The interpretation and use of weather radar displays in aviation

Final report

Prepared by: Dr Mark Wiggins
MARCS Auditory Laboratories
University of Western Sydney

April 2005

The Interpretation and Use of Weather Radar Displays in Aviation
Sponsored under the Aviation Safety Research Grants Programme 2004-2005
Australian Transport Safety Bureau
PO Box 967
Civic Square ACT 2608

Contents

CONTENTS III

ACKNOWLEDGEMENTS V

EXECUTIVE SUMMARY 1

1. INTRODUCTION 2

2. WEATHER-RELATED DECISION-MAKING IN AVIATION 3

2.1 Weather Radar and Weather-Related Decisions 3

2.2 Weather Radar: Design and Training 4

2.3 Empirical Perspectives of Expertise 5
2.3.1 Expertise Defined 6
2.3.2 An Expertise Model for the use of Weather Radar in Aviation 8
2.3.3 The Current Project 9

3. STUDY ONE 10

3.1 Aim 10

3.2 Data Sources 10
3.2.1 Procedure 10

3.3 Results 11
3.3.1 Raters 11
3.3.2 Aviation Safety Reporting System 11
3.3.3 Federal Aviation Administration (FAA) Accident/ Incident Data System 12
3.3.4 National Transportation Safety Board Database 13

3.4 Discussion 13
3.4.1 Limitations 14

4. STUDY TWO 15

4.1 Cognitive Interview 15

4.2 Method 16
4.2.1 Participants 16
4.2.2 Interview Protocol 16

4.2.3 Procedure 17

4.3 Results and Discussion 18
4.3.1 Perception 18
4.3.2 Information Management 19
4.3.3 Information Acquisition 19
4.3.4 Option Generation 21
4.3.5 Option Evaluation 22
4.3.6 Analogous Situations 23
4.3.7 General Comments 23

4.4 Conclusion 23
4.4.1 Limitations 24

5. STUDY THREE 25

5.1 Method 25
5.1.1 Participants 25
5.1.2 Materials 26
5.1.3 Procedure 27

5.2 Results 27
5.2.1 The Use of Weather Radar Displays 27
5.2.2 Weather Radar-Related Error Analysis 27
5.2.3 The Interpretation of Weather Radar Displays 29

5.3 Discussion 44
5.3.1 Limitations 47

6. GENERAL DISCUSSION 48

6.1 Conclusion 51

7. REFERENCES 52

ATTACHMENTS 58

Attachment A: Summary of Expertise Literature 59
Features of Expertise 60

Attachment B: Cognitive Interview Protocol 71

Attachment C: Selected Pages of the Weather Radar Survey 76

Attachment D: Participants' Statements for Scenarios 2 and 8 82

Acknowledgements

This research was made possible through a grant provided by the Commonwealth of Australia, through the Australian Transport Safety Bureau, Department of Transport and Regional Services, and through the assistance of Dr Sandra Bollwerk, Brett Molesworth, CADRE, and Eastern Airlines (Qantaslink). The author would also like to acknowledge all of those pilots who generously gave up their time to participate in the research process.

EXECUTIVE SUMMARY

This project investigated the use of weather radar displays in commercial aviation. Three studies are described.

The first study used an expertise model of the use of weather radar displays to classify aircraft accident and incident reports. The three data sources used were the Federal Aviation Administration Accident/Incident Data System, the National Transportation Safety Board Accident and Incident Database, and the Aviation Safety Reporting System. Although generalisation of the outcomes is limited, the results provide some evidence to suggest that where the use of weather radar was implicated in an aircraft accident or incident, the error was most likely to be associated with a failure to recognise and/or interpret the information on the display.

Study Two involved a cognitive interview of experienced commercial pilots and their use of weather radar displays to assist in the management of flight. The results revealed a relatively consistent response which emphasised the timely and accurate interpretation of radar 'paints' as the basis for successful performance. It was apparent that, for some pilots, the process involved the development and application of 'rules-of-thumb' and that these rules had been acquired through experience.

The results of Studies One and Two provided the basis for the development of a survey that was distributed to pilots both in hard-copy and on-line via the internet. Respondents were asked to provide information about their use of weather radar displays, describe an incident involving the use or misuse of weather radar displays, and give their interpretations of a series of 12 simulated weather radar 'paints'. In the case of the incidents described by the respondents, the results indicated that the majority of cases were related to the recognition and interpretation of the information on displays, consistent with the outcomes of Study One.

In relation to the simulated weather radar 'paints', the results indicated that while the interpretation of some of the displays was relatively consistent, the responses to other displays were less consistent between respondents. These differences were not due to demographic features such as age or experience, but appeared due to the level of ambiguity associated with the displays. Specifically, for some displays, it appeared that the key cues necessary for the successful interpretation of the information were either difficult to interpret or were absent. This outcome forms the basis for a number of recommendations concerning improvements in training, education and the design of weather radar displays.

1. INTRODUCTION

This report summarises the outcomes of a project to investigate the use of weather radar displays in aviation. The specific aims of the project were to:

(a) Develop an expertise model of the use of weather radar displays to acquire information about the role of the weather radar displays in aircraft accident causation;

(b) Apply the expertise model to a series of aircraft accident and incident reports with the purpose of establishing the utility of the approach as a method of information acquisition for accident investigation purposes;

(c) Establish the comparative rate of failures across a range of aircraft incidents and accidents;

(d) Identify the strategies that pilots engage in response to information from weather radar displays; and

(e) Recommend strategies to improve the design of instructional systems in relation to weather radar displays.

2. Weather-Related Decision-Making in Aviation

Despite significant advances in the technology related to the prediction and reporting of weather conditions, the safety and efficiency of a flight remains dependent upon the pilot making an accurate and expeditious decision concerning the impact of the conditions reported. These so-called ‘weather-related decisions’ remain the province of the operator and, therefore, are subject to the vagaries of human performance.

Errors in relation to weather-related decision-making are difficult to establish for a number of reasons, not least of which is the fact that a significant proportion of these accidents, especially in general aviation, result in fatalities. For example, in the 1999 calendar year, the National Transportation Safety Bureau (NTSB 2003) recorded that, of the 106 general aircraft accidents involving Instrument Meteorological Conditions, 54.7% resulted in fatalities.

The rate of weather-related decision errors is also difficult to establish due to the process of summarising aircraft accident and incident statistics amongst investigative authorities. In many cases, it is not clear whether aircraft accidents or incidents that occurred in poor weather were due to poor decision-making on the part of the pilot or due to some other factor, such as mechanical failure. The result is a possible underestimation of the significance of weather-related decision errors in aircraft accident and incident causation.

2.1 Weather Radar and Weather-Related Decisions

In addition to weather reports and forecasts, the pilots of advanced technology aircraft now have available, weather radar systems that display a vast array of weather-related information in real-time. It is assumed that the provision of this information has the potential to improve weather-related decision-making by enabling pilots to recognise changes in the weather conditions at a relatively early stage of the flight and thereby take appropriate action. However, it is not clear precisely how or when these types of decisions should take place to safeguard the integrity of the aviation system.

The experience of pilots in commercial environments is one of safeguarding the passengers and aircraft while, simultaneously, ensuring the expeditious arrival of the aircraft at the destination. This balance between safety and efficiency is particularly evident in relation to decisions about weather. The difficulty associated with weather conditions is that, despite an increase in the amount of information available to pilots, notions of severity and the extent of the impact of a particular weather pattern, remain both uncertain and dynamic.

Reference to the interpretation and use of weather radar systems remains conspicuously absent from the vast majority of aircraft accident and incident reports involving weather. Arguably, this is due to the fact that the use of the weather radar display was not implicated in the accident or incidence sequence, and/or that the weather radar was simply not identified as a significant factor in the occurrence. The issue becomes slightly more complex when issues pertaining to design and training are considered in relation to weather radar displays.

2.2 Weather Radar: Design and Training

There are a number of different models of weather radar that are available commercially, and each functions by interrogating the moisture content in the atmosphere. Where there is significant moisture content, the information is displayed in a reddish hue, consistent with the Anglo-Celtic population stereotype of danger. Where the moisture content is low, the information is displayed in green or blue. Progressive changes between the blue/green and red hues indicate progressively different levels of moisture content (Barr, 1993).

The level of moisture in the atmosphere can be associated with a level of meteorological activity so that light rain may be associated with visibility of greater than five nautical miles, while severe rain may be associated with visibility of between 0.5 and 1 nautical mile. Further, the different levels of rain activity are often associated with different levels of turbulence. In combination with the shape of the distribution of rainfall and the speed and direction of the movement of the weather activity, it is possible to derive the type of meteorological activity that is likely to be experienced. However, it is important to note that these conclusions are typically derived from the information displayed and are not necessarily directly interpretable from the instrument.

While there is a degree of standardisation associated with the design of different weather radar displays, access to different types of information within different displays is managed using a range of devices from push-buttons to knobs, and the information is normally arranged in the form of ‘pages’ through which the user scrolls to obtain the appropriate information (Barr, 1993; National Research Council 2002). The interpretation of the display is normally assisted by a legend that is either available during training or which may be displayed on the instrument, together with the radar ‘paint’.

Training in the use of weather radar displays typically consists of exposure to manuals, computer-based activities, and/or video recordings. The information to which operators are exposed during training is generally limited to the operation of the system and the basic interpretation of the information. Organisational procedures will dictate subsequently, the process through which this information is interpreted in terms of its impact on the course of an operation.

The focus on the development of declarative (factual) knowledge during training is useful insofar as it provides a context for the learning process. However, there is strong evidence to suggest that optimal performance in the use of displays is dependent upon the acquisition and application of an appropriate skill-base. The extent to which exposure to declarative knowledge facilitates the acquisition of skills is a matter of some debate, and recent research suggests that active participation in the performance of a task is necessary for the development of skilled performance (Wiggins & O’Hare 2003a).

To identify the skills that are likely to be necessary for the efficient and effective use of weather radar displays, it is important to establish a framework against which the use of the weather radar displays can be assessed. One mechanism for the development of this framework involves integrating a model of information processing with the characteristics of experts. By establishing the characteristics of expert performance across a range of domains, it becomes possible to generate some fundamental principles of performance that occur, irrespective of the domain in which they are engaged.

2.3 Empirical Perspectives of Expertise

From an empirical perspective, isolating the origins of expertise is probably best achieved through a longitudinal analysis in which the learner is examined at frequent intervals during the transition from novice to expert. Unfortunately, this type of approach requires a vast investment in both time and resources to reach a satisfactory conclusion. Moreover, this type of approach is often conducted in an applied environment that may lead to difficulties in experimental control.

On the basis of these arguments, researchers have typically avoided the longitudinal approach in favour of a cross-sectional methodology in which comparisons are made between individuals at different stages during the transition towards expertise. In recent years, there has been a plethora of such expert-novice comparisons in a number of fields as diverse as squash (Abernethy, 1990), map reading (Gilhooly, Wood, Kinnear & Green, 1988), medical science (Patel & Groen, 1991), physics (Adelson, 1981), troubleshooting (Johnson, 1988), fire fighting (Klein, 1989; Klein & Klinger, 1991), political science (Voss & Post, 1988), sonar operation (Kirschenbaum, 1992), and aviation (Wiggins & O'Hare 2003b). This contributes further to the vast collection of previous research involving the identification of features that differentiate experts from novices.

Undoubtedly, the primary impetus for this type of research has been the drive to develop more efficient and more effective training systems. The underlying assumption is that, through the characterisation of expertise, training programmes might be developed that emphasise these expert-related features amongst novices (Abernethy, 1994; Edwards & Ryder, 1991; Olsen & Biolsi, 1991; Patel & Groen, 1991). The aim is to encourage novices to adopt these practices and, thereby, facilitate the relatively rapid progression towards expertise (Abernethy, 1990).

A secondary, though no less significant impetus for expert-novice comparisons involves the development of expert systems. In this case, the aim is to characterise the cognitive structure of expertise in order to construct expert systems that mimic and, in some cases, improve upon expert performance, particularly under conditions of uncertainty (Cooke & McDonald, 1986; Shanteau, 1988). However, this process has met with only limited success due to the inherent difficulties associated with eliciting procedural knowledge from experts (Hayes-Roth, Waterman & Lenat, 1983).

Irrespective of the motivation associated with expert-novice comparisons, a great deal of knowledge has been acquired concerning the differences that exist between the two groups. Consistent with a singular view expertise, much of this research has been devoted to isolated analyses of task-related performance, psychomotor skills, cognitive skills, or affective processing. Rarely has a pluralistic approach been adopted, and few researchers have endeavoured to synthesise the outcomes of research arising from different domains (for exceptions, see Glaser & Chi, 1988; Shanteau, 1988). Even fewer researchers have attempted to examine the implications of such research for a theoretical analysis of the transition towards expertise. The result is a somewhat limited view of the differences that exist between experts and novices.

In spite of the lack of detailed empirical data concerning the transition from novice to expert and, therefore, the origin of expertise, some conclusions may be drawn through an analysis of the common features that have been reported as characterising the nature of expertise. Table A.1 (see Appendix A) provides the details of a number of expert-novice comparisons and the features that have been identified as associated with expert performance within a variety of domains. This list provides the basis from which common cognitive features might be identified that characterise the nature and, through inference, the origin of skilled performance.

In reviewing the list provided in Table A.1, it is important to note that various authors may have utilised different labels for what are essentially, identical cognitive constructs. Moreover, the characteristics of expertise identified by various authors generally relate to distinct features of a similar cognitive process such as perception, information acquisition or option generation. Figure 1 provides a diagrammatic summary of the perceptual, cognitive, and behavioural features identified in Table A.1. This three-stage approach is consistent with a generalised, three-stage model of information processing, comprising a perceptual element, a response selection element, and a response execution element (Wickens & Flach, 1988). It should be interpreted as an explicit, rather than an implicit, model of the nature of expertise.

2.3.1 Expertise Defined

On the basis of Figure 1, expertise can be conceptualised as a complex process involving a number of perceptual, cognitive and behavioural features depending upon the nature of the task under investigation. For example, in some situations, expertise appears goal-directed and efficient, while in other circumstances, it appears variable and individualised. In the case of option generation, expertise is characterised as involving either forward reasoning or a recognition-primed approach, and may, in some cases, require a detailed process of plan generation and the integration of information towards a solution. Finally, expertise appears to be moderated by a detailed mental representation of the problem environment, combined with a structured and task-specific knowledge base, and a range of memory retrieval strategies that facilitate the recall of information efficiently and accurately.

The outcomes of the synthesis of empirical research outlined in Figure 1 are somewhat consistent with theoretical notions of expertise, particularly in terms of the variability of performance. For example, Wickens, Gordon and Liu (1998) suggest that human information processing can occur at a number of different levels of cognition, depending upon the nature of the task. Three stages of information processing are hypothesised that are consistent with the distinction between skill-based, ruled-based, and knowledge-based behaviour proposed by Rasmussen (1983). At the skill-based level of performance, responses are relatively rapid and are directed towards to the application of solutions. However, in the case of a less familiar situation, information tends to be processed at a higher level of cognition, and a process of mental simulation and hypothesis generation occurs in which a single option may be examined in detail until it is either confirmed or rejected (Wickens et al., 1998). The success of this approach is dependent upon a detailed mental representation of the problem environment, a well organised knowledge-base, and is generally driven by a forward reasoning strategy towards problem resolution. These perspectives are entirely consistent with the analysis of the empirical outcomes that are outlined in Figure 1.

On the basis of Figure 1, the skills that are likely to be required for the effective and efficient use of weather radar systems include the capacity to:

- Recognise spatial patterns;
- Conduct an accurate and efficient situation assessment;
- Conduct an efficient search strategy;
- Encode information so that the information is solution-centred;
- Generate a plan;
- Perceive and respond to inconsistencies; and
- Conduct a mental simulation of expected options and outcomes.

The successful application of these skills is likely to be dependent upon:

- A detailed mental representation of the environment;
- The capacity to develop accurate predictions on the basis of the information available; and
- An extensive task-specific knowledge base.

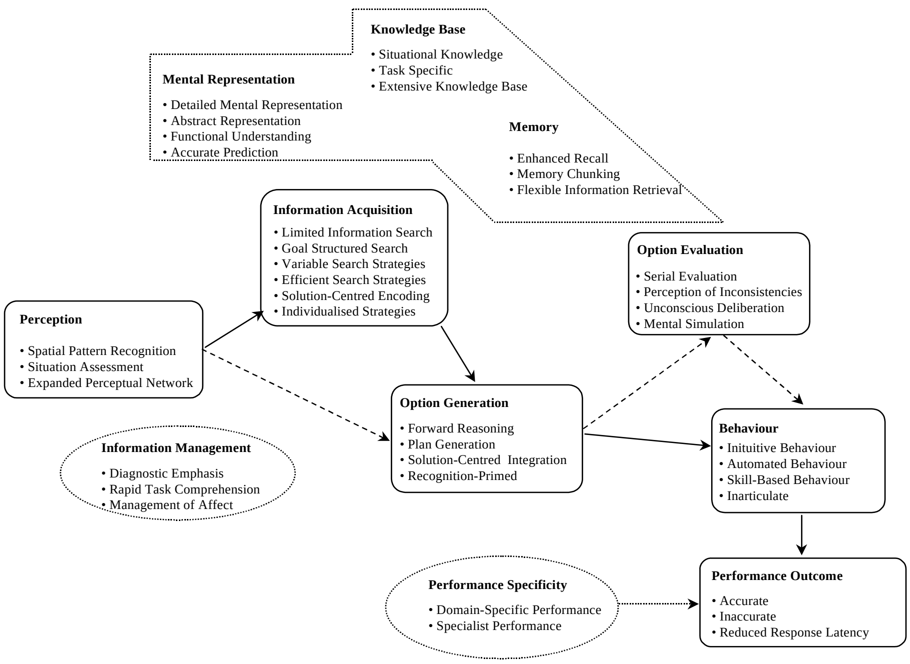


Figure 1. A diagrammatic summary of the perceptual, cognitive and behavioural features of expertise identified in Table A.1. Dashed arrows indicate an alternative information-processing route that may be activated under specific conditions. Information contained within the dotted lines is applicable at all levels of cognitive processing.

From the perspective of human performance, the skills that are necessary for the accurate and efficient interpretation and response to weather radar displays can be regarded as primarily cognitive in nature. These skills are drawn from each of the information processing stages that, theoretically, precede the application of behaviour. However, it is important to note that while these skills are presumed necessary, the fact that an error has occurred in relation to the use of weather radar displays does not necessarily reflect a lack of skill. Indeed, it may be the case that, despite an extensive skill base, the design of the system was such that the successful application of skills was not possible. This interdependence between the design of the system and the capabilities of the user is an important principle associated with the optimal use of weather radar systems.

2.3.2 An Expertise Model for the use of Weather Radar in Aviation

From the perspective of accident investigation, there are a number of factors that might be identified as precipitating a failure in information processing. However, by decomposing human performance into a sequence of activities, albeit cognitive, it becomes possible to isolate the specific features of an error and, thereby, formulate more specific recommendations. While there has been some criticism of such taxonomic approaches to the analysis of human behaviour (see Dekker 2001), it remains a useful strategy to ensure that the opportunity to improve human performance in the workplace is maximised.

A taxonomic approach to the analysis of weather radar-related errors is founded on a clear understanding of the nature of expert human performance (see Figure 2), and serves to highlight specific areas within the system where further investigation is required and where, ultimately, tangible improvements could be made. However, the taxonomy is also particularly useful from the perspective of trend analysis, and overcomes some of the difficulties associated with statistical summaries of aircraft accidents and incidents that involve broad categories. In decomposing and analysing events on the basis of the use of the systems involved, it becomes possible to establish the extent to which design and/or training initiatives need to be improved in respect of the use of these systems.

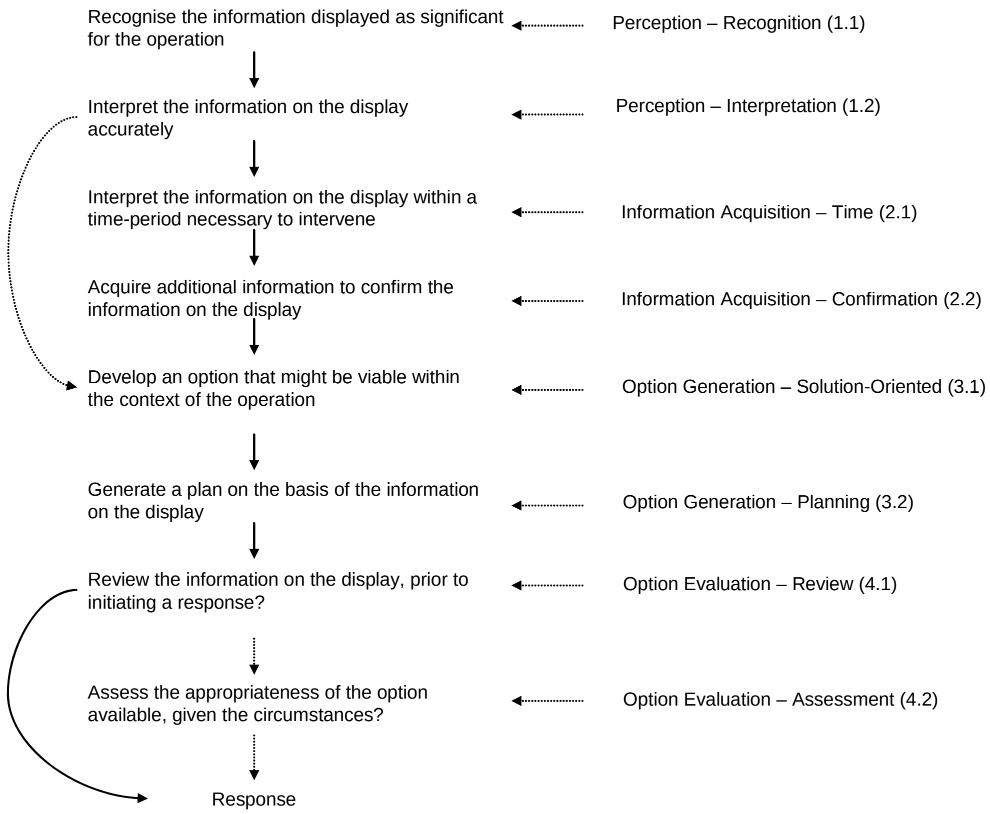


Figure 2. Expertise model (left) and associated taxonomic classifications (right) for the use of weather radar displays in aviation.

2.3.3 The Current Project

The aim of the current project was to test various aspects of the validity of the weather radar investigative taxonomy using a combination of existing aircraft accident and incident summaries, a cognitive interview with expert pilots, and a questionnaire that sought pilots' responses to a number of simulated weather radar displays that might be observed during flight.

3. Study One

3.1 Aim

The aim of Study One was to summarise contemporary aircraft accident and incident records from a number of different perspectives, including the frequency with which weather conditions were involved, and the frequency with which weather radar was cited as either being available or as being a significant factor associated with the occurrence.

3.2 Data Sources

The three primary data sources for the present study were the Aviation Safety Reporting System (ASRS) Database, the Federal Aviation Administration (FAA) Accident/ Incident Data System (AIDS), and the United States National Transportation Safety Board Aircraft Accident and Incident Database. These sources were accessed due to their availability, and the fact that the data encapsulate a relatively large aviation system.

Despite their utility, there are a number of limitations associated with the data sources used in the present study. Most importantly, the information included in accident and incident reports is collected post-hoc, and the conclusions are derived based upon the information available to the investigating authority. As a result, important information may be overlooked and/or may not have been considered relevant to an occurrence. In the particular case of the ASRS database, the information is submitted confidentially by operational personnel and has not necessarily been subjected to the rigours of investigation. Therefore, the information that is ultimately included in this database is often framed within a particular perspective.

Although there are a number of limitations associated with the use of these data sources as the basis for definitive conclusions, the data do, nevertheless, provide a basis for the development of subsequent research initiatives that may be tested within an experimentally controlled environment.

3.2.1 Procedure

In the case of the ASRS Database, a report set was generated that included a random sample of 50 reports that occurred between June 2002 and January 2003, and where there was reference to encounters with weather during flight. Reports were included from both general aviation and airline operations, since many general aviation aircraft in the United States are equipped with weather displays.

Data from the FAA AIDS Database were drawn from a search of the narrative text that included the term ‘weather radar’ for cases that occurred between January 1 1994 and January 1 2004, This ten year period was selected on the assumption that: (a) it would provide a dataset of reasonable size; and (b) it would cover a significant period during which weather radar systems were being developed and installed in aircraft. The search of the AIDS Database resulted in 15 reports of a total of 23305 cases involving both general aviation and airline operations. By contrast, 641 reports were generated from a search of the narrative text that was restricted to the term ‘weather’.

Consistent with the process of data acquisition adopted for the analysis of the AIDS Database, the search of the NTSB Accident Database was restricted to cases that occurred between 1 January 1994 and 1 January 2004. Of a total of 21708 records involving both general aviation and airline operations, 148 were returned using the term ‘weather radar’, while 6158 records were generated using the term ‘weather’.

3.3 Results

3.3.1 Raters

Consistent with all taxonomies, the successful application of the expertise model is dependent upon a clear understanding of the various categories and sub-categories and how these features might manifest within the environment. Consequently, three raters were used in the initial assessment of cases, of whom two had operational experience in the aviation environment. A sub-set of cases was examined to determine the reliability of raters and where there were discrepancies, the case was discussed and an overall classification was resolved. Overall, the three raters achieved a reliability coefficient of 0.76 which is regarded as an acceptable level of reliability, given the difficulties associated with the amount of information available in many of the reports. It is anticipated that greater levels of reliability will be achieved during the process of accident investigation, since investigators will have access to relevant information first hand.

3.3.2 Aviation Safety Reporting System

In interpreting the following results, it is important to note that these data were self-reported accounts and, therefore, they were not subjected to the rigour associated with an aircraft accident investigation. Moreover, the search term used for the extraction of the data was ‘weather’, rather than the more specific search term of ‘weather radar’ that was used for subsequent extractions of data. The use of the broader search term resulted in a distribution in which only 15 of the 50 cases referred to the use of the weather radar during the operation.

Of the 15 cases that referred to the weather radar, five could be described as examples in which the information arising from the system was interpreted, examined, and acted on appropriately. The 10 remaining cases involved failures that could be classified according to taxonomy.

Despite the relatively small sample size and the inherent difficulties associated with the nature of the data, the information that could be derived from the ASRS data was considered a necessary part of a broader examination of the use of the weather radar in aviation operations. Nevertheless, it was necessary to limit the analysis to a purely descriptive level and limit the conclusions that might be drawn.

The analysis of the data was conducted by summarising, initially, the taxonomic category within which the case was coded. The taxonomy comprised four categories including Perception, Information Acquisition, Option Generation and Option Evaluation. A distribution of the frequency with which the case was coded within a particular category revealed a pattern in which Information Acquisition was cited most frequently, and Option Evaluation was cited least frequently.

Table 1. Distribution of the frequency with which ASRS cases were coded within each of the four taxonomic categories.

	Frequency	**Percentage**
Perception	2	20
Information Acquisition	4	40
Option Generation	3	30
Option Evaluation	1	10

To investigate the pattern of results more fully, those categories with the greatest frequency of cases were re-classified into the relevant sub-categories. In the case of Information Acquisition, the two sub-categories related to the time period within which the information was interpreted (Information Acquisition - Time) and whether additional information was sought to confirm the information on the display (Information Acquisition – Confirmation). The two sub-categories of Option Generation included the extent to which the interpretation of the display was oriented towards the development of solutions (Option Generation – Solution-Oriented), and the extent to which a plan was developed on the basis of the information available (Option Generation – Planning).

An inspection of the distribution indicated that, for Information Acquisition, three of the four cases were classified as occurrences in which additional information was not sought to confirm the information arising from the weather radar display. In the case of Option Generation, two of the three cases were classified as occurrences in which the operator/s had not developed a plan in response to the information arising from the display.

While it is difficult to develop conclusions based on the frequency counts associated with these data, it, nevertheless, provides a baseline against which subsequent analyses can be compared. Moreover, it establishes, in some sense, the utility of the taxonomy in its ability to differentiate between the different stages of the perception, interpretation and response to weather radar displays.

3.3.3 Federal Aviation Administration (FAA) Accident/ Incident Data System

Of the 15 reports that were retrieved from the FAA Accident/Incident Data System, only one of the reports referred to the pilot's use of the weather radar in managing a weather-related situation. In this case (20010710018469C), the pilot appeared to use the system appropriately to diagnose and respond to the changes in the weather conditions en-route. Of the remaining 14 cases, three referred to situations in which storm cells or turbulence were not detected on the weather radar, nine referred to a weather radar failure, and in two cases, the role of the weather radar was not specified. On the basis of these data, it was not possible to apply the taxonomy in the diagnosis of failures involving the use of the weather radar.

3.3.4 National Transportation Safety Board Database

Overall, 21 of the 151 National Transportation Safety Board reports that referred to weather radar were suitable for coding using the weather radar taxonomy. The remaining accounts either referred to the weather radar as a peripheral issue in the occurrence, or involved the appropriate and successful use of the weather radar to manage the situation.

A frequency distribution of those reports in which the use of the weather radar could be coded using the taxonomy indicated that 42.8% of accounts were associated with the perception of information, 19% of accounts were associated with the acquisition of information, 23.8% of accounts were associated with the generation of options, and 14.4% of accounts were associated with the evaluation of options in response to the information presented on the display. A more detailed analysis indicates that the sub-categories of the taxonomy that were cited most frequently in the analysis included the recognition of the information presented on the weather radar display, the interpretation of the information perceived on the display, and the extent to which information was sought to confirm the information presented on the display (see Table 2).

Table 2. Distribution of the frequency with which NTSB cases were coded within each of the six taxonomic sub-categories.

	Frequency	**Percentage**
Perception – Recognition	10	58.8
Perception – Interpretation	3	17.6
Information Acquisition – Time	1	5.8
Information Acquisition – Confirmation	2	11.7
Option Generation – Solution-Oriented	0	0
Option Generation – Planning	0	0
Option Evaluation – Review	1	5.8
Option Evaluation – Assessment	0	0

3.4 Discussion

Overall, the results associated with this analysis reveal a number of important issues concerning the nature of weather radar displays in aircraft accident causation. The most important of these issues is the relative dearth of information in the accident and incident narratives where weather conditions were a significant factor and where the aircraft involved were equipped with weather radar displays. This may reflect the fact that the pilots' use of the weather radar was considered, but was subsequently dismissed as a significant factor. However, the lack of information may also reflect an assumption that, unless the weather radar was not used or was inoperable during the event, that the pilots' interpretation and response to the information displayed was sound. The pattern of results that emerged from the analysis of the accident and incident databases suggests that this assumption may not necessarily be accurate.

Given that there were no cases eligible for analysis within the Federal Aviation Administration Accident/ Incident Data System, it was not possible to consider the implications of the taxonomy in terms of these data. However, narratives from the Aviation Safety Reporting System Database and the National Transportation Safety Board Aircraft Database were able to be considered in relation to the application of the taxonomy.

In relation to the categories of the weather radar taxonomy, the data arising from the ASRS database revealed a pattern in which occurrences were associated with each of the four categories at a relatively similar frequency. It is important to note that the sample size was such that the application of additional statistical comparisons was not possible. Nevertheless, the application of the taxonomy to these data did highlight the utility of the approach as a means of further diagnosing the source/s of error associated with weather radar systems.

The sample size arising from the NTSB accident/incident database enabled a more detailed analysis of the data than was possible using the ASRS database. In the case of the NTSB data, a relatively distinctive pattern did emerge in which the majority of occurrences were associated with the perceptual stage of information processing, either in terms of the recognition of the events as significant, or in terms of the interpretation of the information displayed. The phase that was cited least frequently in the analysis was option generation.

When the data were reclassified according to the weather radar taxonomy sub-categories, it was clear that the recognition of information (perception – recognition) on the weather radar display featured most frequently during the occurrences. The interpretation of information and the confirmation of information arising from the weather radar display featured less frequently in the accounts. Although it is difficult to draw conclusions on the basis of these data, it does provide some evidence to suggest that it is the recognition and interpretation of the information arising from the weather radar displays that are most likely to be associated with aircraft accidents and incidents in which the weather radar is cited as a significant factor.

3.4.1 Limitations

Clearly, there are a number of limitations associated with this stage of the project that constrain the conclusions that can be drawn and the extent to which the outcomes can be generalised. In particular, the relatively sample size and the difficulties associated with the detail of aircraft accident and incident reports are such that the weather radar taxonomy ought to be subjected to further analysis prior to its implementation within the framework of aircraft accident investigation. As part of this process, the next stage of this project involved an interview with subject-matter experts that sought to confirm the classification process that is embodied within the expertise model of the use of weather radar displays in aviation.

4. Study Two

The outcomes of Study One provided a useful, retrospective analysis of the impact of weather radar in accident and incident occurrences. However, these data need to be interpreted with some degree of caution, since the sample is not randomised. Moreover, the reports selected for analysis are the basis of investigations, the outcomes of which may or may not reflect the range of factors associated with an occurrence.

To further investigate the use of weather radar displays amongst pilots, Study Two involved a series of task-related interviews with subject-matter experts. Unlike Study One, the aim of Study Two was to investigate the processes that experts use in acquiring, interpreting, and responding to information arising from weather radar displays. This information was designed to form the basis of a survey to be administered during Study Three.

4.1 Cognitive Interview

Given the nature of the data required for Study Two, a cognitive interview technique was considered appropriate, since it enabled the acquisition of both behavioural and cognitive information pertaining to the performance of a task. The cognitive interview technique has been employed for a wide range of purposes including the development of instructional systems, the analysis of errors, and the design of expert systems (Hoffman, 1987). In essence, however, the cognitive interview is simply a mechanism that enables the acquisition of information from the perspective of the agent (operator). This type of information is presumed to be advantageous in developing systems that either complement the performance of the agent or imitate the performance of the agent.

The process of conducting a cognitive interview begins with the development of a semi-structured interview protocol (Cooke, 1994). This protocol is developed on the basis of assumptions concerning the nature of the performance of the agent. However, the semi-structured nature of the protocol is testament to the fact that the acquisition of information is led by the agent, rather than the interviewer. This is a particularly important principle, since the assumptions made by the interviewer may not necessarily reflect the underlying cognitive structures that determine performance at a particular level of skill acquisition.

Having developed the interview protocol, operators are asked to respond to the questions and the information is examined using one or more of a number of qualitative procedures for data management. The interpretation of information can occur at a number of philosophical levels, from ‘grounded theory’, in which the interpretation is driven by the nature of the data to ‘deterministic’ in which the interpretation is driven by an over-arching theoretical perspective (Hoffman, Crandall, & Shadbolt, 1998).

In the case of Study Two, a deterministic approach to the interpretation of the data was employed, since part of the aim was to further establish the validity of the expertise model of the use of weather radar displays as a basis for the investigation of aircraft accidents and incidents in which weather radar systems may be implicated. Inevitably, this theoretical perspective also influenced the development of the interview protocol that was employed in the study.

4.2 Method

4.2.1 Participants

The participants consisted of five subject-matter experts, each of whom had accumulated greater than 1500 hours using weather radar systems. All of the participants were male, were resident within Australia and held a Commercial Pilots Licence (CPL, 1 participant) or an Airline Transport Pilots Licence (ATPL, 4 participants).

The average experience, as measured by the number of hours accumulated, is summarised in Table 3.

Table 3. Summary of the experience, as measured by the number of hours accumulated, of participants involved in Study Two.

	Mean	Standard Deviation
Total Experience	9448.20	3608.01
Pilot in Command Experience	5606.80	2466.87
Instrument Experience	5190.00	5758.52
Recent Experience (last 90 days)	119.00	80.65
Experience using Weather Radar	5593.20	5036.19

The type of weather radar systems on which the participants had accumulated experience included the Rockwell Collins Multiscan™ (2 Participants[1]), and the Bendix-King RDS81 (2 Participants) and RDS82 (1 Participant). One participant indicated that he was engaged primarily in general aviation operations, while two participants indicated that their most recent experience (within the previous six months) involved short-haul operations, and the remaining two participants indicated that their most recent experience involved long-haul operations. All of the participants operated primarily within the Asia-Pacific Region.

4.2.2 Interview Protocol

The interview protocol was developed using a cognitive interview methodology incorporating the Critical Decision Method (CDM) of information acquisition. CDM is a well established methodology in which participants who are engaged in a cognitive interview are asked to recall a specific instance as the basis for responding to subsequent questions. The main

[1] One participant indicated that he had accumulated experience on more than one system during the six months preceding testing.

advantage associated with this approach is that it is presumed to engage a level of reflection that accesses those cognitive and behavioural processes that underscore the performance of the operator.

In the current study, the participants were asked to recall a situation in which they were forced to rely on weather radar to assist the management of a flight. The intention of this directive was to reduce the scope of the interview and provide a concrete foundation for the acquisition of information. Consistent with this perspective, each of the subsequent questions was framed so that it related to the situation that participants were asked to bring to mind.

The questions that comprised the interview protocol were based on the features of the expertise model of the use of weather radar displays that was developed as part of Study One. Questions pertaining to each feature were developed to acquire information in relation to each of the underlying cognitive features, including:

Perception

- Situation Assessment
- Expanded Perceptual Network

Information Management

- Diagnostic Emphasis
- Task Comprehension

Information Acquisition

- Limited Information Search
- Goal Structured Search
- Search Strategies
- Accuracy

Option Generation

- Forward Reasoning
- Plan Generation
- Mental Representation

Option Evaluation

- Serial Evaluation
- Analogous Situation
- Task-Specific Knowledge

4.2.3 Procedure

The participants were self-selected in response to advertisements placed in a range of aviation-related media. Initially, they were sent a copy of the information sheet and consent form and were asked to nominate a time during which they would be available for a telephone interview. This method of interview was used to ensure that geographic isolation did not restrict the involvement of participants.

During the follow-up telephone interview, participants confirmed their consent to participate in the study and the interviewer read a standard briefing statement. Having determined that there were no further questions, the interviewer began by seeking responses to a series of demographic questions that sought to establish the age, gender and experience of participants.

Prior to the cognitive interview questions, participants were asked to recall an incident in which they were forced to rely on weather radar to assist the management of a flight. They were advised to use this experience as the basis for their subsequent responses to the interview questions.

The interview questions were asked in order, although the protocol was semi-structured so that additional information could be sought if necessary. Having completed the interview, participants were thanked for their participation and the taped recordings were transcribed.

4.3 Results and Discussion

Given the relatively small sample size, and bearing in mind that the intention of this study was to provide the basis for a more detailed survey, it was considered appropriate to examine the data derived from the interviews from a descriptive perspective. In describing the nature of the data, it was considered appropriate to interpret the information in terms of the themes that emerged from the accounts of participants. Thematic analysis is a relatively well established technique for the interpretation of qualitative data that provides an opportunity to capture the ‘richness’ of data, while enabling comparisons to be made across a cohort (Purkitt & Dyson, 1990).

4.3.1 Perception

For the majority of participants, it appeared that those features of the situation that indicated the requirement for a reliance of weather radar were related to both the nature of the situation pre-flight, and the characteristics of the environment during flight. As might be expected, the main pre-flight indicator consisted of the weather forecast. However, at least one participant considered the nature of the flight itself as an indication of a potential reliance on weather radar. Specifically, traversing the tropic zone was noted as a particular route that might increase the likelihood of a reliance on weather radar.

During flight, the indicators consisted of meteorological reports, reports from other operators, and the returns from the weather radar. One participant suggested that visual reference from the cockpit also represents an important indicator: “…of course, once you are airborne, and you are looking outside the window…you can see that there are storm cells” [S3]. In combination, these indicators represent not only an indication of the requirement for the use of weather radar, but also provide a basis for the interpretation of the information derived from weather radar displays.

To examine the issue of interpretation in more detail, participants were asked to explain those specific features of the indicators that they were seeking when establishing the need to use weather radar to manage a flight. The responses revealed a number of issues, including the rate at which the weather conditions were changing over time, the direction of movement of the weather systems, the type of clouds associated with the weather system (cumulus, nimbus and lightening in particular), the proposed track of the aircraft (reference to seasonal variations in conditions), and whether it was possible to "visually see around [storm cells]" [S3]

4.3.2 Information Management

Having established the requirement for some level of reliance on weather radar, participants were asked to reflect on the case that they reported at the outset and explain the most important issues that they were seeking to resolve as part of their interpretation of the display. At a preliminary level, the participants indicated that the most important issue was the extent to which radar returns were actually being displayed by the system. Once it was clear that the system was functioning appropriately, participants indicated that they sought to establish the "extent of the build-ups" [S5], "the locations of the cells" [S1], "where you had to divert off…track" [S2], and the "intensity, distance, size, colour and adjacent systems" [S4]. The use of these features as the basis for diagnosis appears to be based, at least in part, on the previous experience of the operators in terms of their ability to interpret the information accurately and reliably.

The importance of operational experience as the basis for the interpretation of the information arising from weather radar displays is further illustrated in participants' explanations of the process of developing a mental picture of the situation. In the case of one participant, the development of an accurate mental picture occurs "…from seeing where the weather is, or also just seeing where it wasn't" [S5]. Other participants referred to manipulation of the 'tilt' and 'distance' functions as a basis for establishing the dimensions of the weather system. Finally, a number of participants referred to the relationship (location and movement) between the pattern of cells against the aircraft track as a basis for establishing a mental picture of the situation.

4.3.3 Information Acquisition

In identifying the most appropriate response to the information arising from the weather radar display, it appears necessary, initially, to clearly establish those features of the display that are likely to impact the development of solutions. When asked to describe the features of the display with which they were most engaged, the participants generally referred to the location of the cells relative to the track. However, there was also reference to "the intensity, the extent of the build-ups, the size of the thunderstorm, the combination of that plus the wind reported at the current level or a different level" [S5]. One participant was more specific and referred to "…particularly red or magenta or even yellow to indicate that there are areas of turbulence" [S1].

The association between the features of the display and weather behaviour was highlighted by one participant who noted that:

Black is no return, then it goes green, yellow, red, magenta. And so, if you are looking at a storm for instance and it goes green, yellow, red, magenta very quickly over a very short space of time, then you know that …you've got very, at least, very heavy rainfall, possibility of a lot of turbulence associated with that...You are also looking for the size of a cell…the height of the cell. Keeping an eye on say it's direction of movement, I mean you have an idea of the direction of movement from your forecast and things, but also if it's not on your track, it is going be on your track in 10 minutes for instance, things like that. [S3]

The significance ascribed to the information arising from the weather radar displays was considered by the participants at a number of different levels. One participant responded by considering the information in terms of the fuel available for the flight. In this case, the information available would be interpreted as more or less significant, depending upon the capacity to 'hold' the aircraft or divert to an alternate as necessary [S5]. At another level, the significance of the information was interpreted by one of the participants in terms of the intensity of the return. Where returns are interpreted as 'intense', there is a relatively greater need to consider a diversion [S2]. Finally, two participants considered the significance of the information in terms of the shape of the weather system and the rate at which it changes over time [S3, S4]. In the case of the shape, there was an implication that 'tails' and 'hooks' could be associated with significant levels of turbulence and that these features could regarded as a warning. Similarly, the rate of change of the weather system was expressed in terms that were more qualitative so that a rapid progression from 'green' to 'magenta' would be expected to be associated with significant levels of rainfall [S3].

In searching for significant information arising from the weather display, one participant referred to an elaborate process in which the range of the display would be reduced as the aircraft approached the destination. He also "maximised the sensitivity" and "adjusted the tilt…to see the vertical extent of it" [S5]. Similarly, another participant "changed the range of the radar regularly [and] tilted the area up and down…getting up to 8+ degrees" [S1]. This level of interrogation of the system is further illustrated by an account from one participant who noted that:

…we would start scanning at 300 miles away…and if you could see a storm at 300 Miles you knew it was a big one…and as you were approaching you lowered the range of the radar beam, so that you could get a better view of it. Then you can use the tilt on the radar, on the radar beam, to determine the heights of the tops of the storms as you approaching them and you can find out the maximum intensity using the tilt. [S2]

Although the information provided by the participants enabled the development of an understanding of the various indicators that are used by different pilots in the interpretation of weather radar displays, it was also important establish the basis of this knowledge. In response to the question pertaining to the acquisition of task-related information, the majority of participants referred to a combination of initial learning and operational experience as the basis of their knowledge. Some of the specific learning strategies that were identified included "studying all the handbooks available"

[S1], "learning about how to use the radar" [S3], and acquiring knowledge "from the material supplied by the manufacturer" [S4].

For the majority of participants, it was this combination of knowledge and experience that provided them with the capacity to establish the accuracy of their interpretation of the information from the weather radar display. As one participant noted "I think that it is just experience" [S5], while another referred to "experience with that radar set and…overall experience" [S2]. Two participants observed that there was no way of knowing that their interpretation of the display was entirely accurate [S1, S3]. However, one participant provided a level of qualification to this statement adding that:

…if you're flying along and you haven't got storm cells per se, but you've just got green areas…not associated with thunder storms, just wide spread areas of rain, sometimes you get an image, where you just have…sort of a fuzzy line that extents out about 20 Miles from you and it never changes…When you got that sort of thing you just think: for what ever reason it is not really giving me a true indication of what's out there…But if you get a nice, good picture of a cell and you can see things, not necessarily directly behind it, because it might be hiding things, but things further or storm cells further away than the one that you actually paying attention to, then you can be pretty sure you have got a very good picture of what is going on in front of you. [S3].

Taking a longer term perspective, one of the participants referred to the acquisition of experience as involving:

…two ways: By observing the system in daylight, and compar[ing it] to the radar return, and also by accidentally flying into turbulence around systems…going through them or not skirting them enough. For example, you might look at something that looks benign on the radar or doesn't look too bad on the radar, so you don't go around it very much and you hit into quite a bit of turbulence [S4].

4.3.4 Option Generation

Once the process of information acquisition had been explored, the participants were asked to consider the options that were available in relation to the case that they recalled at the outset of the interview, and how these options were generated under the circumstances. In establishing the goals associated with the generation of options, one participant noted that: "you are trying to achieve a successful landing at the destinations where passengers pay to go" [S5]. However, in achieving this goal, there was also reference to a process of cost-benefit analysis in which safety, comfort, and the potential for delays were considered [S5]. Consistent with this view, other participants referred to the need to consider the safety and comfort of passengers while managing issues such as fuel and potential delays [S2, S3]. The remaining participants considered the issue at a more specific level referring simply to the need to avoid adverse weather conditions [S1, S4].

In establishing the range of options that were available under the circumstances, a number of participants adopted a detailed analytical strategy that reflected a considerable degree of risk assessment. For example, one participant described a process of:

…scanning out in front of you…further out and backwards…if you turn the aircraft left or right 20 degrees for a brief while and scan out there, you get a picture, because, of course, the radar image itself I think goes out …I think you have a 45 degree arc either side the centre. Of course, if you want to have a look at more than 45-degrees left or right, you've got to actually physically turn the aeroplane…But also you have got things like the radar controllers…they can give you some help with storm avoidance. But they're also keeping you clear of traffic and updating you on the location of other aircraft. So you know…[you are not] going to give yourself a hazardous situation by flying into another aircraft [S3].

In another case, a participant reported that:

…when we departed XXXX…and flew down to XXXX, we knew the weather was fine in XXXX for return. The weather forecast for XXXX was fine as well. So, you are looking for other options pre-departure in XXXX and also en-route. So, that you are looking at all the various options before you actually got there and even picked up on weather radar …It's just experience. You just rely on as many other inputs is possible, whether it is talking to Air Traffic Control to see what other aircraft are doing diversion- wise or whether they're getting into XXXX or not…talking to the company to see whether aircraft were departing, what the weather was like and the likelihood of getting...and then also what their preferred option was…whether it was to turn around mid-way to XXXX and go back to XXXX or whether it was to go down to XXXX, hold if we get in, and if we can't get in, return to XXXX. [S5]

To further examine the process of option development, participants were asked to specifically identify those features of the weather radar display that they had considered. While some participants sought to identify areas of potential hazard, such as "any area from yellow through to magenta" [S1] and "looking at the poor weather" [S5], others appeared to search for areas where the hazards was lesser, such as "nice gaps in the radar cells" [S3] or "an alternative route" [S4]. This distinction is possibly a reflection of individual experience in which some pilots are oriented towards the avoidance of poor weather as a goal, while others are focussed on operational requirements to negotiate a route to ensure the safety and comfort of passengers.

4.3.5 Option Evaluation

The evaluation of options is an important part of the process of ensuring the accuracy of outcomes, but it also appears to be quite difficult. At least two participants noted that they did not know what was the best option under the circumstances [S1, S3], despite the information available. Other participants referred to their experience as the basis for determining the most appropriate option under the circumstances [S5, S4].

In identifying the options that were not appropriate under the circumstances, one participant reported that: "you can just see it across the track and just know that you have to move" [S4]. At a more general level, another participant noted that "…you just try and look at the whole picture of what is out there in front of you" [S3]. Finally, one participant used a combination of the known limitations on the flight (fuel) and the information derived from other sources such as air traffic control [S5].

4.3.6 Analogous Situations

It was clear that the participants had developed a repertoire of experiences in relation to the use of weather radar that they drew upon to assist their performance. In one case, a participant described, in some detail, a situation in which the aircraft was struck by lightning, despite the fact that they were flying between cells. There was an implication that this was unexpected, despite the assiduous use of weather radar to navigate around the cells, and that this case had resonated for a number of years [S5]. Similarly, another participant referred to the general experience of having to divert around storm cells or lines of cells as part of operations in particular parts of the world [S3]. In each case, the remaining participants were able to recall situations that were analogous to the situation that they described at the outset of the interview.

4.3.7 General Comments

One of the most significant themes that emerged as a result of the interviews was the apparent lack of operational training and experience in the use and interpretation of weather radar. One participant noted that "…there is no real training available other than books. When I got in an aircraft with weather radar on it, I got all the books that I could study…and [it] left me a little lost" [S1]. Consistent with this view, another participant observed that: …a lot of people never had a proper training course on it. They just say turn the radar on and set it as is" [S3]. It should be noted, however, that this was an observation and was not the personal experience of this participant.

Confirming the difficulties associated with training in the use of weather radar systems, there appears to be an emphasis on "learning by experience on the line" [S4]. One participant reported that:

…when you do simulator training it's very rare…[that you get] returns on the weather radar and you have to deviate around weather…You might be given a forecast or you might be given a ATIS saying what the weather is, but there is no real training about… or there is no inclusion of weather on a weather return during training or during normal flight cyclical checks when you do a sim. [S5]

4.4 Conclusion

The questions that were included as part of the cognitive interview were designed, in part, to examine aspects of the expertise model of the use of weather radar displays. This model provided a structure for the subsequent discussion and a basis for the development of the weather radar survey in Study Three.

Overall, the results of the interviews provided a reasonably consistent picture of the use of weather radar displays within commercial aircraft operations. It was generally agreed that timely responses to weather radar system information are dependent upon an awareness of the requirement to use weather radar at an early stage of the flight. Some participants referred to an expectation that was developed prior to the flight (weather forecast), while others used the information derived from in-flight sources such as the radar itself, other pilots, and air traffic control.

In acquiring information pertaining to the nature of the situation, the participants considered the colours on the weather display, the distribution of the 'paints' (vertical and horizontal), and the implications in terms of the track of the aircraft. It was clear that some pilots saw their goal in relation to the weather radar display as primarily avoiding areas of potential turbulence, whereas others saw it as negotiating a route to ensure the efficiency of the flight while maintaining the safety and comfort of passengers. Although there is no direct evidence, this distinction in goal-orientation may impact the interpretation of weather radar displays, particularly when the information depicted is relatively ambiguous. This is an issue that is examined in greater detail in Study Three.

The accurate interpretation and timely response to the information arising from weather radar displays appears to be dependent upon a combination of experience and training, although the former appears to be perceived as most significant. This experience appears, for some pilots, to have resulted in a series of relatively clear 'rules-of-thumb' concerning the interpretation and response to weather radar displays. For example, the rapid change of colours represents an important indicator concerning the strength of the turbulence that might be expected. Other rules of thumb related to the use of the tilt and range to establish the vertical and horizontal displacement of the weather system.

Despite the fact that experience appears to be perceived as an important determinant for the successful use of the weather radar display, all but one of the participants referred to an apparent lack of experiential training in the use of weather radar.

4.4.1 Limitations

By its very nature, this study was designed to be descriptive in nature and inform the development of the survey to be included as part of Study Three. The conclusions derived are based upon a relatively small sample and, therefore, should be interpreted with some degree of caution. Nevertheless, the participants were all very experienced, both in terms of the operational qualifications and in terms of their direct experience with the use of weather radar displays. Moreover, there was a degree of consistency in the responses.

5. Study Three

Study Three was designed to examine the extent to which some of the issues raised as an outcome of the preceding research were evident across a broader sample of pilots. In particular, the outcomes of Study One suggested that errors associated with the use of the weather radar were associated with the recognition and interpretation of information, rather than the generation or evaluation of options once the information had been acquired.

Where Study One focussed on the outcomes of weather radar-related errors, Study Two was designed to examine the process of managing weather radar-related information prior to, and during flight. Study Three involved the development and distribution of a survey, the aim of which was to validate the outcomes arising from the previous studies.

5.1 Method

5.1.1 Participants

The participants comprised 109 pilots, all but one of whom were male, ranging in age from 18 to 66 ($\bar{X}$ = 43.56, *SD* = 10.47). Overall, 55% of respondents recorded their country of residence as Australia/ New Zealand, 12% resided in North America, 5.5% resided in Europe, while the remaining participants failed to note their country of residence. The majority of respondents held an airline transport licence (90.8%), while 5.5% and 1.8% of respondents held commercial and private pilots licences respectively. They had made an estimated mean 726 (*SD* = 2201.62) in-flight decisions and their mean operational flying experience in terms of the number of hours accumulated is listed in Table 4.

Table 4. Summary of the experience, as measured by the number of hours accumulated, of participants involved in Study Three.

	Mean	*SD*
Total Experience	10096.44	6270.84
Pilot in Command Experience	6998.76	8561.29
Instrument Experience	4032.75	4616.73
Cross-Country Experience[2]	7592.58	5762.14
Recent Experience (last 90 days)	138.79	76.43
Cross-Country Recent Experience	121.24	81.41
Experience using Weather Radar	6279.10	5614.09

2 Defined as in excess of 20 nautical miles from an airport.

5.1.2 Materials

The materials for Study Three consisted of a survey that was distributed both in a paper-based form and through the world-wide web (see Appendix C). The aim of this strategy was to increase the sample size and ensure a broad-based response.

The survey comprised three sections, the first of which sought participants' demographic characteristics, including their age, gender, country of residence, operational flying experience and their experience using weather radar systems. The second section of the survey was designed to establish the extent to which the respondents had made an error involving the use of a weather radar display during the six months preceding the completion of the survey. Participants were asked to reflect on the incident and respond to a series of questions that related to the nature of the incident and the consequences.

The final section of the survey consisted of a series of 12 simulated, static weather radar displays. Each of the displays depicted a weather radar return in which different levels and patterns of precipitation were evident (see Figure 3). An example display and a legend were also included as an aid to the interpretation of information. For the scenarios in the paper-based survey, the following conditions were specified for each of the 12 scenarios: The aircraft is currently heading 340^{o} at FL230. The wind is currently 250^{o} at 22 knots and the tilt is set at 0 degrees. For the web-based survey, the tilt was set at -1.5 degrees. This difference was designed to establish whether the level of tilt specified impacted the responses of participants.

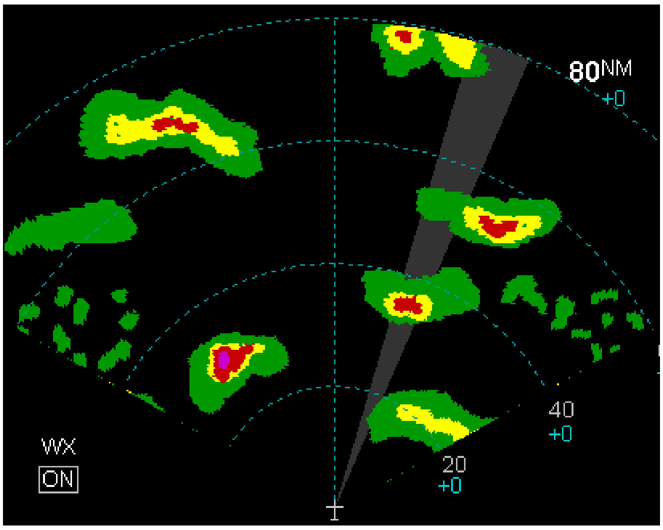


Figure 3. An example of one of the weather radar displays to which pilots were asked to respond in Study Three.

The participants were asked to review each of the 12 simulated displays and rate their confidence in being able to continue the flight for 80 nautical miles at the present track and altitude. They were also asked to indicate the features of the displays that led to this conclusion and rank the level of turbulence, updrafts and downdrafts that might be expected. No time limit was prescribed for the completion of the survey.

5.1.3 Procedure

The distribution of paper-based surveys occurred though a direct mail-out to pilots at a major regional airline, and to respondents to an advertisement in the Civil Aviation Safety Authority's Flight Safety magazine. Prospective participants received a copy of the questionnaire, an information sheet, and a reply-paid envelope in which to return the completed questionnaire. They completed the survey in their own time, there was no inducement, and participants were advised that completion of the survey was both voluntary and confidential.

The survey was also available for completion through a website that displayed the survey in a form consistent with the paper-based survey. Participants were directed to the website through a series of on-line advertisements to a range of organisations. Prior to completing the survey, participants were asked to read an information sheet to ensure that they were familiar with the expectations of the task[3]. Having read the information sheet, participants were able to complete the survey by entering responses using their computer keyboard. Data were automatically recorded in a spreadsheet once the various sections of the survey had been completed.

5.2 Results

The data arising from the survey were examined from a number of different perspectives to determine: (a) Pilots' use of weather radar displays; (b) The types of errors that might be associated with the use of weather radar displays; and (c) Those features that are associated with the successful interpretation and management of information from weather radar displays.

5.2.1 The Use of Weather Radar Displays

It was clear from the responses that the majority of pilots rely on the weather radar display at all times (35%) or the majority of the time (53%) during the course of a flight. Similarly, the majority of pilots expressed trust in the accuracy of weather radar displays at all times (22.4%) or the majority of the time (70.6%). Nevertheless, 59.5% of respondents indicated that they have experienced situations in which the information displayed on the weather radar display was incorrect. This suggests that, although weather radar systems may, at times, be perceived as erroneous, they are generally regarded by the pilot population as a reliable piece of equipment on which they rely significantly for the safe and efficient conduct of the flight.

5.2.2 Weather Radar-Related Error Analysis

Study Three was designed to examine both the nature of errors associated with the use of weather radar displays and the elements of successful performance. In the case of errors, the participants were asked to recall an incident within the

3 It was also possible to print and retain the information sheet.

six months preceding the completion of the survey, during which they had committed an error associated with the use of a weather radar display. Overall, 53.2% of respondents indicated that they had committed an error associated with the use of a weather radar display. Of these respondents, 77.5% were the pilot flying at the time during which the error occurred.

A distribution of the stage of flight during which the error occurred indicated that errors were most likely during the cruise stage of the flight (see Figure 4). According to the respondents, 32% of the errors reported were recognised in less than one minute of their occurrence, while a further 58% of the errors reported were recognised within one to five minutes of the occurrence. Of the remaining errors reported, 8% were recognised between six and 15 minutes following the occurrence, and 2% were recognised greater than 15 minutes after the occurrence.

In reflecting on the errors that they reported, the respondents were asked to indicate which of a series of statements best reflected the nature of the error that occurred. These statements corresponded to the elements of the expert model of the use of weather radar displays that was developed as part of Study One. A frequency distribution of the responses indicated that the majority of errors (42%) were perceived to be associated with the failure to interpret accurately the information arising from the weather radar display (See Table 5).

In identifying that an error had been made, pilots reported a number of indicators including turbulence, the penetration of cumulo-nimbus cloud, icing, passenger discomfort, difficulty in controlling the aircraft, a requirement for a missed approach, windshear, and visual confirmation of cloud formations, often behind ‘shadowing’ weather. These indicators are broadly consistent with the indicators that were identified as part of the interviews reported in Study Two.

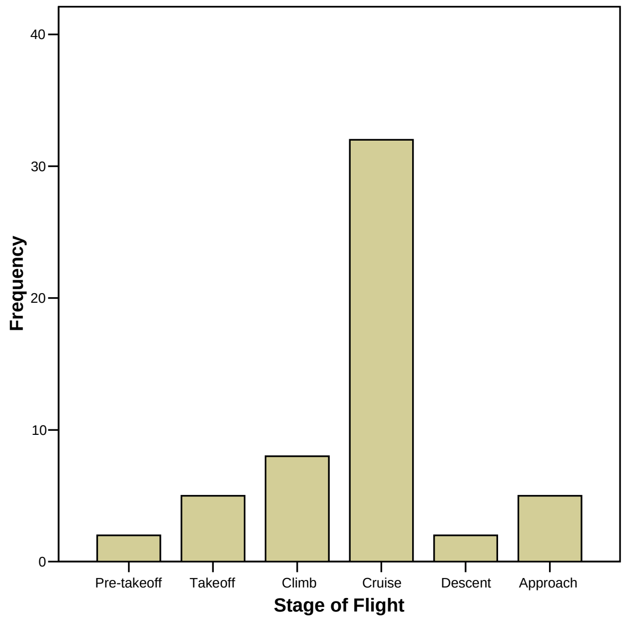


Figure 4. Frequency with which weather radar-related errors occurred, distributed across the stage of flight.

5.2.3 The Interpretation of Weather Radar Displays

Of significant interest in Study Three was the consistency with which pilots interpreted a series of weather radar displays. They were asked to consider 12 static, simulated radar displays and, on the basis of the information available, rate their confidence in continuing the flight on the current track and at the current altitude for a further 80 nautical miles. They were also asked to identify those indicators that led to their conclusion and rate the conditions that might be expected should the aircraft continue along the planned route.

Prior to any comparisons between respondents, it was important to establish whether the two different levels of 'tilt' specified for the paper-based and web-based forms the survey (0 degrees and -1.5 degrees) impacted pilots' interpretations of the displays. To test the impact of the different levels of tilt, a one-way, repeated measures Analysis of Variance was conducted with the type of survey completed as a between-groups factor and the 12 scenarios as a repeated-measures factor. The confidence in continuing the flight for 80 nautical miles without an alteration in track or altitude was included as the dependent variable. The results revealed a non-significant interaction between the type of survey completed and the mean levels of confidence for each of the 12 scenarios, F (11, 924) = 0.46, p = 0.93. A similar, non-significant interaction was evident for the levels of turbulence expected should the aircraft continue along the planned route, F (11, 803) = 0.83, p = 0.61. Combined, these results suggest that the differences in the level of tilt specified for the two surveys failed to impact the responses to the displays and that the participants could be considered a single cohort for the purposes of subsequent analyses.

Table 5. Distribution of the frequency with which respondents coded weather radar-related errors within each of the six taxonomic sub-categories.

	Frequency	**Percentage**
Perception – Recognition	4	10.5
Perception – Interpretation	16	42.1
Information Acquisition – Time/ Confirmation	2	5.3
Option Generation – Solution-Oriented	1	2.6
Option Generation – Planning	1	2.6
Option Evaluation – Review	3	7.9
Option Evaluation - Assessment	4	10.5
Other	7	18.4

Overall, the results pertaining to the confidence of pilots indicated that, for some displays, the responses were relatively consistent, while the responses to other displays were much more variable (see Figures 5a-5l). Therefore, it was concluded that the level of consistency amongst pilots' responses is dependent, to some extent, on the particular characteristics of the displays. For example, in the case of Scenario 12 (see Figure 5l), the vast majority of participants were not at all confident that they would be able to continue the flight as planned. However, in the case of Scenario 8 (see Figure 5h), a similar proportion of participants were very confident that the flight could continue as were not all confident that the flight could continue. This suggests that the information displayed in Scenario 8 was more difficult to interpret than the information presented in Scenario 12.

It should also be noted that the displays for Scenarios 1 (see Figure 5a) and 8 (see Figure 5h) were identical. A Pearson's product-moment correlation confirmed that a statistically significant, moderate correlation exists between the levels of confidence expressed in response to the two scenarios, $r\ (92) = 0.66$, $p < .000$. This result suggests that participants were responding genuinely and consistently in response to the information presented in the displays. Therefore, the differences that appear to exist between the responses must be explained by factors other than random responses to the displays.

To explain the differences between pilots' responses to the scenarios, it was assumed, initially, that their responses, in terms of confidence, would be a reflection of their expectation of the weather conditions that might experienced if the aircraft continued on the present track and at the present altitude. Therefore, respondents were asked to indicate, for each scenario, the extent of the turbulence, updrafts and downdrafts given that the aircraft maintains the current track and altitude for 80 nautical miles.

A comparison between the patterns of results in Table 6 suggests that where participants considered the likelihood of turbulence, updrafts and/or downdrafts as severe, they were also less likely to have confidence that they would be able to continue the flight on the current track and at the current altitude. This is perhaps an indication that respondents interpret the radar returns displayed as a reflection of a particular level of weather activity, and it is this interpretation that determines the confidence that they express in the continuation of the flight.

To establish whether the responses from pilots were a product of individual differences in terms of their anticipation of the turbulence that might be experienced during a particular scenario, a series of Spearman non-parametric correlations[4] was conducted between the scenarios in terms of the perceived turbulence associated with continuing the flight on the current track and at the current altitude[5]. The results revealed a pattern of behaviour in which responses to some scenarios were consistent across individual participants, while responses to other scenarios were more variable (see Table 7).

4 Spearman's correlational analyses were undertaken in preference to Pearson's product-moment correlations due to the ordinal nature of the data.

5 Critical alpha was set to .01 due to the number of analyses conducted.

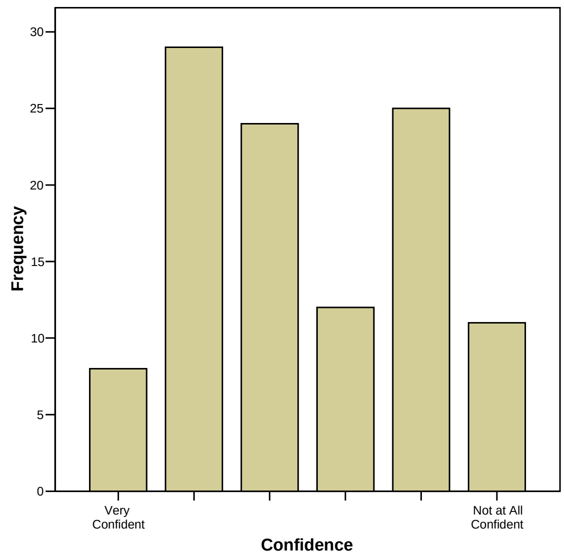


Figure 5a. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 1 for the next 80 miles without an alteration in track or altitude.

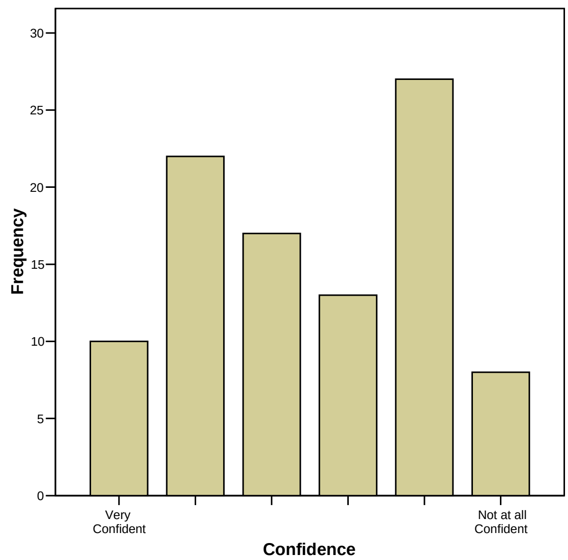


Figure 5b. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 2 for the next 80 miles without an alteration in track or altitude.

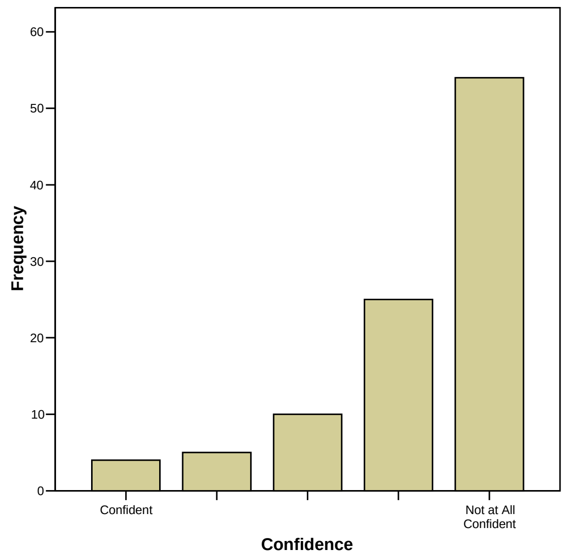


Figure 5c. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 3 for the next 80 miles without an alteration in track or altitude.

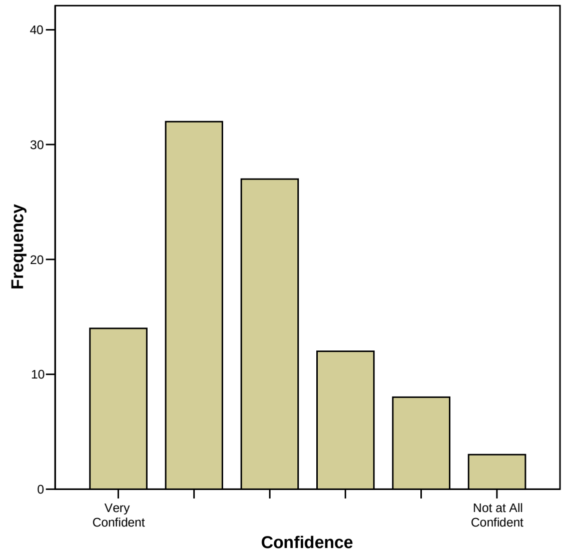


Figure 5d. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 4 for the next 80 miles without an alteration in track or altitude.

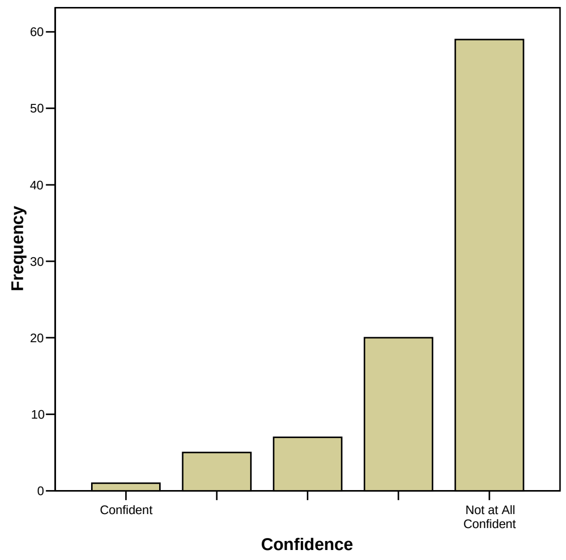


Figure 5e. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 5 for the next 80 miles without an alteration in track or altitude.

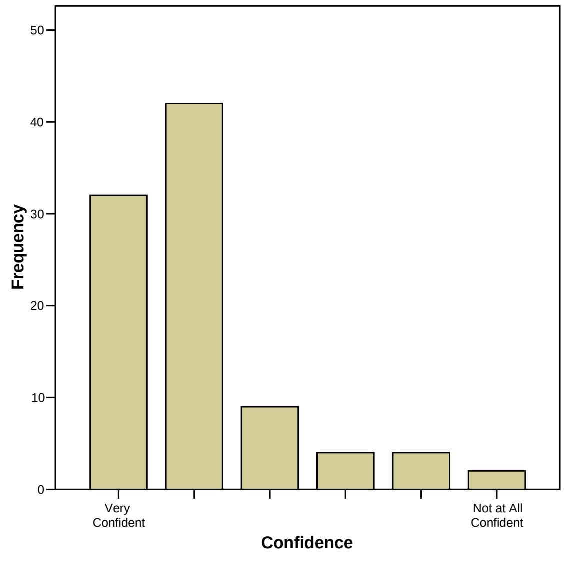


Figure 5f. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 6 for the next 80 miles without an alteration in track or altitude.

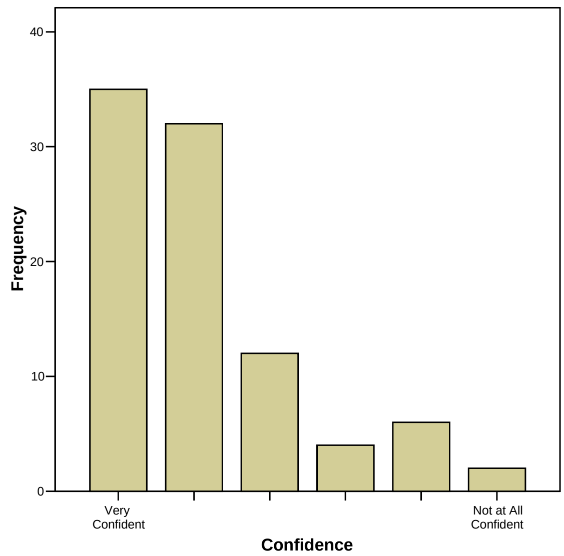


Figure 5g. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 7 for the next 80 miles without an alteration in track or altitude.

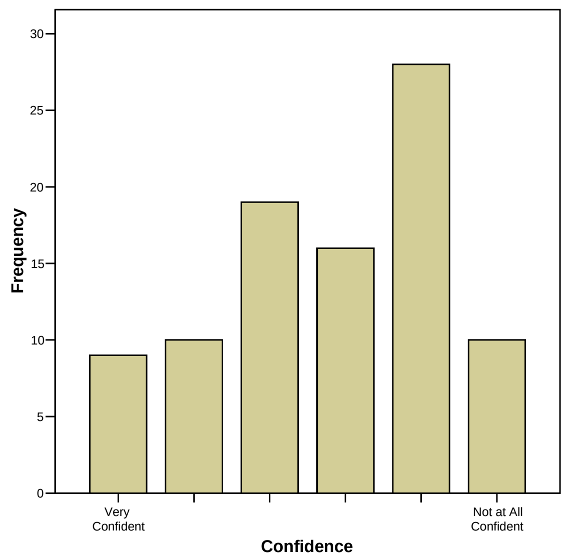


Figure 5h. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 8 for the next 80 miles without an alteration in track or altitude.

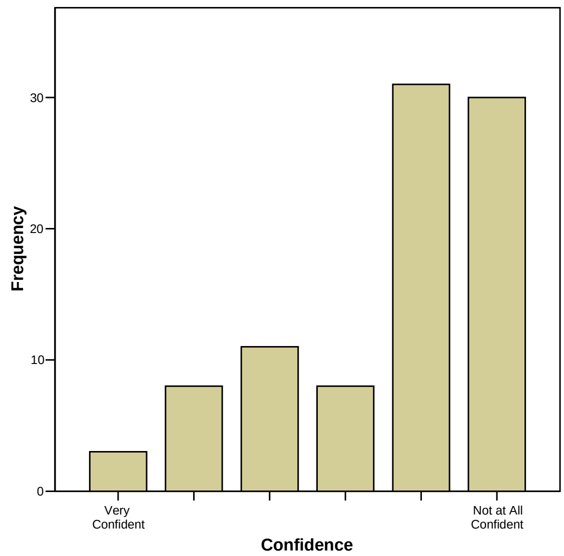


Figure 5i. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 9 for the next 80 miles without an alteration in track or altitude.

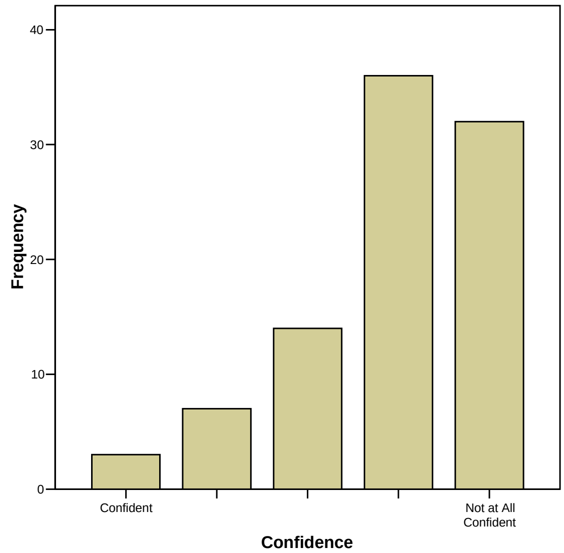


Figure 5j. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 10 for the next 80 miles without an alteration in track or altitude.

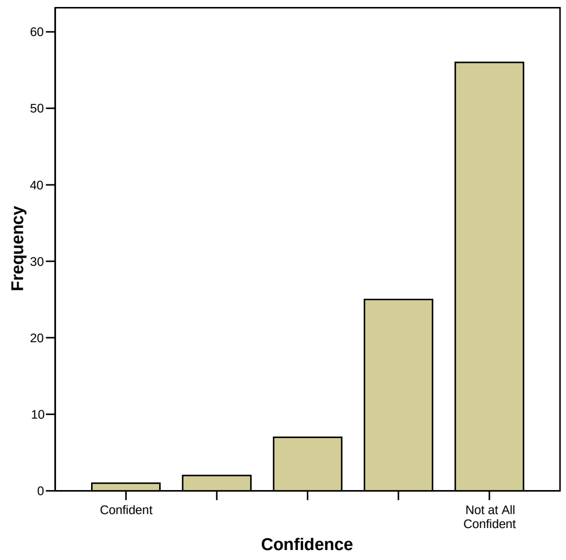


Figure 5k. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 11 for the next 80 miles without an alteration in track or altitude.

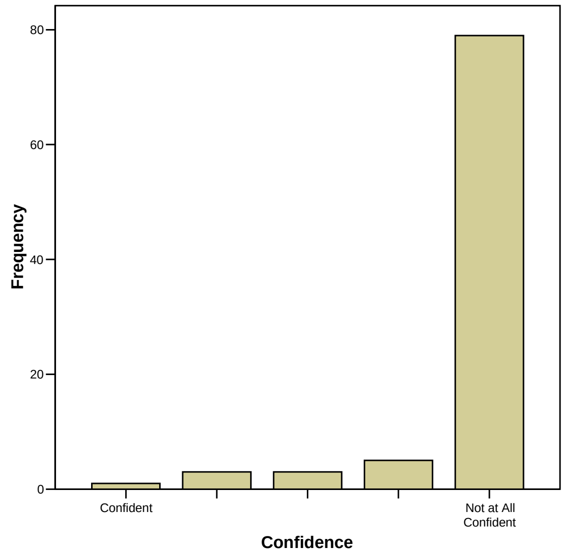


Figure 5l. Frequency of respondents and their level of confidence that they will be able to continue the flight for Scenario 12 for the next 80 miles without an alteration in track or altitude.

Table 6. Summary of the mean levels of confidence for each of the 12 scenarios in Study Three, distributed across the ratings (low, moderate, severe) for turbulence, updrafts and downdrafts that might be expected if the aircraft was to maintain the current track and altitude for 80 nautical miles.

		Low	**Moderate**	**Severe**
Scenario 1	Turbulence	2.67	3.83	4.93
	Updrafts	3.02	4.24	4.75
	Downdrafts	2.98	4.07	4.86
Scenario 2	Turbulence	2.59	3.66	5.12
	Updrafts	2.84	4.40	5.17
	Downdrafts	2.79	4.33	5.17
Scenario 3	Turbulence	3.56	5.18	5.90
	Updrafts	4.00	5.43	5.92
	Downdrafts	4.25	5.37	5.96
Scenario 4	Turbulence	2.22	3.26	5.25
	Updrafts	2.62	3.46	5.50
	Downdrafts	2.62	2.28	5.50
Scenario 5	Turbulence	5.14	5.13	5.86
	Updrafts	5.27	5.36	5.84
	Downdrafts	5.60	5.27	5.91
Scenario 6	Turbulence	1.89	2.93	3.00
	Updrafts	1.94	3.33	3.00
	Downdrafts	1.95	3.27	2.12
Scenario 7	Turbulence	1.91	3.13	6.00[6]
	Updrafts	1.98	3.44	6.00[7]
	Downdrafts	2.00	3.30	6.00[8]

[6] One respondent.
[7] One respondent.
[8] One respondent.

Table 6 (cont'd). Summary of the mean levels of confidence for each of the 12 scenarios in Study Three, distributed across the ratings (low, moderate, severe) for turbulence, updrafts and downdrafts that might be expected if the aircraft was to maintain the current track and altitude for 80 nautical miles.

		Low	**Moderate**	**Severe**
Scenario 8	Turbulence	2.76	4.11	4.94
	Updrafts	3.12	4.64	5.13
	Downdrafts	3.03	4.57	5.11
Scenario 9	Turbulence	2.93	4.83	5.54
	Updrafts	3.44	5.16	5.50
	Downdrafts	3.81	5.15	5.50
Scenario 10	Turbulence	4.22	4.71	5.47
	Updrafts	4.27	4.95	5.52
	Downdrafts	4.35	4.98	5.48
Scenario 11	Turbulence	5.00[9]	4.76	5.78
	Updrafts	4.50	5.09	5.91
	Downdrafts	4.57	5.03	5.89
Scenario 12	Turbulence	4.50	5.26	6.00
	Updrafts	4.23	5.57	6.00
	Downdrafts	5.00	5.46	6.00

The results of the correlational analyses suggest that participants' responses were not due to a predisposition towards higher or lower assessments on the part of individual pilots. Rather, the fact that there were, in some cases, very limited associations between the responses to different scenarios, suggests that they were interpreting the information on a case-by-case basis and that the relative levels of turbulence anticipated differed based on these interpretations. The question now arises as to the basis of these interpretations and whether these responses are predicted, at least in part, by the demographic features of the participants involved.

Although there were differences amongst participants in the levels of turbulence that they anticipated in response to the various displays, it is also important to note that there was also a degree of consistency amongst respondents for at least some of the displays. Table 8 summarises the

9 Two respondents.

frequency of respondents who rated the anticipated turbulence as ‘low’, ‘moderate’ or ‘severe’, across the twelve scenarios.

Table 8. Frequency of respondents who rated the anticipated turbulence as ‘low’, ‘moderate’ or ‘severe’, across the twelve scenarios.

		Low	**Moderate**	**Severe**
Scenario 1	Turbulence	45	47	14
Scenario 2	Turbulence	35	44	17
Scenario 3	Turbulence	9	57	30
Scenario 4	Turbulence	50	38	4
Scenario 5	Turbulence	7	45	37
Scenario 6	Turbulence	73	15	1
Scenario 7	Turbulence	69	15	1
Scenario 8	Turbulence	25	44	18
Scenario 9	Turbulence	14	64	13
Scenario 10	Turbulence	9	41	36
Scenario 11	Turbulence	2	25	58
Scenario 12	Turbulence	2	27	61

The frequency distribution in Table 8 suggests that some scenarios are more likely than others to elicit variability amongst respondents. For example, there is a greater level of consistency in the responses to Scenarios 6 and 7 than there is to Scenarios 2 and 8. As a consequence, these scenarios were selected as the basis for more detailed analysis. At one level, it was important to determine those features that may have contributed to the level of consistency amongst respondents. At another level, it was important to establish those features that may have contributed to the differences in the responses.

Table 7. Summary of the Spearman correlation (Rho) analyses between the 12 scenarios for the level of turbulence that might be expected if the aircraft was to maintain the current track and altitude**.**

		Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7	Scenario 8	Scenario 9	Scenario 10	Scenario 11
Scenario 2	Rho	.646**										
	Sig. (2-tailed)	.000										
	N	96										
Scenario 3	Rho	.241	.212									
	Sig. (2-tailed)	.018	.039									
	N	96	95									
Scenario 4	Rho	.362**	.504**	.053								
	Sig. (2-tailed)	.000	.000	.618								
	N	92	92	91								
Scenario 5	Rho	.426**	.212	.333**	.170							
	Sig. (2-tailed)	.000	.047	.002	.115							
	N	88	88	88	87							
Scenario 6	Rho	.341**	.272	.229	.192	.159						
	Sig. (2-tailed)	.001	.010	.032	.073	.143						
	N	88	88	88	88	86						
Scenario 7	Rho	.173	.240	.115	.286**	.141	.533**					
	Sig. (2-tailed)	.115	.028	.299	.008	.204	.000					
	N	84	84	84	84	83	85					
Scenario 8	Rho	.606**	.596**	.103	.471**	.315**	.311**	.222				
	Sig. (2-tailed)	.000	.000	.341	.000	.003	.004	.043				
	N	88	88	87	87	86	85	84				
Scenario 9	Rho	.170	.129	.365**	.274**	.302**	.214	.276	.114			
	Sig. (2-tailed)	.110	.227	.000	.009	.004	.045	.011	.292			
	N	90	90	90	89	88	88	85	88			
Scenario 10	Rho	.307**	.239	.234	.135	.270	.010	.180	.387**	.123		
	Sig. (2-tailed)	.004	.025	.029	.213	.012	.931	.105	.000	.252		
	N	88	88	87	87	85	85	82	86	88		
Scenario 11	Rho	-.004	.057	.370**	-.095	.289**	-.045	.064	.077	.026	.386(**)	
	Sig. (2-tailed)	.973	.599	.000	.386	.008	.686	.571	.486	.809	.000	
	N	86	86	85	85	83	83	80	84	86	85	
Scenario 12	Rho	.164	.092	.252	.009	.456**	-.049	.039	.107	.257	.360(**)	.406(**)
	Sig. (2-tailed)	.127	.392	.018	.937	.000	.653	.727	.328	.015	.001	.000
	N	88	88	88	88	86	86	83	86	89	87	86

** Correlation is significant at the 0.01 level (2-tailed).

The examination of the scenarios selected for further analysis began with a comparison between the frequency with which pilots of different ranks interpreted the likelihood of 'low', 'moderate', or 'severe' turbulence if there was no alteration in track or altitude for Scenarios 2, 8 6, and 7. The results indicated that, although there were was a broad level of consistency within ranks for Scenarios 6 and 7 (see Tables 9c and 9d), no such consistency was evident for Scenarios 2 and 8 (see Tables 9a and 9b). This evidence suggests that factors other than rank per se, are associated with interpretations of the likelihood of turbulence following the interrogation of weather radar displays.

To further examine the role of individual differences in predicting the outcomes associated with Scenarios 2 and 8, the demographic characteristics of respondents were summarised according to the levels of turbulence that were anticipated (see Tables 10a and 10b). The demographic characteristics included age, total hours flying experience, the number of flight hours accumulated as pilot-in-command, the number of hours accumulated under instrument flight rules, the number hours flown in excess of 20 nautical miles from an airport, the number of flight hours accumulated in the 90 days prior to completing the survey, the number of hours flown in excess of 20 nautical miles from an airport in the 90 days prior to completing the survey, the estimated number of in-flight decisions made to avoid weather, and the number of hours use of weather radar. An inspection of the means across the levels of expected turbulence failed to reveal any systematic difference between respondents, suggesting that the lack of consistency in the interpretation of the weather radar displays was not necessarily related to the demographic characteristics of pilots.

Table 9a. Frequency with which second officers, first officers, line captains, and check captains interpreted the likelihood of 'low', 'moderate', or 'severe' turbulence if there was no alteration in track or altitude for Scenario 2.

	Second Officer	First Officer	Line Captain	Check Captain
Low	1	10	16	5
Moderate	1	11	18	10
Severe	0	5	6	6

Table 9b. Frequency with which second officers, first officers, line captains, and check captains interpreted the likelihood of 'low', 'moderate', or 'severe' turbulence if there was no alteration in track or altitude for Scenario 8.

	Second Officer	First Officer	Line Captain	Check Captain
Low	1	7	14	2
Moderate	0	10	21	10
Severe	1	6	4	6

Table 9c. Frequency with which second officers, first officers, line captains, and check captains interpreted the likelihood of 'low', 'moderate', or 'severe' turbulence if there was no alteration in track or altitude for Scenario 6.

	Second Officer	First Officer	Line Captain	Check Captain
Low	1	21	34	15
Moderate	0	3	7	2
Severe	1	0	0	1

Table 9d. Frequency with which second officers, first officers, line captains, and check captains interpreted the likelihood of 'low', 'moderate', or 'severe' turbulence if there was no alteration in track or altitude for Scenario 7.

	Second Officer	First Officer	Line Captain	Check Captain
Low	1	19	31	16
Moderate	0	3	8	1
Severe	0	1	0	0

Although there was no evidence to suggest that the differences evident in the interpretations of the displays in Scenarios 2 and 8 were related to demographic features, it remains the case that the interpretation of one group of respondents differed from another, especially in terms of their expectations of turbulence. To identify the basis of these differences, a comparative assessment was conducted using comments that participants recorded that related to their confidence to continue the flight for the next 80 miles without an alteration in track or altitude. The comments related to the specific features of the displays and were distributed across the responses to the three levels of turbulence.

In analysing the comments, they were initially coded on the basis of whether a statement referred to the intensity of the storm cells, the distance between the aircraft and the storm cells, the drift of the storm cells, the shape of the storm cells, and/or the track of the aircraft relative to the storm cells (see Appendix D, Figures D.1 and D.2). These features were derived from the outcomes of Study Two and appeared to represent key indicators for the interpretation of weather radar displays.

The raw data were subsequently recalculated as a proportion of the total number of statements recorded for each of the three levels of turbulence (low, moderate, severe). A comparison between the types of statements that were recorded for Scenarios 2 and 8 revealed a pattern in which those participants who expected the turbulence to be 'severe', consistently made a proportion of statements relating to the intensity of the storm cells greater than those participants who expected the turbulence to be either 'moderate' or 'low' (see Figures 6a and 6b). Similarly, participants who expected the turbulence to be 'severe' also made a consistently greater proportion of statements relating to the distance between the aircraft and the storm cells and the drift of the storm cells. In the case of both scenarios, the relative proportion of statements that related to the track of the aircraft, relative to the storm cells, was least for those participants who expected the turbulence to be 'moderate'.

It is important to note that while there were differences in participants' expectations of the level of turbulence, the aim of this study was not to establish the relative veracity of one interpretation over another. The fact that there are differences in expectations suggests that, for particular types of weather radar displays, the interpretation of this information is difficult and is likely to be related to the relative importance that is ascribed to particular features of the information displayed.

Clearly, for most of the pilots, the track of the aircraft relative to the storm cells was of some significance to the process of interpretation, irrespective of their expectation of the level of turbulence.

Those differences that did occur (Intensity, distance, drift) tended to be related to features about which there was a greater level of uncertainty. Consequently, these represent significant areas where improvements in performance can be achieved, either through the redesign of weather displays, or through an improvement in training for operational personnel.

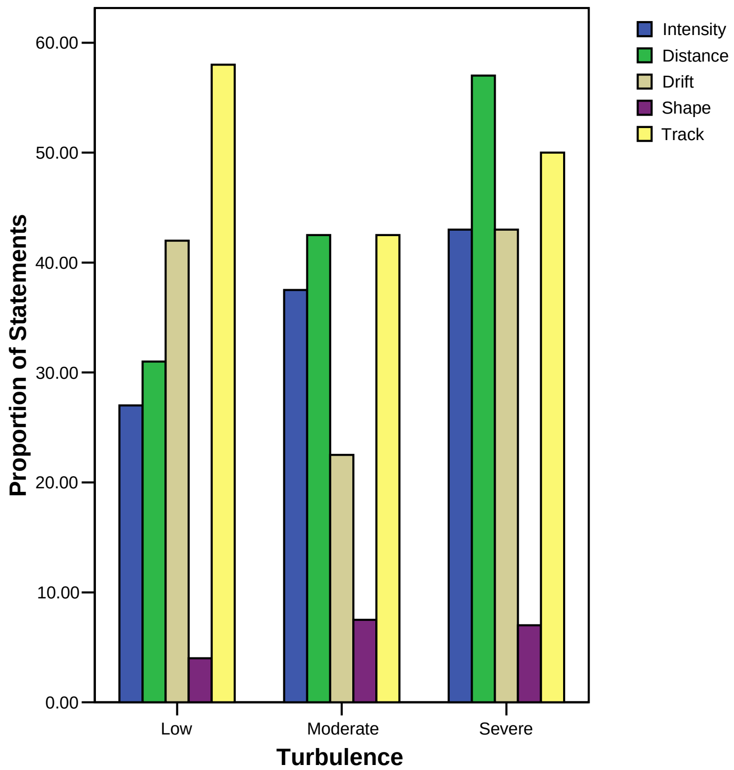


Figure 6a. Proportion of statements relating the intensity of the storm cells, the distance between the aircraft and the storm cells, the drift of the storm cells, the shape of the storm cells, and/or the track of the aircraft relative to the storm cells, distributed across participants' expectation of the turbulence associated with Scenario Two.

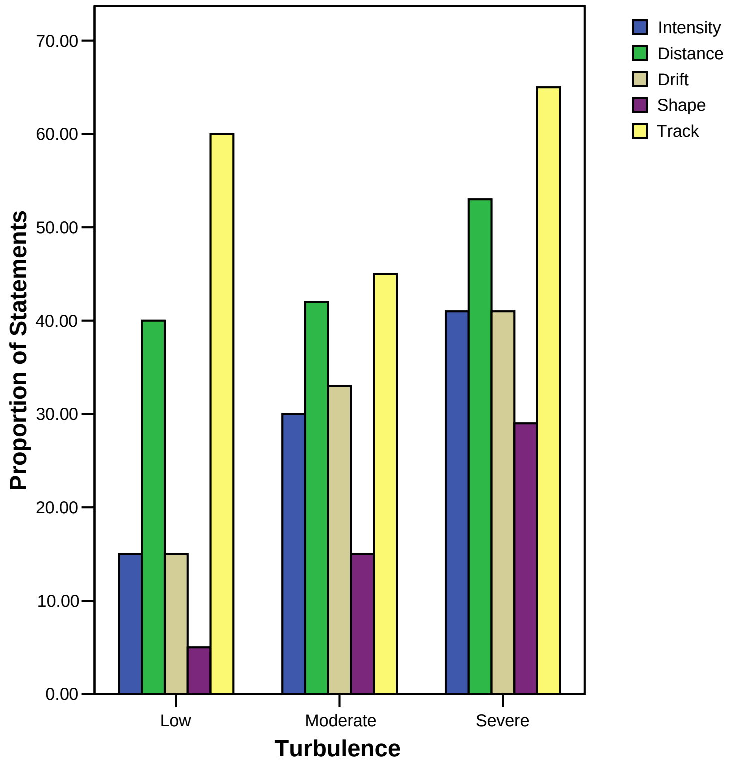


Figure 6b. Proportion of statements relating the intensity of the storm cells, the distance between the aircraft and the storm cells, the drift of the storm cells, the shape of the storm cells, and/or the track of the aircraft relative to the storm cells, distributed across participants' expectation of the turbulence associated with Scenario Eight.

5.3 Discussion

The broad aim of Study Three was to explore and extend some of the research outcomes of Studies One and Two. Consequently, there were a number of different elements to the survey, each of which related to a specific feature associated with the use of weather radar displays amongst pilots. At the outset, it is important to note that the study was descriptive in nature, and used a number of analytical strategies to consider the data and draw meaningful conclusions. Therefore, some of the conclusions are not definitive. Rather, they provide the basis for future research endeavours while, at the same time, providing some valuable information concerning the investigation of aircraft accidents and the development of strategies to improve pilot performance in relation to the use of weather radar displays during flight.

Table 10a. Demographic characteristics of respondents summarised according to the level of turbulence expected for Scenario 2

Expected Turbulence in Scenario 2		Age	Total Hours	Pilot in Command Hours	Instrument Hours	Cross-Country Hours	Hours in the last 90 days	Cross-Country Hours in the last 90 days	Number of In-Flight Decisions Made	Use of weather radar
Low	Mean	43.03	10885.24	9127.71	3151.48	7249.16	156.97	138.90	589.79	6361.88
	N	34	34	34	33	31	34	31	29	33
	Std. Deviation	12.51	6131.28	12942.92	3756.78	5470.86	78.86	89.34	945.99	4761.06
Moderate	Mean	42.77	9372.37	6107.16	3875.93	6900.28	122.05	95.63	660.27	5611.07
	N	43	43	43	43	43	43	43	41	42
	Std. Deviation	11.25	6286.15	5423.25	4695.10	5529.59	76.46	75.52	1257.78	5410.61
Severe	Mean	43.24	10252.94	6028.18	3068.56	8606.25	145.59	144.29	182.07	6694.12
	N	17	17	17	16	16	17	17	15	17
	Std. Deviation	9.66	6282.03	5400.42	2485.52	6218.20	71.63	71.01	265.18	6488.40

Table 10b. Demographic characteristics of respondents summarised according to the level of turbulence expected for Scenario 8

Expected Turbulence in Scenario 8		Age	Total Hours	Pilot in Command Hours	Instrument Hours	Cross-Country Hours	Hours in the last 90 days	Cross-Country Hours in the last 90 days	Number of In-Flight Decisions Made	Use of weather radar
Low	Mean	44.56	12101.52	10039.68	5286.63	6814.27	158.00	125.36	823.50	7210.08
	N	25	25	25	24	22	25	22	18	24
	Std. Deviation	13.546	6363.593	14712.404	5335.310	6160.610	94.249	105.015	1129.645	5162.809
Moderate	Mean	43.23	9084.53	6052.35	2765.74	7247.26	128.30	116.49	499.12	5000.35
	N	43	43	43	43	43	43	43	43	43
	Std. Deviation	10.177	5310.492	4975.316	2989.947	4483.081	65.748	73.067	1136.592	4335.434
Severe	Mean	40.18	10202.18	7447.47	2074.69	9480.00	146.82	146.82	356.67	7993.75
	N	17	17	17	16	16	17	17	15	16
	Std. Deviation	10.944	7547.276	6588.443	2523.866	7529.776	56.489	56.489	498.464	7583.444

One of the most significant outcomes of Study Three was some degree of confirmation for the retrospective analysis of weather radar-related aircraft accidents and incidents that emerged as an outcome of Study One. Participants in Study Three were asked to recall an incident in which they committed a weather radar-related error and were asked to classify the incident based on the taxonomy that was developed as part of Study One. The results of Study Three provided some degree of confirmation that, when examining the factors that precipitate weather radar-related errors, those factors that are most frequently cited tend to fall within the perception category of the expertise model of the use of weather radar displays. While there was a difference between the two studies in the frequency of incidents within the sub-categories (Study One showed a greater frequency for Perception – Recognition, where Study Three showed a greater frequency for Perception – Interpretation), it appears that when errors occur in relation to the use of weather radar-related displays, the precipitating factors appear to be related to the perception of information. Although this is a descriptive outcome, the pattern of responses, in combination with the information derived from Study Two, tends to lend support to this conclusion.

The second major feature associated with Study Three was an assessment of pilots' responses to a series of static weather radar displays that depicted various conditions during flight. While there was a level of consistency amongst participants for a number of the displays, it was also evident that some displays elicited a more variable response amongst participants in terms of their confidence that they would be able to continue the flight for 80 nautical miles without an alteration in track or altitude. This lack of consistency was also evident when pilots were asked to rate the expected level of turbulence if the flight was to continue along the proposed track and at the proposed altitude.

Given the lack of consistency in participants' responses to some of the displays, two scenarios (Scenarios 2 and 8), in particular, were examined in more detail. These scenarios were selected on the basis of the variability in pilots' expectations of the level of turbulence that would be experienced (low, moderate, severe) should the flight continue for 80 miles without an alteration in track or altitude. These differences were not explained by the demographic features collected, including flight experience, the use of weather radar displays, or age. Therefore, it was assumed that any differences evident were due to differences in the interpretation of the information presented in the displays.

In the case of each display, participants were asked to identify the specific features of the weather display that led to their level of confidence in continuing the flight without a change in track or altitude. These data were used subsequently to investigate differences between respondents on the basis of their expectations of the level of turbulence. The statements were coded using the features that were identified in Study Two as important for the accurate and timely interpretation of weather radar displays. These features included the intensity of the storm cells, the distance between the aircraft and the storm cells, the drift of the storm cells, the shape of the storm cells, and/or the track of the aircraft relative to the storm cells.

A comparison between the proportions of different features evident in the participants' statements revealed a pattern in which those participants who rated the level of turbulence expected as 'severe' for both Scenarios 2 and 8, also made a relatively greater proportion of statements relating to the intensity of the storm cell, the distance between the aircraft and the storm cells, and the drift of the storm cells. It is important to note that, although there were differences observed between participants in terms of the proportion of statements made, it is was not the aim of this study to determine the accuracy of a particular judgement per se. Rather, it was simply designed to establish the level of consistency amongst pilots, given a series of identical weather radar returns.

Given that the differences between the proportion of respondents' statements tended to relate to those features of the display that are subject to a greater degree of ambiguity, the results of Study Three suggest that greater emphasis in both system design and pilot training needs to be directed towards improving the interpretability of displays in relation to the intensity of storm cells, the distance between the aircraft and storm cells, and the drift of cells over a period of time. From the perspective of system design, increases in interpretability could occur through a reconsideration of the way in which weather-related information is displayed, and by incorporating predictive features of the extent of turbulence based on the pattern, shape, and movement of cells.

From a training perspective, it is important to educate pilots using both formal and informal strategies. In the case of formal strategies, pilots might be exposed to a series of real-time weather radar returns and be asked to anticipate the conditions based on their interpretation of the intensity of the cells, their distance from the aircraft, and the movement of the cells. Feedback should be provided subsequently, as a means of enabling pilots to develop their own strategies for the interpretation of weather radar displays. In the case of informal strategies, pilots need to be reminded that all flights are opportunities for learning and that active anticipation on the flightdeck, followed by a review of the outcomes, can provided as meaningful, if not more meaningful opportunities for learning than could be provided in the classroom environment.

5.3.1 Limitations

One of limitations concerning the use of surveys as a basis for understanding human performance is that they may not capture the complexity and dynamic nature of the behaviour that occurs within the operational context. In Study Three, the weather radar displays were static and, therefore, it might argued that pilots were unable to develop their mental representation of the weather conditions as might occur within the operational environment. This limitation may have impacted the responses, particularly in terms of the interpretation of the displays. Nevertheless, it should be noted that for the majority of the scenarios, the responses were relatively consistent, and it was only in some conditions that variable responses emerged. Indeed, it may have been the case that the use of static displays may have accentuated any differences that might otherwise exist within the operational environment.

6. General Discussion

The overall aim of this series of studies was to develop an understanding of the use of weather radar displays amongst pilots. Study One was designed to examine the outcomes of aircraft accidents as a basis for understanding the stage during information processing at which the majority of errors associated with the use of weather radar displays emerges. Aircraft accident and incident data from the Aviation Safety Reporting System and the National Transportation Safety Bureau were examined using an expertise model of the use of weather radar displays developed as part of Study One. Overall, the accident and incident reports reflected a dearth of information relating to the use, non-use, or misuse of the weather radar as a factor associated with the occurrence. This is despite the fact that all of the aircraft accidents and incidents examined were weather-related.

Those aircraft accidents and incidents in which weather radar was cited as significant factor were all coded using the expertise model of the use of weather radar. This model distinguished between eight stages of the use of weather radar displays from ‘perception’ to ‘response’. The results revealed a pattern in which the majority of aircraft accidents and incidents were coded as failures at the perception stage of the model. In particular, it was determined that the pilots either did not recognise the significance of the information derived from the display, or failed to interpret the information accurately.

Despite the apparent utility of the expertise model as a basis for classifying failures in the use of weather radar, it should be noted that the process of coding was based on, in many cases, limited information from aircraft accident and incident reports. In addition, it is not clear whether the outcomes were simply an artefact of the aircraft accident and incident investigation process or a random sample of events.

Study Two was designed, in part, to further examine the outcomes of Study One. However, it was also designed to examine some of the broader issues associated with the use of weather radar displays such as training and operational experience. Using the critical decision method as the basis for a semi-structured interview, five subject-matter experts were asked a series of questions relating to their use of weather radar systems, including the specific features of the display that they used to resolve a situation in which they were forced to rely on weather radar. A thematic analysis of the outcomes of the interview revealed that, consistent with Study One, the timely recognition of potential weather hazards represents one of the most important bases for interpreting and managing weather radar-related information successfully. The subsequent interpretation of this information appears to be dependent on rules-of-thumb that pilots have developed concerning the pattern of information, and the subsequent impact on the safe and efficient conduct of the flight. These rules-of-thumb are likely to have been developed on the basis of experience in which pilots were able to assess the accuracy of their interpretations against the conditions that were actually experienced.

In generating and selecting options to resolve the situations that the pilots described in Study Two, it was evident that some pilots appeared more oriented towards avoiding areas of potential turbulence, whereas others appeared more oriented towards negotiating a path to ensure the efficiency of the flight while maintaining the safety and comfort of passengers. However, it is not clear how this difference in goal-orientation might influence the interpretation and subsequent response to weather radar displays. For example, it might be the case that where there is a level of ambiguity associated with the interpretation of a weather radar display, those pilots who are motivated to avoid areas of potential turbulence may err on the side of caution, anticipating relatively higher levels of turbulence should the aircraft continue on the present track and altitude. By contrast, those participants who are motivated towards negotiating a path, may anticipate relatively lower levels of turbulence should the aircraft continue on the present track and altitude.

In relation to the general issues associated the use of weather radar displays, participants in Study Two emphasised the significance of both operational experience and experiential training as the basis for improvements in performance. In the case of operational experience, there was an emphasis on establishing the accuracy of the interpretations that were made during flight as a basis for establishing rules-of-thumb. Therefore, it was the quality, rather than the quantity of experience that appeared most important as a basis for increasing the rate of skill acquisition. This theme is consistent with other research that emphasises the significance of the quality of experience, in which participants are able to consider their responses and the consequences of their behaviour (see Leake, 1999; O'Hare, 1997; Wiggins, 1997).

An apparent lack of experiential training was a theme that was emphasised by a number of participants in Study Two as a significant problem for pilots in developing the skills necessary to use weather radar displays safely and efficiently. Experiential training is, in many ways, part of the process of accumulating experience, since the emphasis is on the acquisition of experience within a training context. In this case, it is important to avoid reference simply to factual or 'declarative' information in the absence of practical or 'procedural' information that might assist an operator acquire, interpret and respond to weather-related information. It is also important to avoid overly prescriptive approaches to the development of skills and, instead, enable learners to develop their own approaches to the problem within the constraints of the organisational and regulatory framework. This ensures that the skills that are developed are appropriate for the individual learner and are related to an existing skill-base. This type of learning philosophy has been successfully developed and implemented within the aviation context through the 'Weatherwise' series of computer-based training programs (Wiggins & O'Hare 2003a).

Although the outcomes of Study Two provided some useful information in terms of the use of weather radar amongst subject-matter experts, the information was based on a relatively limited sample. To further validate the outcomes of Studies One and Two, participants in Study Three were asked to complete a survey, the first section of which sought demographic information including flight experience and their experience using weather radar displays. The second section of the survey asked pilots to recall an incident, within the six months preceding the completion of the survey, during which they committed an error in relation to the use of the weather radar display. The final section sought pilots' responses to a series of 12 simulated weather radar displays.

The analysis of the incidents that pilots' recalled as part of the survey indicated that, consistent with Study One, where errors occurred in the use of weather radar displays, they were most likely to be associated with the perception stage of the expertise model. However, where Study One indicated that errors were most likely to occur at the recognition substage of perception, the outcomes of Study Three suggested that errors were most likely to be evident at the interpretation stage. Despite these differences, there appears a degree of consistency between the outcomes of the two studies in which perception, rather than information acquisition, option generation, or option evaluation appear most prevalent in aircraft accidents and incidents involving the use of weather radar displays.

In the case of pilots' responses to the 12 weather radar displays, there was a level of consistency amongst pilots for a number of the displays in terms of their confidence in being able to continue the proposed flights for 80 nautical miles without an alteration in track or altitude. Nevertheless, there were also a number of displays in which there was a considerable level of variability amongst respondents. Two of these scenarios (Scenarios Two and Eight) were selected for further examination.

In addition to the level of confidence in the continuation of the flight, participants were asked to rate the level of turbulence that might be expected should the aircraft continue along the proposed route and at the proposed altitude. Consistent with the data pertaining to confidence, participants' estimates of the turbulence were also variable for Scenarios 2 and 8 suggesting that these displays, in particular, were more difficult to interpret than the other displays in the series. A subsequent comparison between the demographic characteristics of pilots, distributed across their ratings of the level of turbulence expected, failed to highlight any differences that might explain the expectations for the two scenarios. This suggests that neither flight experience nor experience using weather radar is sufficient to explain the differences evident in pilots' expectations of the turbulence.

In addition to the ratings for each scenario, pilots were asked to identify the specific features of the display that led to their conclusions regarding their confidence in continuing the flight. These statements provided a basis to explain the differences in pilots' expectations of the turbulence associated with Scenarios Two and Eight. Each of the pilots' statements was coded according to whether they referred to the intensity of the storm cells, the distance between the aircraft and the storm cells, the drift of the storm cells, the shape of the storm cells, and/or the track of the aircraft relative to the storm cells. A descriptive comparison between the responses to the two scenarios revealed a consistent pattern in which those participants who expected the turbulence to be 'severe' also tended to make a relatively greater proportion of statements that referred to the intensity of the storms, the distance between the aircraft and the storm cells, and/or the drift of the storm cells. This suggests that, rather than experience per se, the differences between pilots in their ratings of the turbulence might be explained by their interpretation of key cues associated with the display. This interpretation might become more difficult as the information displayed becomes more ambiguous.

The emphasis on differences in the interpretation of displays that emerged in Study Three is consistent with the outcomes of the previous studies in which the perceptual stage of the expertise model was identified as the issue of most concern in the use of weather radar displays. The results also suggest that training initiatives ought to target both the recognition and interpretation of displays of hazardous weather as a priority to facilitate improvements in pilot performance.

6.1 Conclusion

It is important to note that this series of studies was largely descriptive in nature as it represents one of the first investigations of this type into the use of weather radar displays. Therefore, it is an initial step in the process of optimising the relationship between weather radar displays and human performance. Nevertheless, the results of this series of studies lend support to the proposition that, in some situations, pilots experience difficulty in both recognising and interpreting weather radar displays. This results in a level of inconsistency in performance, despite similar levels of operational experience. Future research needs to examine this issue from two different perspectives, the first of which involves an experimental approach in which pilots are asked to respond to a variety of dynamic displays of weather information. This research would not only serve to test the validity of the outcomes of the present study, but would also provide a level of confirmation for the direction of future training and design initiatives.

The second area of future research concerns the development and evaluation of a training strategy that emphasises experiential learning and situates the process within an operational context. Rather than based on subjective perceptions of the utility of the course, such evaluations should emphasise skill-based outcomes and evaluate these outcomes within the operational environment.

7. References

Abernethy, B 1990, Anticipation in squash: Differences in advance cue utilisation between expert and novice players, *Journal of Sports Sciences*, vol. 8, pp. 17-34.

Abernethy, B 1994, 'The nature of expertise in sport', *International perspectives on sport and exercise psychology*, eds. S Serpa, J Alves, and V Pataco. FIT Inc, Morgantown, pp. 57-68.

Adelson, B 1981, Problem-solving and the development of abstract categories in programming languages, *Memory and Cognition,* Vol. 9, pp. 422-433.

Akin, O 1980, *Models of architectural knowledge,* Pion, London.

Anderson, JR 1982, Acquisition of cognitive skill, *Psychological Review,* vol. 89, pp. 369-406.

Barr, JC 1993, *Airborne weather radar: A user's guide,* Iowa State University Press, Ames.

Bettman, JR, and Park, CW 1980, Effects of prior knowledge and experience and phase of choice process on consumer decision processes: A protocol analysis, *Journal of Consumer Research,* vol. 7, pp. 234-248.

Brown, JS, Burton, RR, and de Kleer, J 1981, 'Pedagogical, natural language, and knowledge engineering techniques in Sophie I, II, and III', *Intelligent tutoring systems.* Eds. D Sleeman and JS Brown, Academic Press, New York, pp. 227-281.

Charness, N 1979, Components of skill in bridge, *Canadian Journal of Psychology,* vol. 33, pp. 1-6.

Charness, N 1981, Search in Chess: Age and skill differences, Journal of Experimental Psychology: *Human Perception and Performance,* vol. 7, pp. 467-476.

Charness, N 1991, 'Expertise in chess: The balance between knowledge and chess', *Toward a general theory of expertise,* eds. KA Ericsson and J Smith, Cambridge University Press, Cambridge, pp. 39-63.

Chase, WG, and Ericsson, KA 1982, 'Skill and working memory', *The psychology of learning and motivation,* eds. GH Bower, Academic Press, New York, vol. 16, pp. 1-58.

Chase, WG, & Simon, HA, 1973, Perception in chess, *Cognitive Psychology,* vol. 4, pp. 55-81.

Chi, MTH, Feltovich, P, & Glaser, R 1981, Categorisation and representation of physics problems by experts and novices, *Cognitive Science,* vol. 5, pp. 121-152.

Christensen-Szalanski, J J J, Beck, DE, Christensen-Szalanski, CM, and Koepsell, TD, 1983, Effects of expertise and experience on risk judgements, *Journal of Applied Psychology,* vol. 68, pp. 278-284.

Cohen, MS 1993, ‘Taking risks and taking advice: The role of experience in airline pilot diversion decisions’, *Proceedings of the Seventh International Symposium on Aviation Psychology,* ed. RS Jensen, Ohio State University, Columbus, OH, pp. 244-248.

Cooke, N 1994, Varieties of knowledge elicitation techniques, *The International Journal of Human-Computer Studies,* vol. 41, pp. 801-849.

Cooke, NM, & McDonald, JE 1986, A formal methodology for acquiring and representing expert knowledge, *Proceedings of the IEEE,* vol. 74, pp. 1422-1430.

de Groot, AD 1966, ‘Perception and memory versus thought: Some old ideas and recent findings’, *Problem-solving: Research, method and theory,* ed. B Kleinmuntz, Wiley, New York, pp. 19-50.

Dekker, SWA 2001, The disembodiment of data in the analysis of human factors accidents, *Human Factors and Aerospace Safety, 1,* 39-57.

Dreyfus, HL and Dreyfus, SE 1986, ‘Why skills cannot be represented by rules’, *Advances in cognitive science 1,* ed. E. Sharley, Ellis Harwood, Chichester, pp. 315-335.

Edwards, BJ and Ryder, JM 1991, ‘Training analysis and design for complex aircrew tasks’, *Instruction: theoretical and applied perspectives,* ed. RF Dillon and JW Pellegrino, Praeger, New York, pp. 150-162.

Egan, DE and Schwartz, BJ 1979, Chunking in recall of symbolic drawings. *Memory and Cognition*, vol. 7, pp. 149-158.

Ericsson, KA and Polsen, PG 1988, An experimental analysis of a memory skill for dinner-orders. *Journal of Experimental Psychology: Learning, Memory and Cognition*, vol. 14, pp. 305-316.

Ericsson, KA and Staszewski, J 1989, ‘Skilled memory and expertise: Mechanisms of exceptional performance’, *Complex information processing: The impact of Herbert A. Simon,* eds. D Klahr and K Kotovsky, Erlbaum, Hillsdale, pp. 235-267.

Feltovich, PJ, Johnson, PE, Moller, JH, and Swanson, DB 1984, ‘LCS: The role and development of medical knowledge in diagnostic expertise’, *Readings in medical artificial intelligence,* ed. WJ Clancey & EH Shortliffe, Addison-Wesley, Reading, pp. 275-319.

Frey, PW and Adesman, P 1976, Recall memory for visually presented chess positions, *Memory and Cognition,* vol. 4, pp. 541-547.

Gilhooly, KJ, Wood, M, Kinnear, PR, and Green, C 1988, Skill in map reading and memory for maps, *Quarterly Journal of Experimental Psychology*, vol. 40A, pp. 87-107.

Glaser, R & Chi, MTH 1988, ‘Overview’, *The Nature of expertise,* eds. R Glaser, MTH Chi, and MJ Farr, Lawrence Erlbaum, Hillsdale, pp. xv-xxvii.

Goldberg, LR 1959, The effectiveness of clinicians’ judgements: The diagnosis of organic brain damage from the Bender-Gestalt test, *Journal of Consulting Psychology,* vol. 23, pp. 25-33.

Goldberg, LR 1968, Simple models or simple processes? Some research on clinical judgements, *American Psychologist,* vol. 23, pp. 483-496.

Gordon, DR 1986, Models of clinical expertise in American Nursing Practice, *Social Science and Medicine,* vol. 22, pp. 953-961.

Hammond, KR, Hamm, RM, Grassia, J, and Pearson, T 1987, Direct comparison of the efficacy of intuitive and analytical cognition in expert judgement, *IEEE transactions on Systems, Man, and Cybernetics,* vol. 17, pp. 753-770.

Hayes-Roth, F, Waterman, DA, and Lenat, DB 1983, *Building expert systems.* Addison-Wesley, Reading.

Hershey, DA, Walsh, DA, Read, SJ, and Chulef, AS 1990, The effects of expertise on financial problem solving: Evidence for goal directed, problem solving scripts, *Organisational Behaviour and Human Decision Processes,* vol. 46, pp. 77-101.

Hinsley, DA, Hayes, JR, and Simon, HA 1977, 'From words to equations: Meaning and representation in algebra word problems', *Cognitive processes in comprehension,* eds. MA Just and PA Carpenter, Erlbaum, Hillsdale, pp. 89-108.

Hoffman, RR 1987, The problem of extracting the knowledge of experts from the perspective of experimental psychology, *AI Magazine,* vol. 8, pp. 53-67.

Hoffman, RR, Crandall, B, and Shadbolt, N 1998, Use of the critical decision method to elicit expert knowledge: A case study in the methodology of cognitive task analysis, *Human Factors,* vol. 40, pp. 254-276.

Hoffman, PJ, Slovic, P, and Rorer, LG 1968, An analysis of variance model for the assessment of configural cue utilisation in clinical judgement, *Psychological Bulletin,* vol. 69, pp. 338-349.

Howe, MJA, and Smith, J 1988, Calender calculating in "idiot savants": How do they do it? *British Journal of Psychology,* vol. 79, pp. 371-386.

Johnson, PE, Duran, AA, Hassebrock, F, Moller, J, Prietula, M, Feltovich, PJ, and Swanson, DB 1981, Expertise and error in diagnostic reasoning, *Cognitive Science,* vol. 5, pp. 235-283.

Johnson, SD 1988, Cognitive analysis of expert and novice troubleshooting performance, *Performance Improvement Quarterly,* vol. 1, pp. 38-54.

Kieras, DE, and Bovair, S 1984, The role of a mental model in learning to operate a device, *Cognitive Science,* vol. 8, pp. 255-273.

Kirschenbaum, SS 1992, Influence of experience on information-search strategies, *Journal of Applied Psychology,* vol. 77, pp. 343-352.

Klein, GA 1989, Recognition-primed decisions (RPD), *Advances in Man-Machine Systems,* vol. 5, pp. 47-92.

Klein, G, & Klinger, D 1991, Naturalistic decision-making, *Crew System Ergonomics Information Analysis Centre Newsletter,* vol. 2, pp. 1-4.

Larkin, JH, McDermott, J, Simon, DP, and Simon, HA 1980, Expert and novice performance in solving physics problems, *Science,* vol. 208, pp. 1335-1342.

Leake, DB 1999, 'Case-based reasoning', *A companion to cognitive science* eds. W Betchell and G Graham, Blackwell, Malden, pp. 465.476.

Leli, DA, and Filskov, SB 1984, Clinical detection of intellectual deterioration associated with brain damage, *Journal of Clinical Psychology,* vol. 40, pp. 1435-1441.

Lesgold, A 1988, 'Problem solving', *The psychology of human thought,* eds. RS Sternberg and EE Smith, Cambridge University Press, New York, pp. 188-213.

Lesgold, AM 1984, 'Acquiring expertise', *Tutorials in learning and memory: Essays in honour of Gordon Bower,* eds. JR Anderson and SM Kosslyn, Freeman, San Francisco, pp. 31-60.

Lesgold, A, Rubinson, H, Feltovich, P, Glaser, R, Klopfer, D, and Wang, Y 1985, *Expertise in a complex skill: Diagnosing x-ray pictures.* University of Pittsburgh Press, Pittsburgh.

Livingston, C, and Borko, H 1990, High school mathematics review lessons: Expert-novice distinctions, *Journal for Research in Mathematics Education,* vol. 21, pp. 372-387.

Logan, GD 1988, Toward an instance theory of automatisation, *Psychological Review,* vol. 95, pp. 492-527.

McKinney, EH 1993, Flight leads and crisis decision-making, *Aviation, Space and Environmental Medicine,* vol. 64, pp. 359-362.

McKeithen, KB, Reitman, JS, Rueter, HH, and Hirtle, SC 1981, Knowledge organisation and skill differences in computer programmers, *Cognitive Psychology,* vol. 13, pp. 307-325.

Means, B, Salas, E, Crandall, B, and Jacobs, TO 1993, 'Training decision makers for the real world', *Decision-making in Action: Models and Methods,* eds. GA Klein, J Orasanu, R Calderwood, CE Zsambok, Ablex, Norwood, pp. 306-326..

Morris, PE, Tweedy, M, and Gruneberg, MM 1985, Interest, knowledge and the memorisation of soccer scores, *British Journal of Psychology,* vol. 76, pp. 415-425.

Mosier, KL 1991, 'Expert decision-making', *Proceedings of the Sixth International Symposium on Aviation Psychology,* ed. RS Jensen, Ohio State University, Columbus, pp. 266-271.

National Research Council 2002, *Weather radar technology: Beyond NEXRAD,* Author, Washington.

National Transportation Safety Board 2003, *Annual Review of Aircraft Accident Data: U.S. General Aviation, Calendar Year 1999,* Author, Washington, NTIS NTSB/ARG-03-02.

O'Hare, D 1997, Cognitive ability determinants of elite pilot performance, *Human Factors,* Vol. 39, pp. 540-552.

Olsen, JR, and Biolsi, K 1991, ‘Techniques for representing expert knowledge’, *Toward a General Theory of Expertise: Prospects and Limits,* eds. KA Ericsson, and J Smith, Cambridge University Press, Cambridge, pp. 240-285.

Ortega, KA 1989, Problem-solving: Expert/novice differences, *Human Factors Society Bulletin,* vol. 32, pp. 1-5.

Patel, VL, and Groen, GJ 1991, ‘The general and specific nature of medical expertise: A critical look’, *Toward a general theory of expertise*, eds. KA Ericsson and J Smith, Cambridge University Press, Cambridge, pp. 93-125.

Premkumar, G 1989, A cognitive study of the decision-making process in a business context: Implications for design of experts systems. *International Journal of Man-machine Studies,* vol. 31, pp. 557-572.

Purkitt, HE, and Dyson, JW 1990, ‘Decision-making under varying situational constraints’, *Contemporary issues in decision-making,* eds. K Borcherding, OI Larichev, and DM Messick, Elsevier, Amsterdam, pp. 353-366.

Randel, JM, Pugh, HL, and Reed, SK 1996, ‘Difference in expert and novice situation awareness in naturalistic decision-making’, *The International Journal of Human-Computer Studies,* vol. 45, pp. 579-597.

Rasmussen, J 1983, Skills, rules, and knowledge: signals, signs, and symbols, and other distinctions in human performance models, *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-13, pp. 257-266.

Rasmussen, J 1993, ‘Deciding and doing: Decision-making in natural contexts’, *Decision-making in action: Models and methods,* eds. GA Klein, J Orasanu, R Calderwood, CE Zsambok, Ablex, Norwood, pp. 158-191.

Reitman, J 1976, Skilled perception in go: Deducing memory structures from inter-response times, *Cognitive Psychology,* vol. 8, pp. 336-356.

Salthouse, TA 1991, ‘Expertise as the circumvention of human information processing’, *Toward a general theory of expertise: Prospects and limits,* eds. KA Ericsson, and J Smith, Cambridge University Press, Cambridge, pp. 286-300.

Shanteau, J 1988, Psychological characteristics and strategies of expert decision-makers, *Acta Psychologia,* vol. 68, pp. 203-215.

Shiffrin, RM, and Dumais, ST 1981, ‘The development of automatisation’, *Cognitive skills and their acquisition,* Erlbaum, Hillsdale, pp. 111-140.

Sloboda, J 1976, Visual perception of music notation. Registering pitch symbols in memory, *Quarterly Journal of Experimental Psychology,* vol. 28, pp. 1-16.

Stokes, AF, Kemper, KL, and Marsh, R 1992, *Time-stressed decision-making: A study of expert and novice aviators,* Aviation Research Laboratory, Urbana-Champaign, Tech. Rep. No. ARL-93-1/INEL-93-1.

Tikhomirov, OK, and Vinogradov, YE 1970, Emotions in the heuristic function, *Soviet Psychology,* vol. 8, pp. 198-203.

Voss, JF and Post, TA 1988, ‘On the solving of ill-structured problems’, *The nature of expertise,* eds. R Glaser and MTH. Chi, Lawrence Erlbaum, Hillsdale, pp. 261-285.

Voss, JF, Vesonder, GT, and Spilitch, GJ 1980, Text generation and recall by high-knowledge and low-knowledge individuals, *Journal of Verbal Learning and Verbal Behaviour*, vol. 19, pp. 651-667.

Wickens, CD, and Flach, JM 1988, ‘Information processing’, *Human factors in aviation,* eds. EL Wiener and DC Nagel, Academic Press, San Diego, pp. 111-156.

Wickens, CD, Gordon, SE, and Liu, Y 1998, *An introduction to human factors engineering,* Longman, New York.

Wiggins, MW 1997, ‘Expertise and cognitive skills development for ab-initio pilots’, *Aviation training: Learners, instruction and organisation,* eds. EL Wiener and DC Nagel, Ashgate, Aldershot, pp. 54-66.

Wiggins, MW, and O’Hare, D 1995, Expertise in aeronautical weather-related decision-making: A cross-sectional analysis of general aviation pilots, *Journal of Experimental Psychology: Applied,* vol. 1, pp. 304-319.

Wiggins, MW, and O’Hare D 2003a, Weatherwise: Evaluation of a cue-based training approach for the recognition of deteriorating weather conditions during flight, *Human Factors,* vol. 45, pp. 337-345.

Wiggins, MW, and O’Hare, D (2003b). Expert and novice pilot perceptions of static in-flight images of weather, *International Journal of Aviation Psychology,* vol. 13, pp. 173-187.

Wiggins, MW, Stevens, C, Howard, A, Henley, I, and O’Hare, D 2002, Expert, intermediate and novice performance during simulated pre-flight decision-making. *Australian Journal of Psychology,* vol. 54, pp. 162-167.

Ye, N, & Salvendy, G 1994, Quantitative and qualitative differences between experts and novices in chunking computer software knowledge. *International Journal of Human-Computer Interaction,* vol. 6, pp. 105-118.

ATTACHMENTS

Attachment A: Summary of Expertise Literature

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed.

Features of Expertise	**Characteristic**	**Field**	**Authors**
Perception	Pattern Recognition	Programming	Adelson (1981)
		Avionics	Chi, Feltovich and Glaser (1981)
		Radiology	Lesgold, Rubinson, Feltovich, Glaser, Klopfer and Wang (1985)
		General	Dreyfus and Dreyfus (1986)
		Troubleshooting	Johnson (1988)
		Music	Sloboda (1976)
		In-Flight Weather	Wiggins & O’Hare (2003b)
	Spatial Pattern Recognition	Chess	Charness (1981)
		Fire Fighting	Klein (1989)
	Situational Assessment	In-Flight Decision-Making	Mosier (1991)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	Characteristic	Field	Authors
Perception	Expanded Perceptual Network	Typing	Salthouse (1991)
	Enhanced Situation Awareness	Electronic Warfare Systems	Randel, Pugh, and Reed (1996)
Information Acquisition	Limited Information Search	Sonar Operation	Kirschenbaum (1992)
		Troubleshooting	Johnson (1988)
		Weather-Related Decision-Making	Wiggins and O'Hare (1995)
		Flight Instructor Decision-Making	Wiggins, Stevens, Howard, Henley, and O'Hare (2002)
	Goal-Structured Search	General	Anderson (1982)
		General	Lesgold (1988)
	Variable Search Strategies	General	Lesgold (1988)
		Finance	Hershey, Walsh, Read, and Chulef (1990)
	Efficient Cue Acquisition	Finance	Hershey, Walsh, Read, and Chulef (1990)
		Sonar Operation	Kirschenbaum (1992)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Information Acquisition	Efficient Cue Acquisition	Weather-Related Decision-Making	Wiggins and O'Hare (1995)
		Weather-Related Decision-Making	Stokes, Kemper, and Marsh (1992)
		Weather-Related Decision-Making	Wiggins and O'Hare (2003a)
	Individualised Strategies	Physics	Howe and Smith (1988)
	Solution-Centre Encoding	Physics	Chi, Feltovich, and Glaser (1981)
		Electronics	Egan and Schwartz (1979)
		Chess	Frey and Adesman (1976)
		Medical Science	Patel and Groen (1991)
		Go	Reitman (1976)
Option Generation	Forward Reasoning	General	Patel and Groen (1991)
	Recognition-Primed	Fire Fighting	Klein (1989)
		Fire Fighting	Klein and Klinger (1991)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Option Generation	Mental Simulation	Fire Fighting	Klein (1989)
		Fire Fighting	Klein and Klinger (1991)
	Plan Generation	Programming	Adelson (1981)
		Physics	Chi, Feltovich, and Glaser (1981)
	Solution-Centred Integration	Physics	Chi, Feltovich, and Glaser (1981)
		Electronics	Egan and Schwartz (1979)
		Chess	Frey and Adesman (1976)
		Medical Science	Patel and Groen (1991)
		Go	Reitman (1979)
Option Evaluation	Serial Evaluation	Chess	Chase and Simon (1973)
		Symbolic Images	Egan and Schwartz (1979)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Option Evaluation	Serial Evaluation	Fire Fighting	Klein (1989)
	Unconscious Deliberation	Fire Fighting	Klein (1989)
		General	Means, Salas, Crandall, and Jacobs (1993)
	Perception of Inconsistencies	Algebra	Hinsley, Hayes, and Simon (1977)
		Medical Science	Feltovich, Johnson, Moller, and Swanson (1984)
		Diagnostic Reasoning	Johnson, Duran, Hassebrock, Moller, Prietula, Feltovivh, and Swanson (1981)
Behaviour	Skill-Based Behaviour	General	Rasmussen (1993)
	Intuitive Behaviour	General	Dreyfus and Dreyfus (1986)
		Engineers	Hammond, Hamm, Grassia, and Pearson (1987)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Behaviour	Intuitive Behaviour	Programming	Larkin, McDermott, Simon, and Simon (1980)
		Nursing	Gordon (1986)
	Automated Behaviour	General	Dreyfus and Dreyfus (1986)
		General	Shiffrin and Dumais (1981)
	Compiled Productions	Consumers	Bettman and Park (1980)
		Programming	Larkin, McDermott, Simon, and Simon (1980)
	Improvisation	Teaching	Livingston and Borko (1990)
Mental Representation	Detailed Representation	Architecture	Akin (1980)
		Troubleshooting	Brown, Burton, and de Kleer (1981)
		Troubleshooting	Johnson (1988)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Mental Representation	Detailed Representation	Chess	Charness (1981)
		Electronics	Egan and Schwartz (1979)
		Radiology	Lesgold, Rubinson, Feltovich, Glaser, Klopfer, and Wang (1985)
		Programming	Lesgold (1984)
		Computer Gaming	Kieras and Bovair (1984)
		Programming	Premkumar (1989)
	Abstract Representation	Political Science	Voss and Post (1988)
	Functional Understanding	Symbolic Images	Egan and Schwartz (1979)
		Physics	Larkin, McDermott, Simon, and Simon (1980)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Mental Representation	Functional Understanding	Programming	McKeithen, Reitman, Rueter, and Hirtle (1981)
		Weather-Related Decision-Making	Stokes, Kemper, and Marsh (1992)
Knowledge-Base	Situational Knowledge	Fire Fighting	Klein (1989)
		Fire Fighting	Klein and Klinger (1991)
	Task-Specific Knowledge	General	Ortega (1989)
	Extensive Knowledge Base	Medical Science	Patel and Groen (1991)
		Teaching	Livingston and Borko (1990)
Memory	Enhanced Recall	Bridge	Charness (1979)
		Chess	Chase and Ericsson (1982)
		Chess	de Groot (1966)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Memory	Enhanced Recall	Chess	Ericsson and Staszewski (1989)
		Restaurant Orders	Ericsson and Polsen (1988)
		Map Reading	Gilhooly, Wood, Kinnear, and Green (1988)
		Programming	McKeithen, Reitman, Rueter, and Hirtle (1981)
		Soccer	Morris, Tweedy, and Gruneberg (1985)
		Medical Science	Patel and Groen (1991)
		Go	Reitman (1976)
		Text Recall	Voss, Vesonder, and Spilich (1980)
	Memory Chunking	Programming	Ye and Salvendy (1984)
	Flexible Information Retrieval	Radiology	Lesgold, Rubinson, Feltovich, Glaser, Klopfer, and Wang (1985)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Memory	Flexible Information Retrieval	Medical Science	Patel and Groen (1991)
Information Management	Diagnostic Emphasis	Finance	Hershey, Walsh, Read, and Chulef (1990)
		Chess	Chase and Simon (1973)
		Sonar Operation	Kirschenbaum (1992)
		Weather-Related Decision-Making	Wiggins and O'Hare (1995)
	Metacognition	Troubleshooting	Johnson (1988)
	Rapid Task Comprehension	Medical Science	Patel and Groen (1991)
		Political Science	Voss and Post (1988)
	Management of Affect	Music	Tikhomirov and Vinogradov (1970)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Performance Specificity	Domain-Specific Performance	Weather-Related Decision-Making	Wiggins and O'Hare (1995)
		General	Logan (1988)
		Medical Science	Patel and Groen (1991)
		In-Flight Diversions	Cohen (1993)
		Problem-Solving	Lesgold (1988)
Performance Outcomes	Accurate Performance	Clinical Diagnosis	Goldberg (1959)
		Sonar Operation	Kirschenbaum (1992)
		Air Combat	McKinney (1993)
		Medical Science	Christensen-Szalanski, Beck, Christensen-Szalanski, & Koepsell, (1983)
		Troubleshooting	Johnson (1988)

Table A.1 Summary of the domain-independent characteristics associated with expertise, including the features of expertise, the characteristics and field in which the features were evident, and the authors to whom the research observations are attributed (cont'd).

Features of Expertise	**Characteristic**	**Field**	**Authors**
Performance Outcomes	Accurate Performance	Weather-Related Decision-Making	Stokes, Kemper, and Marsh (1992)
	Inaccurate Performance	Clinical Diagnosis	Goldberg (1968)
	Inaccurate Performance	Radiology	Hoffman, Slovic and Rorer (1968)
		Differential Diagnosis	Leli and Filskov (1984)

Attachment B: Cognitive Interview Protocol

To be Read to Participants

Thank you for agreeing to participate in this interview.

This study is funded by the Australian Transportation Safety Bureau and is being conducted by Dr Mark Wiggins and Dr Sandra Bollwerk from the MARCS Auditory Laboratories at the University of Western Sydney. It is designed to examine your perceptions of various aspects of weather radar displays.

Please be aware that, for the purposes of the research, your responses are being recorded. These responses will be transcribed and you have the right to withdraw from the research at any time. If you choose to withdraw from the research, any recordings made will be erased and the transcriptions will be destroyed.

Please relax and take your time to answer the questions. There are no right or wrong responses. Do you have any questions before we begin?

Section A Personal Information

The following questions are designed to capture some general information about you.

---

1. Age: ________

2. Gender: Male ☐
   Female ☐

3. Country of Residence: ____________________________

4. Please indicate your highest license: Private ☐
   Commercial ☐
   Airline Transport ☐

5. Please indicate which of the following ratings that you hold: Instructor ☐
   Instrument ☐

6. Please indicate your present rank Second Officer ☐
   First Officer ☐
   Line Captain ☐
   Check Captain ☐

   If you are retired, what year did you retire? ________

7. Please indicate the type of aircraft on which you currently/last operate/d:

   ____

8. Please indicate the type of operations in which you are most often engaged:

   General Aviation ☐
   Airline: Long Haul ☐
   Airline: Short Haul ☐
   Other (Please Specify): ____________ ☐

9. Please indicate the region within which you have operated most frequently during the preceding 6 months (indicate more than one region if necessary):

   Asia ☐
   Pacific ☐
   Europe ☐
   Middle East ☐
   North America ☐
   South America ☐

Section B Flight Experience

The following questions relate to your flying experience. Please estimate these figures as accurately as possible.

1. Number of hours (total) experience: ☐

2. Number of hours (total) as pilot in command: ☐

3. Number of hours (total) actual IFR experience: ☐

4. Number of cross-country (in excess of 20 nm from an airport) hours experience: ☐

5. Number of hours (total) during the previous 90 days: ☐

6. Number of cross-country (in excess of 20 nm from an airport) hours during the previous 90 days: ☐

7. Number of times that you have been forced to make an in-flight weather-related decision: ☐

8. Number of hours (total) experience using weather radar systems: ☐

9. Please indicate the type of weather radar display that you have been using most frequently over the previous 6 months (or before you retired).

   ____

10. Have you ever been in a situation where a weather radar system displayed incorrect information?
    - Yes ☐
    - No ☐

    If yes, how many incidents have you been involved in? ________

11. To what extent do you rely on weather radar systems to assist your navigation in and around systems?
    - At all times ☐
    - Most of the time ☐
    - Some of the time ☐
    - Rarely ☐
    - Never ☐

12. To what extent do you trust the information displayed on weather radar systems?
    - At all times ☐
    - Most of the time ☐
    - Some of the time ☐
    - Rarely ☐
    - Never ☐

Section C Interview Questions

Think of a situation in which you were forced to rely on weather radar to assist your management of a flight. As you answer the following questions, keep this situation in mind.

The following questions will be used as probes to gain more information if required

Perception

2.1 What were the features of the situation that indicated to you that you would need to rely on weather radar to assist you to manage the flight? (Situation Assessment)

2.2 What was it that you were looking for when you determined that you would need to rely on weather radar to assist you to manage the flight? (Expanded Perceptual Network)

Information Management

3.1 In your initial interpretation of the weather radar, what were the most important issues that you were seeking to resolve? (Diagnostic Emphasis)

3.2 Can you describe how you developed a mental picture of the weather radar display? (Task Comprehension)

Information Acquisition

4.1 Can you describe the key features of the display that you were looking for? (Limited Information Search)

4.2 Why was this information important for your interpretation? (Goal Structured Search)

4.3 How did you know that this information was important in your interpretation of the information?

4.4 Can you describe how you searched the display for the key features? (Search Strategies)

4.5 How did you know that your interpretation of the weather radar display was accurate?

Option Generation

5.1 In responding to the information on the weather radar display, can you explain what it was that you were trying to achieve in terms of the management of the flight? (Forward Reasoning)

5.2 How did you know what options were available to you under the circumstances (Plan Generation)?

5.3 What were you looking for in the weather radar display when you were developing the options? (Mental Representation)

Option Evaluation

6.1 How did know what was the best option under the circumstances (Serial Evaluation)?

6.2 How did you identify options that were not appropriate?

Analogous Situations

7.1 Did the situation that you encountered remind you of any previous situations that you had experienced (Task Specific Knowledge)?

Attachment C: Selected Pages of the Weather Radar Survey

The Use of Weather Radar Survey

This research is sponsored by the Australian Transportation Safety Bureau and is conducted by the MARCS Auditory Laboratories at the University of Western Sydney, Australia

This study is a survey of various aspects of your use of weather radar displays during flight. It should take approximately 20 minutes to complete.

Please note that you are under no obligation to complete this study and you may withdraw at any time. Any responses that you provide will be de-identified and individual results will be considered confidential.

When you have completed the survey, please refer to the instructions in the information sheet as to how to return the survey to us.

Please keep a copy of the information sheet for your records and feel free to visit our website for more information about our aviation safety research.

THANK YOU!

Section C - Weather Radar Errors

Think of a situation over the last six months in which you made an error involving the use of a weather radar display (if you cannot think of an example, proceed to Section D). Keep this example in mind as you answer the following questions:

1. Please indicate the nature of the flight during which the error occurred.

   Domestic ○
   International ○

2. Were you the pilot flying?

   Yes ○
   No ○

3. At what stage of the flight did the error occur (please circle)?

   Pre-takeoff   Takeoff   Climb   Cruise   Descent   Approach   Landing

4. At what stage after the event did you realise that you had made an error?

   Immediately (within a minute) ○
   After a short period of time (one to five minutes) ○
   After a moderate period of time (six to fifteen minutes) ○
   After a lengthy period of time (sixteen minutes or longer) ○

5. How did you know that you had made an error?

   ____

   ____

6. What were the consequences of the error (if any)?

   ____

Section C (con't) - Weather Radar Errors

7. Which of the following options best describes the error that you recalled on the previous page in relation to the weather radar (only one response):

   (a). I didn't recognize the significance of the information

   (b). I didn't interpret the information accurately

   (c). I didn't interpret the information quickly enough

   (d). I didn't identify a solution

   (e). I didn't develop a plan to manage the situation

   (f). I didn't review the information before I responded

   (g). I didn't establish whether my response was appropriate

   (h). Other (please specify): ______________________

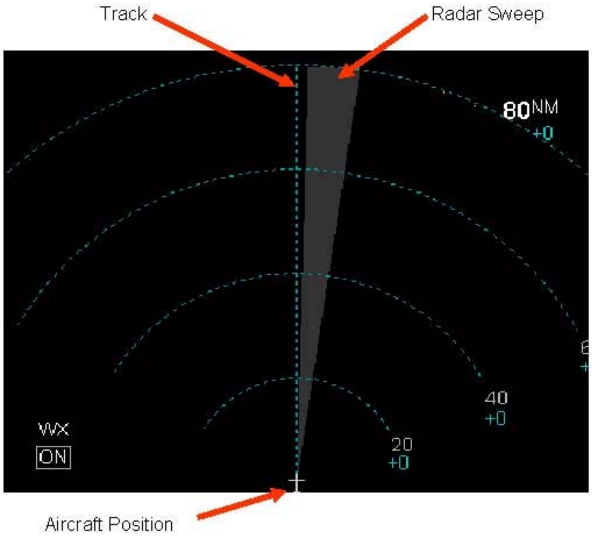
Section D - Interpretation of Weather Radar Displays

In the next (and final) section, you will be asked to respond to a number of in-flight weather radar displays. Please interpret the information in the context of a commercial airline operation.

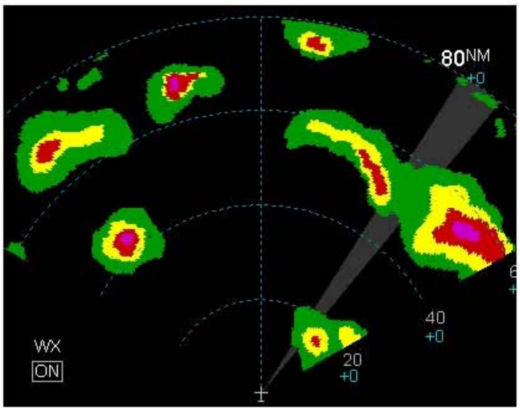
Assumptions:

> The aircraft is currently heading 340° at FL230. The wind is currently 250° at 22 knots and the tilt is set at 0 degrees.

You can use the following legends as a guide to assist your interpretation.



9



1. On the basis of this display, how confident are you that you will be able to continue the flight for the next 80 miles without an alteration in track or altitude (please circle)?

   Very Confident — Not at All Confident

   1 2 3 4 5 6

2. What specific features of the image led you to this conclusion?

   ____

   ____

   ____

3. If you were to continue the flight along the same track and altitude, at what level would you expect to experience the following events (please circle)?

   Turbulence

   Low Moderate Severe

   Updrafts

   Low Moderate Severe

   Downdrafts

   Low Moderate Severe

Attachment D: Participants' Statements for Scenarios 2 and 8

Table D.1 Specific features of the weather radar display in scenario two, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
application of wind effect	3		no apparent build ups along desired flight path	5		potential storm is closer with 90 x-avoid is at 221ct, our used GS of 4nm/min … us … in 12 min. lateral … path to x-track 10nm at 0.3nm/min x12min 1.5nm left of track which is unacceptable for such an intense …	1, 2, 5
cells closer but little movement across track	3, 5		lots of close together returns. Westerly wind would blow tops of cbs across track	3		Finger/high gradients/returns above moderate rain (ie>yellow) Cross wind track through the wx	1, 4, 5
being unable to squeeze past the cell on left at 50nm	2		wind may present an issue with drift into cell at 60nm	3		Large wx painting at 55nm left of track has slight protrusion & a steep gradient	1, 2
cell at 50-55nm should be cleared depending on movement on cells	2, 3		space available	2		"Hook"-type cloud at 60nm range	1, 2, 4
… & heavy rain … & clear path ahead, weak westerly wind	2, 3, 5		Cell at 55nm windward side	2		Need to …. Monitor movement of and build up of cells and divert around as required	1, 3
Right drift will infringe cell at 40nm or 60nm	3		Is similar to previous. Storm at 50nm L of track will not move as far, but due to strong nature would require divertion. Don't know what is beyond 80nm, so may need diversion if storms appear as we continue to track.	1, 3, 5		The last left side cell	2, 5
Flight path still clear & more severe cells downwind	1, 5		The cell at 2o'clock 40nm has a defined Hook formation, too close to track for comfort, probably lightning strike risk	1, 4, 5		Echoes seem isolated CBs, and with that wind I could fly 80 NM with a slight deviation.	3, 5

Table D.1 (cont'd) Specific features of the weather radar display in scenario two, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			Moderate			Severe	
No additional development noted on cells	1		The cell is some distance left off track, particularly heavy & severe rain. If the return is encountered we would experience light possibly moderate rain, tracking of the cell is still warranted	1, 2,		WX returns and wind direction will result in intercept of cells if no heading change by the point of passage severe WX	1, 3
Cells distance from track/heading	2, 5		Returns close to track on the right, with the aircraft drifting that way. The storms might not be drifting at the same rate as the aircraft	3, 5		Cell development at 65 miles 10 degrees left (downwind)	2
We are flying downwind of a cell at 50nm - this cell may drift across our Path	3		A/c appears to be 20nm closer to storms but relative positions of storms remain about the same. Lot of cells in small area > turbulence	1, 2		2 cells close to each other moving downwind, margin will be less than 20 NM on converging track.	2, 5
Again wx is left and right of track	5		Point ca. 20 to 30 degree (L) of HDG will probably be avoided since you will pass it in ca. 15min & it will only have moved ca. 4 to 5nm closer. However, since tilt is at 0 degree, no info avail. about vertical development; overhang & associated turb.	2, 4		Distance between cell number three along left side of track and projected track line based on wind direction and speed	2, 3, 5
May need a slight diversion right between 40 and 60nm	5		Small cells are very close together. A/c will not reach the more distant cells for 15 or 20 minutes	1, 2		Assuming a/c speed 300kt IAS, 420 TAS, 7nm/min, considering the echoes displayed, a/c would take about 11 minutes to reach the 80 nm arc and the clouds would move about 2 nm to the right.	3

Table D.1 (cont'd) Specific features of the weather radar display in scenario two, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
A/C has advanced 10 to 15nm & cells are still in same relative position to track & if this is Australia we know the flanking/feeder system are to the west of the severe storm ctrf, i.e cell to r of track at 40nm	1, 5		Cells do not appear to be closing or growing	1		due the wind coming from 250 and the heading of the aircraft is 340. this cause the clouds set in the path of the airplane	3
No weather on track	5		The large cell at approx 50nm.	2		cells close to track	2, 5
At 500kts/hour 80nm will take approx. 9.5 min. The movement of the storm cells by the wind should not have affected my flight path	3, 5		I cannot see the big picture, so I could maintain this track for maybe 60 miles, but need to change after due some weather ahead. Not enough data to make a conclusion without a deeper scan ahead.	5			
Almost 20nm between buildups - track keeps us clear by ca. 10nm; prevailing wind not strong enough to move CB's across track, no altitude change	3, 5		Knowledge of track (over ground or water and islands) would be useful to help interpret. It appears as though the radar is picking up ground returns at around 60NM which means the tilt may be too low and the scan may not be giving a proper interpretation of the current weather conditions. Also forecast weather and current conditions would also help.... assuming these are in fact clouds... the severity of the currently scanned cloud and the lack of information behind it may indicate some loss of information which may result in some diverting.	1, 5			

Table D.1 (cont'd) Specific features of the weather radar display in scenario two, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
Track looks ok, but again we have less than 20nm distance from storms	2, 5		Cell at 50 nm will drift towards flightpath. A diversion left might be necessary.	3, 5			
At tilt 0 degree I'm concerned what weather may be on track ahead at about 40nm. I would be using tilt down	5		Not clear	5			
Sufficient separation with cell at 50nm	2		Current cell placement and future development due to winds aloft.	1, 2			
Cells to the right are weak(ening).	1		the range of the radar	2			
Weather seems to be more than 20 miles off track with possible turbulence at 60 miles	1, 5		Still 10 miles clearance to closest severe returns. Need to vary tilt and is available sensitivity to check route. Would have to watch display as distance decreases.	2			
Still clear of the cell at 55nm but am still downwind of it.	2, 3		Active cells to close	1			
Hdg and range markers	2		Clear track ahead	5			
track clear of w/x	5		The a/c hdg is clear for the next 80 nm. At 8nm/min, this represents 10 mins. Confident for next 40nm, but would need to check beyond 80nm before committing further. Would also need to do a complete vertical scan.	5			

Table D.1 (cont'd) Specific features of the weather radar display in scenario two, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
The echo at 55NM will drift to the east, even though I would only fly through light rain, I would still be downwind from the cell, therefore I would still have to deviate to the right to maintain proper separation from turbulence and eventually hail (flying certainly under the anvil!). Up to FL250, all echoes must be avoided by at least 10NM.	1, 3, 5		also very many cells with the possibility of hail	1			
Shapes would lead me to believe that the movement is in the same direction as my flight path, and intensity levels would indicate to me that T'storms are, in general, past mature.	1, 3, 4		spacing of cells at 11 and 1 o'clock	2			
			Wx at 11 o'clock hasn't moved since last display & a clear gap exists.	2, 5			
			Track appears clear however reasonable moisture patterns are displayed beyond the weather returns.	1, 5			
			Close formation of 3 cells between 40 - 60 nm ahead. Would adjust tilt angle to approx -0.5 and re observe cells to right of track. May be possible to offset right of track by 10nm for next 50nm before turning left and tracking between last two cells.	1, 5			

Table D.1 (cont'd) Specific features of the weather radar display in scenario two, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
			possible "hook" on return to the right at 60 miles and overhang condition possible on the return at 70 miles on the left.	1, 2, 4			
			the wind/speed and the heading/distance to the cloud on the left	2, 3			
			in the next 60 miles I will be downwind of the cell	3			
			Winds from the left will blow indicated cells into my line of flight	3, 5			
			no returns on track	5			
			Heavy rain at 80nm	2			
			Cells are well defined and separated. Especially those west of track.	1			
			wind 250/22, track 340, it will blow the echo at 11 o`clock 65nm right into me track.	1, 2			
			I cannot rely on the image without first analyzing it by doing some steps. I will have to determine first if the image is land mass or cloud formation by tilting the antenna up and down by about 1 to -1 deg. Next is to make sure the path is clear of precipitation by shifting the antenna sensitivity from minimum to maximum.	5			
			track ahead is clear - steering wind not excessive and we have hdg data on our ND	3, 5			
			11 o'clock 65 miles	2			

Table D.2 Specific features of the weather radar display in scenario eight, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
maybe require diversion because of target at 11clock 70nm	2		pass close build ups	2, 5		cell at 68nm to left shows severe rain & is moving west to east	1, 3
Cell distance from track	2, 5		cell at 80 nm may be due to drift	3		looks like a line of storms at 60nm, may be able to get through gap straight ahead by continuing right of track	1, 5
Nil wx on track	5		application of wind effect	3		Cells moving onto track, anvil shapes to left & right	4, 5
may require a slight deviation to the right at 60 to 70nm	5		cells left of track at 70nm	2, 5		Finger/high gradients/ returns above moderate rain (ie>yellow) Cross wind track through the wx	1, 3, 4, 5
1) 20nm separation 2) isolated cells ahead 3) severe system has past to the east	1, 2, 3		cells at 65-70nm may cross our track if moving with prevailing winds	3, 5		Wx left of track at 70nm	2
The returns are close to track in the direction of drift	2, 5		Drift into cells at 60/80nm	3		The storm at 70nm to the left of track. Magenta returns at 70nm means this is a very intense cell and will start to be on track by the time the aircraft gets there. Because of the widespread nature of the cell, you would have to scan out beyond 80nm occasionally to see what else lies ahead and there would probably be more storms ahead. The intense storm with a high rain gradient would contain severe turbulence, so the storm would have to be given a wide birth.	1, 2, 3, 4, 5
No weather on track Weather at 65 to 70nm may be getting close to track	2, 5		Significant severe cells upwind & downwind of track	1, 5		"Hook"-type cloud at 60nm range	2, 4

Table D.2 (cont'd) Specific features of the weather radar display in scenario eight, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
Keep our eye on cell at 70 miles as it has a finger and moving in direction of flight Path	3, 4, 5		Cell at 60nm on left indicating strong development	1, 2		The weather at 75nm would have moved right on track	2, 3, 5
Plenty of space here maintain course maintain altitude	2		Hook formation/ double cell at 2o'clock 55nm will have cloud area not shown in return, a zigzag left at 40nm, then right at 60nm will be required	4, 5		Storms showing at roughly 65NM will be moving into the aircraft track. A closer look is required for a proper assessment. Use tilt.	2, 3, 5
Cell at 60nm to be negotiated	5		At ca. 70nm a well defined cell is too our left moving at 20kts, light, mod, heavy, severe rain is displayed its possible we may encounter all 4 & possible hail	1, 3, 4		Cells at 60nm will probably move into flightpath	2, 3, 5
Weather return at 60nm & right of track also the returns at 65nm and left of track	5		Red & pink returns and 'fingers' & "moon" shapes > very unstable air. Aircraft lateral clearance not great (5 10 10 nm)	1, 2, 4		Too many echoes,if we do not deviate	1, 5
At 230, I need to avoid by at least 10 miles. It is advisable to do so upwind, however, I could still do so downwind, experiencing some turbulence.	2		Turn may be advisable to provide sufficient buffer around paint at 70nm	2		Shape and level of activity along route	1, 4, 5
Active cell at 55 miles could move left	2, 3		Cells are not defined.	4		WX at 10- 11 oclock and 70 NM will be on proposed flight path at time of aircraft passage	2, 3, 5
weather approaching track at 80 nm and wind data	5		prevailing wind at the development left of track is concerning	3		close spacing of cells at 11 and 1 o'clock	1
track clear of w/x	5		Large cell at 70nm	1, 2		Cell at 55 miles and 75 miles may affect the track but cell on left of track at 70 miles may interfere with track	2, 5

Table D.2 (cont'd) Specific features of the weather radar display in scenario eight, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
no returns on track	5		Downwind of cell at 60nm off to left. Its movement is onto track. Insufficient separation	2, 3, 5		11 o'clock 75 miles	2
It will depend how fast echo 11.5 o'clock 65nm dissipates.	1		Potential of growth and movement of cells at 70nm and 75nm.	1, 2, 3			
wx on track	5		the aircraft track is out of cells	5			
Heavy rain in cell probably some turb downwind from the cell	1		track clear	5			
			The a/c hdg is clear for the next 80 nm. At 8nm/min, this represents 10 mins. Confident for next 40nm, but would need to check beyond 80nm before committing further. Would also need to do a complete vertical scan.	5			
			track clear of cells, cells very strong but isolated	1, 5			
			Wx close to track with strong returns	1, 2, 5			
			Close cell spacing with marked contours and possibility of cell drift onto track	3, 5			
			70 Nm range will encapsule, possible lightning activity between clouds	1			
			Best course although close formation of cells at 60,70,75nm may cause moderate turb. Cell at 70nm could be a problem with this wind	1, 3, 5			
			too close to cell on right at 60 miles and little room between cells beyond that	2			

Table D.2 (cont'd) Specific features of the weather radar display in scenario eight, distributed across participants' ratings of the turbulence expected as 'low', 'moderate', 'severe' and coded as to whether the statement referred to the intensity of the storm cells (1), the distance between the aircraft and the storm cells (2), the drift of the storm cells (3), the shape of the storm cells (4), and/or the track of the aircraft (5).

Low			**Moderate**			**Severe**	
			That strong echo on the left at 70nm is moving towards a/c track. It probably required a small deviation to the right to keep at least 20nm distance to the core.	2, 3			
			Cell placement in and between 60 and 80 nm	2			
			hdg and range markers plus cell at 70 nm.	2, 5			
			Many cells at the 60 nm mark. These would cause a diversion	5			
			cell 65 NM ahead and 20 NM left of track	2, 5			