (which creates the requirement for option generation and evaluation).

DESIDE was evaluated as being the best ADM method to use when making response selection decisions and non-diagnostic proce-

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

**Table 2**

**Instructors' rankings on the four dimensions of suitability of the five ADM mnemonic methods in the six basic decision making scenarios. (Data taken from Li and Harris[41].**

| Scenarios | ADM Method | Situation assessment | Risk management | Response time | Applicability | Overall ranking for summated scale |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| Go/no go | SHOR | 3 | 4 | 2 | 2 | 4= |
| decisions | PASS | 1 | 2 | 1 | 1 | 1 |
| | FOR-DEC | 4 | 3 | 4 | 4 | 2 |
| | SOAR | 5 | 5 | 3 | 5 | 4= |
| | DESIDE | 2 | 1 | 5 | 2 | 3 |
| 2 | | | | | | |
| Recognition- | SHOR | 1 | 1 | 1 | 1 | 1 |
| primed | PASS | 4 | 3 | 3 | 2 | 2 |
| decisions | FOR-DEC | 1 | 4 | 4 | 5 | 4 |
| | SOAR | 3 | 4 | 2 | 4 | 3 |
| | DESIDE | 5 | 2 | 5 | 3 | 5 |
| 3 | | | | | | |
| Response | SHOR | 3 | 5 | 3 | 3 | 2 |
| selection | PASS | 4 | 3 | 4 | 4 | 3 |
| decisions | FOR-DEC | 2 | 4 | 5 | 2 | 5 |
| | SOAR | 5 | 2 | 2 | 4 | 4 |
| | DESIDE | 1 | 1 | 1 | 1 | 1 |
| 4 | | | | | | |
| Resource | SHOR | 1 | 3 | 1 | 3 | 3= |
| management | PASS | 5 | 5 | 5 | 5 | 3= |
| decisions | FOR-DEC | 3 | 2 | 1 | 1 | 1 |
| | SOAR | 4 | 4 | 1 | 4 | 5 |
| | DESIDE | 1 | 1 | 4 | 2 | 2 |
| 5 | | | | | | |
| Non-diagnostic | SHOR | 4 | 5 | 3 | 4 | 5 |
| procedural | PASS | 3 | 3 | 2 | 2 | 3 |
| decisions | FOR-DEC | 1 | 1 | 4 | 3 | 2 |
| | SOAR | 5 | 4 | 5 | 5 | 4 |
| | DESIDE | 2 | 2 | 1 | 1 | 1 |
| 6 | | | | | | |
| Creative | SHOR | 3 | 2 | 1 | 3 | 1 |
| problem- | PASS | 4 | 2 | 2 | 5 | 5 |
| solving | FOR-DEC | 1 | 1 | 3 | 1 | 2 |
| | SOAR | 5 | 5 | 4 | 1 | 4 |
| | DESIDE | 2 | 2 | 5 | 4 | 3 |

dural decisions (see Table 2). These are non-emergent situations and have no immediately dangerous threats. Pilots have time to think more extensively about their actions. The qualitative data elicited from the instructor pilots' showed that DESIDE had the necessary characteristics to deal with non-urgent situations as a result of its good situation assessment and risk management dimensions. It prompted a comprehensive approach in terms of the number of factors that it encompassed in the decision making process; it was regarded as providing a specific and clear approach to analyse a situation and it possessed a logical order that was easy to remember. However, it did require more time to undertake the required steps and analyse and respond to a changing situation. As a result the instructor pilots advised that practice using DESIDE on a simulator was important if the technique was to be of use when dealing with a real, in-flight situation.

The results from study two suggest that just two mnemonic-based methods had the potential to form a suitable basis for a decision making training that encompass all the requirements of these six basic decision making situations. Instructor opinions indicate that SHOR[35] was regarded as being the best for time-limited and urgent situations; DESIDE[39] was regarded as being superior for guiding knowledge-based decisions needing more comprehensive consideration. To optimise decision making training effectiveness it was regarded as being necessary to instruct pilots both with regard to the operation of the individual techniques and in recognising which technique was most appropriate to apply in a given circumstance.

### 3.5 ADM training course

SHOR and DESIDE mnemonic-based decision making methods formed the basis of the ADM training programme. The training programme commenced with an introduction to ADM theory. This was followed by a description of the content and method of appli-

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

cation of the SHOR and DESIDE ADM mnemonic-based methods and advice concerning which approach was most suitable in any given situation. Following this, participants practiced in the classroom the application of SHOR and DESIDE in flight situations exemplifying the six basic types of decision making scenario described by Orasanu[34]. Finally, the application of ADM in military aviation was described and the participants were de-briefed. The ADM training programme lasted approximately four hours in total.

# 4.0 EVALUATION THE EFFECTIVENESS OF ADM TRAINING INTERVENTION

## 4.1 Introduction

It is vital that a training programme is assessed to determine if it is achieving its goals[42]. The third study evaluated the effectiveness of the short ADM training course developed. This was undertaken by the evaluation of trainees' decision making process based on the underpinning decision making components of situation assessment and risk management. Both the decision making process and the product of the decision making were evaluated, the former through the use of a simple, knowledge-based pencil and paper based test and the latter from observation and instructor evaluation of trainee performance in a series of simulator trials.

In the dynamic tactical environment, effective decision making is highly dependent on situation awareness. Situation awareness and decision making are closely related and in both instances, situation assessment is a fundamental precursor[26,43]. Situation assessment has been described as the process by which the state of situation awareness is achieved and has been identified as a critical component of decision making[44].

For military pilots operating in a hostile environment, the normal hazards of aviation are further compounded by the enemy's intent for the destruction of the aircraft. Risk and time pressure are situational variables that further influence the decision making process, as risk and time pressure may call for an immediate response from a pilot irrespective of whether or not the problem was fully understood[45]. Risk management should be a key part of the decision making process[46]. To manage threats, pilots must first assess the risks associated with them. Risk assessment feeds into decision making in two ways: during the assessment of the precipitating threats and in evaluating potential courses of action[47].

Time pressure has several important implications for decision making. Decision makers will often experience high levels of stress and their thinking will often shift toward the direction of using less complicated reasoning strategies[48]. Military pilots often make important decisions using ambiguous information under great risk and time pressure. Therefore, decision aids and training are required to provide pilots with the necessary skills to make quick and accurate situation assessments[49].

There were several approaches to the evaluation of training that differ in terms of aims, criteria, and methods. The traditional approach to evaluation has as its goal to identify whether training meets its training objectives and if it doesn't how the training programme should be modified[50]. The role of evaluation is to correct and manage training design and this is integral to the instrumental system development (ISD) approach. Feedback from evaluation may result in revision to either or both of the training objectives or training design. The criteria used for this type of evaluation concern both the processes of training development and the products of training as manifested in performance, at both individual and organisational levels. To re-iterate slightly, the purpose of this phase was to evaluate the effectiveness of the ADM training programme through a flight simulator-based experiment and by a knowledge-based evaluation of trainees using paper and paper tests.

## 4.2 Method

### 4.2.1 Participants

A total of 41 male cadet pilots from ROC Air Force Tactical Training Wings participated in the study. The flying experience of participants was between 220 and 354 hours with an average of 292 hours. Participants were randomly divided into two groups, 21 pilots in experimental group which underwent training on the short ADM programme, and 20 pilots in control group who did not receive any training.

### 4.2.3 Evaluation scenarios for pre-training and post-training evaluation

As a result of time and resource constraints only three different types of decision making scenario were assessed in this phase. These were scenarios requiring participants to make recognition-primed decisions, non-diagnostic procedural decisions and engage in creative problem solving.

To develop scenarios for the simulator and pencil and paper evaluations, three focus groups were conducted, one for each decision making situation. The purpose of these focus groups was to verify that the scenarios, three pre-ADM training and three post-training, developed from the ROCAF accidents database, corresponded to the appropriate types of decision making situation and were of equivalent difficulty. Each focus group comprised of one human factors specialist and three senior instructor pilots. To negate practice effects, different (but equivalent) scenarios were used in the evaluations pre- and post ADM training. These focus groups also ensured that enough detail was available in the scenarios to meaningfully evaluate the decision making performance of pilots. These scenarios developed were as follows.

#### 4.2.3.1 Recognition-primed decision making scenario

Pre-training: As described in study two
Post-training: F-5E solo, after taking off at 500 feet, pilot hears two unusual sounds from the engines and feels the aircraft shake. Engine exhaust gas temperature is increased, and RPM decreased.

#### 4.2.3.2 Non-diagnostic procedural decision making scenario

Pre-training: As described in study two
Post-training: When an F-5E pilot is finishing basic fighting manoeuvre training, the ground intercept controller reports that home base weather is worsening. Surplus fuel is down to only 1,400lb. The pilot asks for weather conditions at alternative airports.

#### 4.2.3.3 Creative problem solving decision making scenario

Pre-training: As described in study two.
Post-training: When lowering the landing gear while on the downwind leg the landing gear shaft warning light illuminates, indicating the nose landing gear is abnormal.

### 4.2.4 Research design

Participants were randomly divided into two groups, an experimental group and a control group. Both groups participated in an initial set of baselining trials in both the flight simulator trials and in the knowledge-based pencil-and-paper tests to evaluate their decision making skills. After these initial trials the experimental group attended the four-hour 'ADM training programme for military pilots'. Both groups then participated in a further set of trials to further evaluate their in-flight decision making performance.

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

**Table 3**

**Summary of statistical analyses of the main effects of before/after training; main effects of trained/untrained groups and interaction effects on both the simulator trials and pencil-and-paper tests on the dimensions of situation assessment, risk management, and response time across three of decision making scenarios. All analyses of variance have 1 and 39 degrees of freedom. Statistically significant results ($p < 0.05$) are entered in bold type. Results approaching statistical significance $p < 0.10$) are entered in an italic typeface. *na* indicates not applicable**

| Six basic types of decisions | Dimensions of evaluation | Main effect of before/after training | | Main effect of trained/untrained | | Interaction effects | |
|---|---|---|---|---|---|---|---|
| | | S trials | P-P trials | S trials | P-P trials | S trials | P-P trials |
| Recognition -primed decisions | SA | $F = 3.520$ $p = 0.068$ | $F = 0.927$ $p = 0.342$ | $F = 6.904$ $p = 0.012$ | $F = 1.337$ $p = 0.225$ | $F = 1.735$ $p = 0.195$ | $F = 9.555$ $p = 0.004$ |
| | RM | $F = 12.467$ $p = 0.001$ | $F = 0.141$ $p = 0.710$ | $F = 6.736$ $p = 0.013$ | $F = 2.900$ $p = 0.097$ | $F = 2.248$ $p = 0.142$ | $F = 3.266$ $p = 0.078$ |
| | RT | $F = 2.778$ $p = 0.104$ | na | $F = 0.013$ $p = 0.910$ | na | $F = 3.890$ $p = 0.056$ | na |
| Non-diagnostic procedural decisions | SA | $F = 8.216$ $p = 0.007$ | $F = 8.216$ $p = 0.007$ | $F = 2.484$ $p = 0.123$ | $F = 3.593$ $p = 0.065$ | $F = 4.237$ $p = 0.046$ | $F = 19.540$ $p = 0.000$ |
| | RM | $F = 6.761$ $p = 0.013$ | $F = 0.067$ $p = 0.797$ | $F = 3.316$ $p = 0.076$ | $F = 1.887$ $p = 0.177$ | $F = 7.743$ $p = 0.008$ | $F = 3.266$ $p = 0.078$ |
| | RT | $F = 3.266$ $p = 0.078$ | na | $F = 1.753$ $p = 0.193$ | na | $F = 0.692$ $p = 0.411$ | na |
| Creative problem-solving | SA | $F = 5.364$ $p = 0.026$ | $F = 10.320$ $p = 0.003$ | F = 3.063 $p = 0.088$ | $F = 0.187$ $p = 0.668$ | $F = 2.993$ $p = 0.092$ | $F = 2.393$ $p = 0.130$ |
| | RM | $F = 6.617$ $p = 0.014$ | $F = 5.885$ $p = 0.020$ | $F = 0.669$ $p = 0.418$ | $F = 0.162$ $p = 0.690$ | $F = 4.278$ $p = 0.045$ | $F = 2.393$ $p = 0.130$ |
| | RT | $F = 3.185$ $p = 0.082$ | na | $F = 6.164$ $p = 0.017$ | na | $F = 2.132$ $p = 0.152$ | na |

The simulator trials and pencil-and-paper tests used the recognition-primed decision, non-diagnostic procedural decision, and creative problem solving scenarios. ADM performance was evaluated on three dimensions in the simulator; situation assessment, risk management, and response time, but only on the former two aspects in the pencil-and-paper knowledge-based tests. Evaluations were undertaken in the Northrop F-5E simulator from the ROCAF Tactical Wing. After finishing the simulator trials, the participants then undertook the pencil-and-paper knowledge tests.

As described previously, to negate practice effects, different scenarios were used pre- and post ADM training. To eliminate order effects, the three scenarios were presented in a randomised order during both the pre-test and post-test trials. Pilots' performance was evaluated on three dimensions (situation assessment, risk management and response time) by a qualified flight simulator instructor. These dimensions were derived from the earlier study to select the most appropriate ADM training mnemonic methods. Each aspect of performance was rated using a nine-point Likert-type scale (with a high score of 9 and a low score of 1, or in the case of the dimension of response time a score of 1 indicating the fastest and, 9 is the slowest). To control inter-rater reliability, the same instructor was used for all participants.

## 4.3 Results

### *4.3.1 Recognition-primed decisions*

With regard to performance in the flight simulator there was an effect approaching statistical significance on situation assessment performance before and after ADM training (Table 3). Pilots tended to perform better during the second set of trials (Table 4). However, there was no corresponding significant difference observed on the pencil-and-paper based tests. The group that had received ADM training also significantly outperformed the group that had not received training on situation assessment performance in the flight simulator but there was again no corresponding significant difference between the groups on the pencil-and-paper tests. The interaction term between the trained/untrained group and before training/after training trials was non-significant on the simulator tests but there was a significant interaction on the pencil-and-paper tests. The group that had received ADM training showed significantly greater gains in rated situation assessment performance during the second set of trials compared to the untrained group.

There was a significant difference on the dimension of risk management before and after ADM training on the simulator trials, however there was no significant difference on the pencil-and-paper

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

**Table 4**

**Means and standard deviations in performance scores in the recognition primed decision making scenario on both simulator trial and pencil-and-paper trial. The S-prefix denotes simulator trial; the P&P prefix denotes pencil-and-paper test**

| Recognition-primed decisions | | Group | N | Mean | Standard Deviation |
|---|---|---|---|---|---|
| Situation assessment | Pre-test | S-Trained | 21 | 5·00 | 1·703 |
| | | P&P-Trained | | 5·43 | 1·121 |
| | | S-Untrained | 20 | 4·35 | 1·599 |
| | | P&P-Untrained | | 5·55 | 1·234 |
| | Post-test | S-Trained | 21 | 5·86 | 1·526 |
| | | P&P-Trained | | 6·10 | 0·944 |
| | | S-Untrained | 20 | 4·50 | 1·051 |
| | | P&P-Untrained | | 5·20 | 1·436 |
| Risk management | Pre-test | S-Trained | 21 | 4·48 | 1·537 |
| | | P&P-Trained | | 5·29 | 1·189 |
| | | S-Untrained | 20 | 4·05 | 0·945 |
| | | P&P-Untrained | | 5·30 | 1·128 |
| | Post-test | S-Trained | 21 | 5·71 | 1·309 |
| | | P&P-Trained | | 5·86 | 0·727 |
| | | S-Untrained | 20 | 4·55 | 1·146 |
| | | P&P-Untrained | | 4·95 | 1·191 |
| Response time | Pre-test | S-Trained | 21 | 4·67 | 1·592 |
| | | P&P-Trained | | na | na |
| | | S-Untrained | 20 | 5·35 | 1·872 |
| | | P&P-Untrained | | na | na |
| | Post-test | S-Trained | 21 | 5·86 | 0·964 |
| | | P&P-Trained | | na | na |
| | | S-Untrained | 20 | 5·25 | 1·446 |
| | | P&P-Untrained | na | na | |

tests. There was a significant difference in performance between the trained and untrained group on simulator trials and a result verging on significance on the pencil-and-paper tests. In both cases the group that had received ADM training outperformed the other group. There was no significant interaction term between the trained/untrained group and trial on the simulator trials, however, there was a result approaching significance on the pencil-and-paper tests. The trained group showed greater gains in risk management performance in the latter trials compared to the untrained group

There was no significant difference on rated speed of response between trials on the flight simulator and there was also no difference between the trained and untrained group. There was, however, a result verging on significant with respect to the interaction term. The group that had received ADM training tended to exhibit longer response times in the second trial compared to the untrained group.

### *4.3.2 Non-diagnostic procedural decisions*

There was a significant difference in pilots' situation assessment performance before and after ADM training on the simulator trials (Table 3). Performance was better during the second set of trials (Table 5). However, there was no significant difference on the pencil-and-paper tests. There was no significance between the trained and untrained groups on the simulator trials, however, there was an effect verging on a significant difference on pencil-and-paper tests. The trained group tended to outperform the untrained group. There were significant interactions terms on both the simulator trials and pencil-and-paper tests. The group receiving ADM training showed significantly greater gains in performance in the second trial compared to the untrained.

There was a significant difference on flight simulator performance on the dimension of risk management before and after training. The results indicated pilots' performance was significantly better on the second trials. There was, however, no significant difference in risk management performance before and after training on the pencil-and-paper tests. There was a result tending toward statistical significance on the simulator trials between the trained and untrained groups with the group that had received ADM training exhibiting better risk management performance than the group that had not. There was no significant difference between the trained and untrained group on the pencil-and-paper tests, though. Furthermore, there was a significant interaction between the trained/untrained group and trial on the flight simulator trials and a result verging on

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

**Table 5**

**Means and standard deviations in performance scores in the non-diagnostic procedural scenario on both simulator trial and pencil-and-paper test. The S-prefix denotes simulator trial; the P&P prefix denotes pencil-and-paper test**

| Non-diagnostic procedural decisions | | Group | N | Mean | Standard deviation |
|---|---|---|---|---|---|
| Situation assessment | Pre-test | S-Trained | 21 | 4·38 | 1·774 |
| | | P&P-Trained | | 5·00 | 1·304 |
| | | S-Untrained | 20 | 4·55 | 0·826 |
| | | P&P-Untrained | | 5·30 | 1·218 |
| | Post-test | S-Trained | 21 | 5·90 | 1·480 |
| | | P&P-Trained | | 6·19 | 1·123 |
| | | S-Untrained | 20 | 4·80 | 1·196 |
| | | P&P-Untrained | | 4·55 | 1·638 |
| Risk management | Pre-test | S-Trained | 21 | 4·29 | 1·454 |
| | | P&P-Trained | | 4·95 | 1·161 |
| | | S-Untrained | 20 | 4·55 | 0·826 |
| | | P&P-Untrained | | 5·25 | 1·070 |
| | Post-test | S-Trained | 21 | 5·76 | 1·375 |
| | | P&P-Trained | | 5·71 | 0·956 |
| | | S-Untrained | 20 | 4·50 | 1·192 |
| | | P&P-Untrained | | 4·60 | 1·465 |
| Response time | Pre-test | S-Trained | 21 | 4·33 | 1·494 |
| | | P&P-Trained | | na | na |
| | | S-Untrained | 20 | 4·20 | 0·894 |
| | | P&P-Untrained | | na | na |
| | Post-test | S-Trained | 21 | 5·38 | 1·244 |
| | | P&P-Trained | | na | na |
| | | S-Untrained | 20 | 4·80 | 1·152 |
| | | P&P-Untrained | | na | na |

significance on the pencil-and-paper tests. In both cases the group that had received ADM training showed greater gains in performance on the second trial compared to the untrained group.

There was a significant difference on the dimension of rated speed of response between trials showing that pilots' response times during the second trial in the flight simulator were significantly longer than the first trial. However, there was no significant difference between the trained and untrained group in this respect and there was also no significant interaction term.

### *4.3.3 Creative problem-solving*

There was a significant difference on the dimension of situation assessment before and after decision making training on both the simulator trials and pencil-and-paper tests (Table 3). The results showed that pilots' situation assessment was significantly better after ADM training (Table 6). There was no significant difference between the groups on either the simulator trial or the pencil-and-paper tests, although the former was verging on significance. There was also a result approaching significance for the interaction term on the simulator trials, however, there was no significant interaction term on the pencil-and-paper tests.

There was again a significant difference on the dimension of risk management before and after ADM training on both the simulator trials and pencil-and-paper tests. Pilots in the group receiving training exhibited superior risk management skills. There was no significant difference in this regard between the trained and untrained group on either the simulator trial or the pencil-and-paper trial. There was a significant interaction term between the trained/untrained group and trial on the simulator trials, however, there was no significant interaction term between the trained/untrained group and trial on pencil-and-paper trial. The group that had received ADM training showed greater gains in performance in the simulator trials (after training) compared to the untrained group.

There was a result verging on significant difference on the dimension of response time between trials. Pilots' response times during the second trial were longer than the first trial. There was also a significant difference between the trained and untrained group. The group that had received ADM training tended to be rated as having significantly longer response times than the group that had not received training. There was no significant interaction term between the main effects.

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

**Table 6**

**Means and standard deviations in performance scores in the creative problem-solving scenario on both simulator trial and pencil-and-paper test. The S-prefix demotes simulator trial; the P&P prefix denotes pencil-and-paper test**

| Creative problem-solving | | Group | N | Mean | Standard deviation |
|---|---|---|---|---|---|
| Situation assessment | Pre-test | S-Trained | 21 | 5·05 | 1·830 |
| | | P&P-Trained | | 4·71 | 1·347 |
| | | S-Untrained | 20 | 5·00 | 1·686 |
| | | P&P-Untrained | | 4·90 | 1·483 |
| | Post-test | S-Trained | 21 | 6·43 | 1·076 |
| | | P&P-Trained | | 5·71 | 1·007 |
| | | S-Untrained | 20 | 5·20 | 1·704 |
| | | P&P-Untrained | | 5·25 | 1·020 |
| Risk management | Pre-test | S-Trained | 21 | 4·86 | 1·682 |
| | | P&P-Trained | | 4·71 | 1·347 |
| | | S-Untrained | 20 | 5·15 | 1·843 |
| | | P&P-Untrained | | 4·95 | 1·761 |
| | Post-test | S-Trained | 21 | 6·24 | 1·375 |
| | | P&P-Trained | | 5·67 | 0·966 |
| | | S-Untrained | 20 | 5·30 | 1·380 |
| | | P&P-Untrained | | 5·15 | 1·226 |
| Response time | Pre-test | S-Trained | 21 | 5·19 | 1·778 |
| | | P&P-Trained | | na | na |
| | | S-Untrained | 20 | 4·80 | 1·735 |
| | | P&P-Untrained | | na | na |
| | Post-test | S-Trained | 21 | 6·19 | 1·167 |
| | | P&P-Trained | | na | na |
| | | S-Untrained | 20 | 4·90 | 1·021 |
| | | P&P-Untrained | | na | na |

### 4.4 Discussion

Overall, results from both the simulator-based trials (which assessed the product of the ADM training programme) and the pencil-and-paper test (which assessed the process that the trainees applied) show gains being made in both situation assessment and risk management skills which are attributable to the decision making training course. Perhaps the most direct indication of the efficacy of the ADM training course, though, lies in the significant interaction effects obtained. These interaction terms indicate disproportionate gains in situation assessment and risk management performance on the second trial (post ADM training) in the participant group that received ADM instruction (see Table 3).

Further data obtained in the structured de-briefing of the trainee pilots after each simulator trial and the results of the pencil-and-paper tests clearly showed that the vast majority of pilots who had received ADM training applied the most appropriate ADM mnemonic method in a given scenario. The SHOR mnemonic was most often applied in the recognition-primed decision making scenario (which required a fast response) and DESIDE was most commonly used in the non-diagnostic procedural decision making scenario and in the creative problem-solving situation. In both the latter scenarios, there were no immediate time or fuel pressures on the pilots, allowing a more comprehensive consideration of the situation using the DESIDE mnemonic. De-briefing of the participants and analysis of the pencil-and-paper tests also showed that the ADM mnemonics had been applied in the correct manner. However, situations develop quickly in a high performance jet aircraft and with only a single pilot on board there is also often the requirement to make a quick (but not a 'hasty') decision. These factors, coupled with the often 'individualist' nature of fighter pilots may pre-dispose some to utilise the quicker to apply SHOR mnemonic method in all instances. It may be useful to undertake further research in a multi-crew context to investigate if there is a pre-disposition in transport pilots to use the more comprehensive DESIDE ADM mnemonic method as a result of having fewer time or workload pressures in some situations.

Improvements in situation assessment and risk management performance were, however, sometimes obtained at the expense of a decreased speed of response. In the recognition-primed decision making scenario, which required a rapid response, after training pilots responses were rated as being slower than before but their situation assessment and risk management decisions were improved. This may be a reasonable result, as the trainees avoided a rushed, ill-considered response, even in this situation. Similar results were also

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

obtained in the other two scenarios. Unlike the recognition-primed decision making situation, though, in these cases an immediate decision was not required, hence sacrifices in speed of response were not so critical. However, as before, it can be suggested that 'better' decisions were being made (in terms of situation assessment and risk management) at the expense of these longer response times. The results from the pencil-and-paper tests would support this observation. It is also possible that further training and practice may speed this process.

The nature of a decision depends on the requirements of the tasks and the conditions of the surrounding environment. If the evolving situation is not coincident with established SOPs or previous experience, pilots must to conduct an analysis of the problem and devise a solution. These results add support to the results of earlier research that suggested that ADM was trainable[21,22,26,29,31,33]. In the second phase of this study the opinions of instructor pilots identified SHOR[35] as potentially being the best ADM mnemonic in a time-limited situation and DESIDE[39] was thought to be superior for more complex, knowledge-based decisions where more time was available. This last phase provides supports the conclusions of the earlier survey and provides empirical evidence to demonstrate that pilots trained in the use of these techniques actually make better in-flight decisions in critical situations.

In this study decision making was considered as a standalone skill. However, aspects of ADM, such as risk management, are also key components in many CRM courses. Indeed the FOR-DEC mnemonic-based method[38] was developed from the contents of a CRM. Although this study concentrated on pilots of single-set fighter aircraft, teamwork (in the broadest sense) is also a fundamental skill required of these pilots. Although they may fly a single seat aircraft, they never fly alone. It is suggested that improvements in ADM skills would also produce concomitant benefits in CRM skills, both in single and multi-crew aircraft.

## 5.0 OVERALL DISCUSSION AND CONCLUSION

The simple, short, cost-effective training programme in the appropriate use of ADM mnemonic methods devised in this study produced significant gains in decision making performance. Such a course may easily be integrated into current CRM or simulator-based training programmes.

This paper demonstrates how through conducting a structured human factors assessment of accidents and incidents a well-targeted training programme may be constructed. The use of the interservice procedures for instructional system development (IPISD) training development model also ensures that the ADM course developed is properly evaluated at all stages to further guarantee its effectiveness. Kirkpatrick[51] suggests that training effectiveness should be examined by both proximal and distal measures. In this case the proximal measures show significant gains in the underpinning skills of decision making, however, it still needs to be established if these performance gains continue to be evident at a later date during actual operations. This will be established in the near future, however, the initial results would seem to be most promising.

## REFERENCES


1. ORASANU, J. and CONNOLLY, T. The reinvention of decision making, 1993, *Decision Making in Action: Models and Methods*, KLEIN, G.A., ORASANU, J., CALDERWOOD, R. and ZSAMBOK, C.E. (Eds), pp 3-20, Ablex, Norwood, NJ, USA.
2. BRANSON, R.K., RAYNER, G.T., COX, J.L., FURMAN, J.P., KING, F.J. and HANNUM, W.H. Interservice procedures for instructional system development, 1975, Report N613.39-73-C-O150, Naval Education and Training Command, Florida, USA.
3. PATRICK, J. Training, 2003, *Principles and Practice of Aviation Psychology: Human Factors in Transportation,* TSANG, P.S. and VIDULICH, M.A. (Eds), pp 397-434, Lawrence Erlbaum, New Jersey, USA.
4. WIEGMANN, D.A. and SHAPPELL, S.A. *A Human Error Approach to Aviation Accident Analysis: The Human Factors Analysis and Classification System*, 2003, Ashgate, Aldershot, UK.
5. DIEHL, A. Human performance/system safety Issues in aircraft accident investigation and prevention, 1989, Fifth International Symposium on Aviation Psychology, pp 838-847, Ohio State University, USA.
6. FEGGETTER, A.J. The development of an intelligent human factors data base as an aid for the investigation of aircraft accidents, 1991, Sixth International Symposium on Aviation Psychology, pp 324-629, Ohio State University, USA.
7. HOLLNAGEL, E. *Cognitive Reliability and Error Analysis Method*, 1998, Elsevier Science, Oxford, UK.
8. HUNTER, D.R. and BAKER, R.M. Reducing accidents among general aviation pilots through a national aviation safety program, 2000, Fourth Australian Aviation Psychology Symposium, Ashgate, Aldershot, UK.
9. DEKKER, S. *The Field Guide to Human Error Investigations*, 2002, Ashgate, Aldershot, UK.
10. REASON, J. *Human Error*, 1990, Cam bridge University, New York, NY, USA.
11. SHAPPELL, S.A. and WIEGMANN, D.A. Applying reason: the human factors analysis and classification system (HFACS), *Human Factors and Aerospace Safety*, 2001, **1**, (1), pp 59-86.
12. SHAPPELL, S.A. and WIEGMANN, D.A. Report DOT/FAA/AM-03/4, 2003, Federal Aviation Administration, Washington, DC, USA.
13. SHAPPELL, S.A. and WIEGMANN, D.A. HFACS analysis of military and civilian aviation accidents: a North American comparison, 2004, pp 2-8, International Society of Air Safety Investigators, Queensland, Australia.
14. WIEGMANN, D.A. and SHAPPELL, S.A. Human factors analysis of postaccident data: applying theoretical taxonomies of human error, *Int J of Aviation Psychology*, 1997, **7**, (1), pp 67-81.
15. WIEGMANN, D.A. and SHAPPELL, S.A. Human error analysis of commerical aviation accidents: application of the human factors analysis and classification system, *Aviation, Space, and Environmental Medicine*, 2001, **72**, (11), pp 1006-1016.
16. WIEGMANN, D.A. and SHAPPELL, S.A. Human error perspectives in aviation, *Int J of Aviation Psychology*, 2001, **11**, (4), pp 341-357.
17. LI, W.C. and HARRIS, D. HFACS analysis of ROC Air Force aviation accidents: reliability analysis and cross-cultural comparison, *Int J of Applied Aviation Studies*, 2005, **5**, (1), pp 65-81.
18. JENSEN, R. and HUNTER, D. General Aviation Aeronautical Decision-making, 2002, Federal Aviation Administration, Washington, DC, USA.
19. LI, W.C. and HARRIS, D. Pilot error and its relationship with higher organizational levels: HFACS analysis of 523 accidents, *Aviation, Space and Environmental Medicine*, 2006, **77**, (10), pp 1056-1061.
20. LI, W.C., HARRIS, D. and YU, C.S. Routes to failure: analysis of 41 civil aviation accidents from the Republic of China using the human factors analysis and classification system, Submitted to *Accident Analysis and Prevention*.
21. BUCH, G., and DIEHL, A. An investigation of the effectiveness of pilot judgment yraining, *Human Factors*, 1984, **26**, (5), pp 557-564.
22. CONNOLLY, T.J., BLACKWELL, B.B. and LESTER, L.F. A simulator-based approach to training in aeronautical decision making, *Aviation, Space, and Environmental Medicine*, 1989, **60**, (1), pp 50-52.
23. DIEHL, A. A workshop on understanding and preventing aircrew error, 1991, Sixth International Symposium on Aviation Psychology, pp 28-39, Ohio State University, USA.
24. KLEIN, G.A. and WOODS, D.D. Conclusions: decision making in action, 1993, *Decision Making in Action: Models and Methods*, KLEIN, G.A. ORASANU, J., CALDERWOOD, R. and ZSAMBOK, C.E. (Eds), pp 404-411, Ablex, Norwood, NJ, USA.
25. O'HARE, D. Aeronautical decision making: metaphors, models and methods, 2003, *Principles and Practice of Aviation Psychology: Human Factors in Transportation*, TSANG, P.S. and VIDULICH, M.A. (Eds), pp 201-237, Lawrence Erlbaum, NJ, USA.
26. PRINCE, C. and SALAS, E. Situation assessment for routine flight and decision making, *Int J of Cognitive Ergonomics*, 1997, **1**, (4), pp 315-324.
27. JENSEN, R. and BENEL, R. *Judgment Evaluation and Instruction in Civil Pilot Training*, 1977, Federal Aviation Administration, Washington, DC, USA.
29. COHEN, M.S. Three paradigm for viewing decision biases, 1993, *Decision Making in Action: Models and Methods*, KLEIN, G.A., ORASANU, J., CALDERWOOD, R. and ZSAMBOK, C.E. (Eds), pp 36-50, Ablex, Norwood, NJ, USA.
29. ENDSLEY, M.R. Measurement of situation awareness in dynamic systems, *Human Factors*, 1995, **37**, (1), pp 65-84.


https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press

30. DRILLINGS, M. and SERFATY, D. Naturalistic decision making in command and control, 1997, *Naturalistic Decision Making*, ZSAMBOK, C.E. and KLEIN, G. (Eds), pp 71-80, Lawrence Erlbaum, Mahwah.
31. KLEIN, G.A. A recognition-primed decision (RPD) model of rapid decision making, 1993, *Decision Making in Action: Models and Methods*, KLEIN, G.A., ORASANU, J., CALDERWOOD, R. and ZSAMBOK, C.E. (Eds), pp 138-147, Ablex, Norwood, NJ, USA.
32. ORASANU, J. and FISCHER, U. Finding decisions in natural environments: the view from the cockpit, 1997, *Naturalistic Decision Making*, ZSAMBOK, C.E. and KLEIN, G. (Eds), pp 343-358, Lawrence Erlbaum, Mahwah.
33. WAAG, W.L. and BELL, H.H. Situation assessment and decision making in skilled fighter pilots, 1997, ZSAMBOK, C.E. and KLEIN, G. (Eds), *Naturalistic Decision Making* pp 247-256, Lawrence Erlbaum, Mahwah.
34. ORASANU, J. Decision making in the cockpit, 1993, *Cockpit Resource Management*, WIENER, E.L., KANKI, B.G. and HELMREICH, R.L. (Eds), pp 137-172, Academic Press, San Diego, CA, USA.
35. WOHL, J.G. Force management decision requirements for air force tactical command and control, 1981, IEEE Transactions on Systems, Mans, and Cybernetics, SMC-11, pp 618-639.
36. MAHER, J. Beyond CRM to decisional heuristics: an airline generated model to examine accidents and incidents caused by crew errors in deciding, 1989, Fifth International Symposium on Aviation Psychology, pp 439-444, Ohio State University, USA.
37. OLDAKER, I. Pilot decision making — an alternative to judgement training, 1996, *Technical Soaring*, **20**, (2), pp 36-41.
38. HORMANN, H.J. FOR-DEC: a perspective model for aeronautical decision making, 1995, FULLER, R., JOHNSTON, R. and MCDONALD, N. (Eds), *Human Factors in Aviation Operations*, pp 17-23, Ashgate Publishing, Aldershot, UK.
39. MURRAY, S.R. Deliberate decision making by aircraft pilots: a simple reminder to avoid decision making under panic, 1997, *Int J of Aviation Psychology*, **7**, (1), pp 83-100.
40. JANIS, I.L. and MANN, L. Decision making: a psychological analysis of conflict, choice, and commitment, 1977, Free Press, New York, NY, USA.
41. LI, W.C. and HARRIS, D. Aeronautical decision-making mnemonics: instructor-pilots evaluation of five alternative methods, *Aviation, Space and Environmental Medicine*, 2005, **76**, (12), pp 1156-1161.
42. Report No Advisory Circular 120-51A, 1993, Federal Aviation Administration, US Department of Transport, Washington, DC, USA.
43. NOBLE, D. A model to support development of situation assessment aids, 1993, KLEIN, G.A., ORASANU, J., CALDERWOOD, R. and ZSAMBOK, C.E. (Eds), *Decision Making in Action: Models and Methods*, pp 287-305), Ablex, Norwood, NJ, USA.
44. ENDSLEY, M.R. and BOLSTAD, C.A. Individual differences in pilot situation awareness, *International J of Aviation Psychology*, 1994, **4**, (3), pp 241-264.
45. FISCHER, U., ORASANU, J. and WICH, M. Expert pilots' perceptions of problem situations, 1995, Eighth International Symposium on Aviation Psychology, pp 777-782, Ohio State University, USA.
46. JENSEN, R.S., GUILKE, J. and TIGNER, R. Understanding expert aviator judgment, 1997, *Decision Making Under Stress: Emerging Themes and Applications*, FLIN, R., SALAS, E., STRUB, M. and MARTIN, L. (Eds), (pp 233-242, Ashgate, Aldershot, UK.
47. ORASANU, J., DAVISON, J. and FISCHER, U. The role of risk in aviation decision making: how pilots perceive and manage flight risks, 2001, Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, pp 58-62, Santa Monica, USA.
48. PAYNE, J.W., BETTMAN, J.R. and JOHNSON, E.J. Adaptive strategy selection in decision-making, *J of Experimental Psychology: Learning, Memory and Cognition*, 1988, **14**, pp 534-552.
49. KAEMPF, G.L. and ORASANU, J. Current and future applications of naturalistic decision making, 1997, *Naturalistic Decision Making*, ZSAMBOK, C.E. and KLEIN, G. (Eds), pp 81-90, Lawrence Erlbaum, Mahwah.
50. PATRICK, J. *Training: Research and Practice*, 1992, Academic Press, London, UK.
51. KIRKPATRICK, D.L. Evaluation of training, 1976, *Training and Development Handbook*, CRAIG, R.L. and BITTEL, L.R. (Eds), pp 18.1-18.27, McGraw Hill, New York, NY, USA.

https://doi.org/10.1017/S0001924000004516 Published online by Cambridge University Press