PROTECTION OF AUTHOR’S COPYRIGHT

This copy has been supplied by the Library of the University of Otago on the understanding that the following conditions will be observed:

1. To comply with s56 of the Copyright Act 1994 [NZ], this thesis copy must only be used for the purposes of research or private study.

2. The author's permission must be obtained before any material in the thesis is reproduced, unless such reproduction falls within the fair dealing guidelines of the Copyright Act 1994. Due acknowledgement must be made to the author in any citation.

3. No further copies may be made without the permission of the Librarian of the University of Otago.
DECLARATION CONCERNING THESIS

Author's full name and year of birth:  
Kerrie Anne Paulley (for cataloguing purposes)

Title of thesis:  
Personal Risk Management in Pilots

Degree:  
PhD

Department:  
Psychology

I agree that this thesis may be consulted for research and study purposes and that reasonable quotation may be made from it, provided that proper acknowledgement of its use is made.

I consent to this thesis being copied in part or in whole for

i) a library

ii) an individual

at the discretion of the Librarian of the University of Otago.

Signature:  

Date:  
27/8/07

Note: This is the standard Library Declaration Form used by the University of Otago for all theses. The conditions set out on the form may be altered only in the most exceptional circumstances. Any restriction on access to a thesis may be permitted only with the approval of -

(i) the appropriate Pro-Vice-Chancellor in the case of a Master's thesis;
(ii) the Deputy Vice-Chancellor (Research), in consultation with the appropriate Pro-Vice-Chancellor, in the case of a PhD thesis

and after consultation with the Director of the University Consulting Group where appropriate.

The form is designed to protect the work of the candidate, by requiring proper acknowledgement of any quotations from it. At the same time the declaration preserves the University's philosophy that the purpose of research is to seek the truth and to extend the frontiers of knowledge and that the results of such research which have been written up in thesis form should be made available to others for scrutiny.

The normal protection of copyright law applies to theses.

September 1998
Personal Risk Management in Pilots

Keryn A. Pauley

A thesis submitted for the degree of Doctor of Philosophy at the University of Otago, Dunedin, New Zealand.

30th April, 2007
Abstract

Risk management is a key component of aeronautical decision-making and one of the possible causes of pilot error (e.g., Jensen, Guilke, & Hunter, 1997). Risk management encompasses risk perception and risk tolerance. Risk perception involves the detection of risks associated with a situation, whereas risk tolerance is the willingness to accept a given degree of risk (Hunter, 2002). Previous studies using flight simulators have found that risk perception and risk tolerance differs between pilots who fly into adverse weather and those who do not (e.g., O'Hare, Owen, Jorgensen, Wiegmann, Hunter, & Mullen, 2007). The aim of this research was to assess risk perception and risk tolerance using scenario-based measures. The measure of risk perception was developed over three studies. Since risk perception is a skill which expert pilots exercise (Jensen et al., 1997), I used the Cochran-Weiss-Shanteau (CWS, Weiss & Shanteau, 2003) index to measure how good pilots were at perceiving aeronautical risks. Weiss and Shanteau assumed that an expert should be able to discriminate between two relevant stimuli, and do so consistently. Participants were presented with flight scenarios and rated the risk involved in each scenario from 0 (low risk) to 100 (high risk). If a valid measure of expertise in risk perception, those with experience in aeronautical decision-making should have been better at this task. In study one the qualified pilots had higher and more variable CWS scores than the non-pilots, suggesting that some pilots were expert at this task, whereas most non-pilots were poor at this task. The focus of study two was shifted to weather-related decision-making (WRDM). Geography students, student pilots, and qualified pilots did not differ in their mean CWS scores, although the qualified pilots were most discriminating, and the geography students were most consistent. To decrease the reliance of the task on memory, study three included a blocking task in between each scenario. While only a small scale study, the results suggested that the blocking task improved the qualified pilots’ performance while the geography students’ performance deteriorated. In study four, I used Lopes’s (1987) theory to measure risk tolerance in pilots. According to Lopes (1987), risk tolerant individuals are motivated by opportunity, or what they can gain from taking risks, whereas risk averse individuals are motivated by threat, or what they can lose from taking risks. Qualified pilots were presented with 36 flight scenarios, varying in the level of threat and opportunity. The pilots rated the likelihood of going on the flights. Multiple regression equations were calculated, measuring the influence of threat and opportunity on each
pilot’s ratings. Pilots were largely risk averse, as their ratings were influenced by threat. The two pilots whose ratings were influenced by opportunity had experienced more aviation incidents compared to the pilots who were not influenced by opportunity. The aim of study five was to assess the relationship between risk management and in-flight WRDM. Qualified pilots completed a simulated flight into adverse weather, and four-computer based measures: the expertise in risk perception measure developed in study three, the risk tolerance measure developed in study four, and two implicit association tests assessing implicit risk perception and anxiousness towards adverse weather. Twelve pilots continued beyond the critical decision point, 18 pilots diverted, and 2 pilots crashed. There was no relationship between in-flight WRDM and expertise in weather-related risk perception. However, the pilots who diverted gave higher ratings of risk during the CWS task compared to the pilots who crashed. The pilots who diverted also tended to be more risk averse and implicitly perceived more risk in adverse weather, compared to the pilots who continued, suggesting a relationship between risk management and decision-making in a simulated flight into adverse weather. These five studies further highlight the role of risk management in pilot decision-making. The tools developed in these studies have potential for measuring risk management in pilots.
Acknowledgements

This thesis is a culmination of a lot hard work and support from a number of people. Thanks first to my supervisor, Associate Professor David O’Hare. His steady professional guidance and unwavering personal encouragement were invaluable to me. I am especially thankful for David’s help over the past 6-months. His helpful comments and incredibly quick turnover of drafts enabled me to complete my thesis within my timeframe.

I am also grateful for the opportunity to work alongside Mark Wiggins, from the University of Western Sydney. Not only did Mark provide me with a workspace and use of his simulator, he put aside time from his busy schedule to offer his expertise. I truly appreciate the help and encouragement Mark offered me during the 2-months that I was in Sydney.

Thank you to my statistical advisor, Brian Niven, as well as Iain Fulcher from the Otago aero club. Brian’s statistical knowledge and enthusiasm for fractional factorial designs was very well received. Thank you to Iain for ensuring that my flight scenarios were realistic. As a non-pilot, the technical knowledge that Iain (and other pilots) provided me with was essential.

I must also thank all of the members (past and present) of the Cognitive Ergonomics and Human Decision Making Laboratory, especially Nadia, Craig, Amy, Karen, Neil, and Cameron. Also, thank you to the members of the MARCS Auditory Laboratory at the University of Western Sydney who made me feel so welcome, especially Graham, Shaun, Paul, Jemma, Renee, Tim, and Arman.

Thanks to all those who participated in my research. A special thanks to those who helped organised participants, especially the University of Otago’s Geography department, Janet and Robert Rutherford from Glenorehy Air, Andrea McDowell and Stewart Graham from conair, Brian and Eric from Sydney Flight Training Centre, John Penno from Mainland Air, and Bryon Jones from Southern Wings.

On a personal level, I would like to thank my friends and family. Thanks especially to all of the ‘psychology girls’, who, while experiencing similar issues on the road to completing their Master’s and PhD’s, offered me support, guidance, and much hilarity (usually in the form of ‘singstar’). I am particularly indebted to my extended family, in particular, my parents, Robyn and Alan Pauley; my brother, Scott Pauley; my grandparents, Margaret Pauley, and Joyce and Ron Sharples; and my partner’s parents,
Nikki and Roger Miller. My family has offered me support and encouragement to fulfil my dreams.

Finally, I would like to thank my partner, Andrew Miller. I will always be grateful for the sacrifices that he made, while providing exemplary support and patience. I truly believe that this thesis would not have been completed without his encouragement and love.
Table of Contents

Abstract .................................................................................................................. ii
Acknowledgements ............................................................................................... iv
Table of Contents ................................................................................................. vi
List of Appendices ................................................................................................. x
List of Tables .......................................................................................................... xi
List of Figures ........................................................................................................ xiv
Chapter One: Aeronautical Decision-Making ......................................................... 1
  Weather-Related Decision-Making ........................................................................ 8
Chapter Two: Risk Management and Expertise ..................................................... 20
  Risk Management ................................................................................................ 20
  Expertise .............................................................................................................. 31
Chapter Three: Expertise in Aeronautical Risk Perception ..................................... 40
  Study One ............................................................................................................ 40
    Method ............................................................................................................... 41
    Participants ..................................................................................................... 41
      Materials ....................................................................................................... 42
      Procedure .................................................................................................... 47
    Results ............................................................................................................. 48
      Data Inspection .............................................................................................. 48
      CWS Analysis ............................................................................................... 49
      Relevance of Cues ......................................................................................... 50
      Cues Used by Pilots and Non-pilots ............................................................... 53
      Number of Cues ............................................................................................. 55
      Average Risk Rating ...................................................................................... 55
      Hazardous Event Scale ................................................................................ 56
      Usual Aeronautical Events ........................................................................... 56
      Personal Minimums ....................................................................................... 56
      Pilot Experience ............................................................................................. 57
      Non-pilot demographics ................................................................................. 57
    Discussion ......................................................................................................... 57
  Study Two .......................................................................................................... 61
Chapter Five: Implicit Associations and Time Pressure

Chapter Six: Risk Management and In-flight Decision-Making
Usual Aeronautical Practices ................................................. 193
Logistic Regression .......................................................... 193
Discussion ........................................................................... 195

Chapter 7: General Discussion ........................................... 210
Overview of Research ......................................................... 210
Limitations of Research ....................................................... 222
Theoretical Implications ...................................................... 224
Practical Implications for the Aviation Industry .................... 226
Future Research ................................................................. 227
Conclusion ........................................................................... 229

References .......................................................................... 230
List of Appendices

Appendix A: Flight scenarios (study one).........................................................250
Appendix B: Extra information sheet (study one).............................................259
Appendix C: Demographic Aeronautical Practices Questionnaire......................260
Appendix D: Demographic Information Questionnaire (non-pilot).................266
Appendix E: Flight scenarios (study two)........................................................267
Appendix F: Flight scenarios (study four, phase one)......................................275
Appendix G: Flight scenarios (study four, phase two).....................................288
Appendix H: Flight information (study five)...................................................322
Appendix I: Colour photos of IMC and VMC pictures (study five)..................325
List of Tables

Table 1.1. CAA legal requirements for VFR flight (CAA, 2006, p 55) .................................................. 9
Table 3.1. Factors, levels of each factor, and relevance of the factors used in the scenarios ................................................................. 43
Table 3.2. Factors, levels, and the simulator settings for the weather at Taieri ........................................... 64
Table 3.3. Factors and levels of en-route weather presented in the scenarios .................................................. 67
Table 3.4. Factors and levels for weather information at the destination airport ........................................... 68
Table 3.5. Raw CWS scores for the pilots and geography students .......................................................... 85
Table 3.6. Mean CWS scores for the qualified pilots and geography students for study two and study three .......................................................... 85
Table 4.1. Rankings for the amount of social approval (social) .............................................................. 95
Table 4.2. Rankings for the level of excitement (excitement) ................................................................. 95
Table 4.3. Rankings for the number of passengers (income) ................................................................. 96
Table 4.4. Rankings for the type of aircraft (career) .................................................................................. 97
Table 4.5. Rankings for the threat within the pilot (pilot) ........................................................................ 98
Table 4.6. Rankings for the threat within the aircraft (aircraft) ........................................................... 99
Table 4.7. Rankings for the threat of en-route weather (environment) .................................................. 100
Table 4.8. Rankings for the threat of time pressure (external) ............................................................ 101
Table 4.9. Facets of opportunity and threat presented in the scenarios .................................................. 102
Table 4.10. Within-subject regression coefficients for overall threat and opportunity ................................................................. 107
Table 4.11. Within-subject regression coefficients for pilot-related threat and opportunity for social approval ................................................................. 109
Table 4.12. Within-subject regression coefficients for aircraft-related threat and opportunity to gain excitement ........................................................................... 111
Table 4.13. Within-subject regression coefficients for weather-related threat and opportunity to earn income ......................................................................... 112
Table 4.14. Within-subject regression coefficients for threat of time pressure and the opportunity for career advancement ................................................................. 115
Table 6.1. Age and experience of participants in the normal and double-speed conditions ................................................................. 144
Table 6.2. Category labels and attribute examples ......................................................................... 147
Table 6.3. The order of the blocks for the Anxiety IAT .............................................. 150
Table 6.4. The order of the blocks for the Risky IAT .................................................. 150
Table 6.5. The compatible and incompatible conditions for the Anxiety and Risky IATs .................................................. 151
Table 6.6. Actions of the pilots in the normal and double-speed conditions after the first weather change .................................................. 153
Table 6.7. Actions of the pilots in the normal and double-speed conditions by the critical decision point .................................................. 154
Table 6.8. Age and experience of the pilots who continued or diverted after the first weather change .................................................. 155
Table 6.9. Flight certification of the pilots who continued or diverted after the first weather change .................................................. 155
Table 6.10. Age and experience of the pilots who continued or diverted by the critical decision point .................................................. 156
Table 6.11. Flight certification of the pilots who continued or diverted by the critical decision point .................................................. 156
Table 6.12. The proportion of pilots who continued and diverted at the first weather change, in each experience group, and when under normal and double-speed conditions .................................................. 157
Table 6.13. The proportion of pilots who continued and diverted by the critical decision point, in each experience group, and when under normal and double-speed conditions .................................................. 157
Table 6.14. Within-subject regression coefficients for overall threat and opportunity .................................................. 158
Table 6.15. Within-subject regression coefficients for pilot-related threat and opportunity for social approval .................................................. 158
Table 6.16. Within-subject regression coefficients for aircraft-related threat and opportunity to gain excitement .................................................. 159
Table 6.17. Within-subject regression coefficients for weather-related threat and opportunity to earn income .................................................. 159
Table 6.18. Within-subject regression coefficients for threat of time pressure and the opportunity for career advancement .................................................. 160
Table 6.19. Spearman’s rank order correlations between measures of experience and age, and implicit risk perception and anxiety..........................189
Table 6.20. Logistic regression for first weather change..........................195
Table 6.21. Logistic regression for the critical decision point.....................196
Table 7.1. Personal minimums for visibility and cloud ceiling across the five studies.................................................................219
### List of Figures

**Figure 1.1.** Decision Process Model (Orasanu & Fischer, 1997) ........................................... 4

**Figure 2.1.** Deery’s model of risk management ................................................................. 24

**Figure 3.1.** An example of a scenario .................................................................................. 43

**Figure 3.2.** Average raw CWS scores for pilots and non-pilots ........................................ 50

**Figure 3.3.** Raw CWS score for relevant and irrelevant cues .............................................. 52

**Figure 3.4.** The percentage of participants who used each cue during replication 1 ........... 54

**Figure 3.5.** The percentage of participants who used each cue during replication 2 ........... 54

**Figure 3.6.** An example of a scenario in study two ........................................................... 63

**Figure 3.7.** An example of an en-route weather forecast. From http://www.ifis.airways.co.nz/ ....................................................................................................................... 66

**Figure 3.8.** En-route weather forecast presented in the scenarios ....................................... 66

**Figure 3.9.** Example of a METAR. From http://www.ifis.airways.co.nz/ ......................... 69

**Figure 3.10.** An example of a METAR at the destination airport ........................................ 68

**Figure 3.11.** Raw mean discrimination scores across the three groups of participants ...... 72

**Figure 3.12.** Raw mean level of inconsistency across the three groups of participants ........ 73

**Figure 3.12.** Baddeley and Hitch’s (1974) working memory model (from Baddeley, 1999) ................................................................................................................................. 79

**Figure 4.1.** Involvement in hazardous aeronautical events for the risk tolerant and risk averse pilots ......................................................................................................................... 114

**Figure 5.1.** Time perception-based model of decision making under stress (Zakay, 1993, p. 68) ................................................................................................................................. 134

**Figure 6.1.** The flight simulator ............................................................................................ 145

**Figure 6.2.** Mean total hours experience for the pilots who flew at normal and double speed and continued and diverted by the critical decision point ........................................... 159

**Figure 6.3.** Mean number of hazardous events experienced by the pilots who continued, diverted and crashed at the first weather change ........................................................................ 161
Figure 6.4. Mean raw CWS score for pilots with a PPL and CPL........................................ 163
Figure 6.5. Mean CWS score for pilots with and without a flight instructor rating.............. 163
Figure 6.6. The average risk rating given by pilots who continued, diverted, and crashed, for replication one and two.............................................................. 165
Figure 6.7. Mean within-subject regression coefficient for threat for the pilots who continued and diverted after the first weather change.................................................. 169
Figure 6.8. Mean within-subject regression coefficients for threat for the pilots who continued, diverted, or crashed by the critical decision point........................................... 170
Figure 6.9. Mean within-subject regression coefficients for pilot-related threat for the pilots who continued, diverted, and crashed by the second weather change.............. 173
Figure 6.10. Mean within-subject regression coefficient for the threat of time pressure for the pilots who continue or divert by the second weather change........................... 180
Figure 6.11. The mean likelihood of going on the flights for the pilots who continued and diverted after the first weather change......................................................... 181
Figure 6.12. Distribution of D-ASIS Anxiety scores............................................................ 183
Figure 6.13. Mean D-ASIS anxiety scores for the pilots who continued, diverted, and crashed by the second weather change............................................................. 184
Figure 6.14. Distribution of D-ASIS Risky scores............................................................... 185
Figure 6.15. Mean D-ASIS risky scores for the pilots who continued, diverted, and crashed by the critical decision point................................................................. 186
Figure 6.16. Mean D-ASIS anxiety for the pilots who had and had not experienced a forced landing................................................................. 188
Figure 6.17. Mean D-ASIS Anxiety for the pilots with and without a flight instructor rating................................................................. 190
Figure 6.18. Mean altitude error for the pilots who continued, diverted, and crashed by the critical decision point................................................................. 192
Figure 6.19. Mean heading error for the pilots in the normal and double-speed conditions................................................................. 193
Figure 7.1. Mean CWS scores of the pilots in study two and study five.............................. 214
Figure 7.2. Deery’s (1999) model of risk management (black lines) extended to include implicit risk perception (blue lines)................................................................. 226
Chapter 1: Aeronautical Decision Making

The majority of New Zealand pilots fly in the General Aviation (GA) sector rather than the commercial airline sector. GA is any aviation operation other than airline and military (Ritchie, 1988). GA pilots fly a range of aircraft, such as fixed wing and rotary, single and twin-engine, and gliders. GA pilots also fly for a variety of reasons, including recreational flying, agricultural work, tourist flight operation, air rescue, air mail, and freight transportation (Airline Owners and Pilot Association (AOPA), 2003; O’Hare, 1999; Ritchie, 1988).

GA flying is often considered more difficult and dangerous than commercial aviation (Ritchie, 1988), with more accidents and fatalities per mile and per hour compared to airlines (International Civil Aviation Organization (ICAO), 1997). Although the overall rate of accidents has declined over the years, the rate of accidents due to pilot error has decreased at a much slower rate (Wiegmann & Shappell, 1997). The majority of GA accidents are associated with pilot performance rather than mechanical or structural failure. This figure ranges from 71% (O’Hare, Wiggins, Batt, & Morrison, 1994) to 85% (Li, Baker, Grabowski, & Rebok, 2001). These accidents are more likely to be fatal than other GA accidents (AOPA, 2006; Li et al., 2001, O’Hare et al., 1994).

According to Murray (1997), the very nature of aviation contributes to the high incidence of pilot error. A complex and dynamic environment, inadequate information, and the goal-orientated nature of pilots can lead to stress and errors in the cockpit. As such, the phases of flight involving complex and important tasks, such as take-off and landing, have the highest proportion of pilot-related accidents (AOPA, 2003).

A large proportion of pilot-related accidents are due to pilots making decisions that end adversely. Decision errors are departures from the normal decision processes which increase the chance of an adverse outcome (Lipshitz, 1997). O’Hare et al. (1994) analysed accidents and incidents that occurred in New Zealand between 1982 and 1991. These researchers found that 71% of the mishaps were attributable to pilot error, of these 35% were due to decision errors. Furthermore, decision errors were more likely to result in a fatality than information or action errors. Of the minor incidents, 30.5% were related to decision errors, while 62.5% of the serious or fatal accidents were related to decision errors.
There are two ways the decision process can fail (Orasanu, Martin, & Davison, 2001). The pilot may err due to a misinterpretation of the situation, or the pilot may correctly interpret the situation, but choose the wrong action. Decision errors often take the form of plan continuation errors (Holbrook, Orasanu, & McCoy, 2003). This is when pilots continue with a plan when faced with evidence that they should revise the plan, for example, continuing flying when the weather is deteriorating.

Decision-making is an important aspect of flying, from the initial flight planning to the final approach (Adams & Ericsson, 2000), and differentiates between competent and expert pilots (Jensen, Guilkey, & Tigner, 1997). Aeronautical Decision-Making (ADM) can be characterised by 3 stages: information acquisition, situation awareness, and choosing a course of action (Federal Aviation Administration (FAA), 1991; Wickens, 1999).

Unlike stick-and-rudder skills, it is difficult to assess pilot decision-making (Hunter, 2003). The motor skills of flying are easily observed, whereas decision-making must be indirectly observed from behaviour (Jensen, 1982; Klein & Thordsen, 1991). As such, outcome is often used as a measure of decision-making. If the pilot makes a decision which ends adversely, he or she has poor decision-making skills. However, while poor decisions often result in bad outcomes, sometimes bad outcomes result from sound decisions. Similarly, poor decisions can result in good outcomes (Kaempf & Klein, 1994; Lipshitz, 1997; Orasanu, et al., 2001). Therefore, rather than simply measuring outcomes, researchers should consider how pilots and other decision-makers make decisions.

There are many different models which describe ADM (e.g., Jensen, Adrian, & Lawton, 1987; Jensen et al., 1997; Klein, 1993; O'Hare, 1992; Orasanu & Fischer, 1997; Stokes, Kemper, & Kite, 1997). The basic components of decision-making are detection; situation assessment, which involves diagnosing the situation, risk assessment, and the time available; and choosing the course of action (Kaempf & Klein, 1994; Orasanu, 1993, Orasanu et al., 2001; Orasanu & Fischer, 1997). Decisions can go wrong at any of these stages: incorrect assessment of the problem, the risk, or the time available; or selecting the wrong course of action (Orasanu et al., 2001).

The decision-making process is context dependent. Different strategies will be used depending on the type of decision being made (Orasanu, 1993), context (e.g., perceived relevance of the information and framing of the question; Rohrbaugh & Shanteau, 1999), personal variables (e.g., pilot experience, training, & current health; Mosier, 1991;
Rohrbaugh & Shanteau, 1999), and the operational environment (e.g., whether civil or military, GA or commercial, & crew size; Mosier, 1991). Decision-making strategies differ according to the experience and expertise of the pilot. Experienced crews and pilots tailor their decision strategies to suit the situation, using heuristics or analytical strategies depending on the amount of time available (Orasanu & Fischer, 1997). Since decision-making is dependent on context and personal factors, training organisations should not prescribe one decision-making strategy to all situations and all pilots (Mosier, 1991; O’Hare, 2003; Orasanu, 1993).

According to Jensen, Adrion, and Lawton (1987), rational judgement and motivational judgement are two important aspects of pilot decision-making. For a pilot to make an appropriate decision, he or she must detect and diagnose changes in the environment, decide whether the change is relevant to the safety of the flight (rational judgement), and be sufficiently motivated to choose and execute a course of action (motivational judgement). There can be errors in rational judgement when the pilot does not detect a change, he or she notices the change but does not believe that it is relevant, or the predicted diagnosis may not be correct. Many factors can influence motivation, such as organisational and social pressures, personality factors (e.g., risk-taking), worrying, anxiety, and euphoria.

Orasanu and Fischer (1997) used simulated flights and incident reports to examine the decision-making of airline crew. Three types of decisions were analysed during the simulated flight: go/no-go decisions, choice problems (finding an alternate airport), and scheduling problems (prioritising the problems). Based on the data from the simulated flights and the incident reports, Orasanu and Fischer developed their Decision Process Model, as depicted in Figure 1.1.
Figure 1.1. Decision Process Model (Orasanu & Fischer, 1997).

This model describes pilot decision-making when faced with a situation requiring a course of action. The pilot will notice a change in the environment. The pilot will then use cues to assess the situation, which involves diagnosis of the problem, and assessment of the risk and time available. If under time constraints, the pilot will prescribe the first course of action which comes to mind. If time is available, the decision-making process depends on whether the problem is understood. If the problem is not understood, the pilot will gather more information, for example, through contact with air traffic control. If the problem is understood, the course of action depends on whether a rule is available, the number of options, and the number of tasks available. So, if there is time available, the pilot may consider each option before choosing a course of action. The risk involved in the situation also affects decision-making. High risk will cause the pilot to be cautious, and spend more time considering the options. Very high risk may cause stress, leading to attentional tunnelling and fixation on unsuitable solutions (Stokes, Kemper, & Kite, 1997).

According to the Decision Process Model, situation assessment is the most important part of ADM. Orasanu and Fischer (1997) examined accident reports where crew error contributed to the accident, and found that crews often had difficulty in interpreting the cues, and were particularly prone to underestimating the risk involved in a course of action. Orasanu et al (2001) found that airline accidents involving crew error were often
associated with situation ambiguity and dynamic risk, both of which can lead to errors in situation assessment. Situation assessment is more important than generating and evaluating the course of action (Kaempf & Orasanu, 1997). Through consultation of past experiences, checklists, and procedures, the pilots will often know what to do as soon as they have recognised their situation.

The importance of situation assessment in decision-making is highlighted in models by Stokes et al. (1997) and Klein (1993, 1997). Stokes, et al. (1997) believe that expert pilots make decisions quickly, through access to Long Term Memory. Experienced pilots will match available cues with any past experience in dealing with the situation. If the pilot has no experience in dealing with the situation, he or she will decide on a course of action based on a meta-cognitive appraisal.

Stokes et al.'s (1997) model is closely related to the Recognition-Primed Decision (RPD) model (Klein, 1993, 1997). Through interviews regarding critical decisions, Klein (1993) found that fire commanders did not generally use analytical decision-making strategies in an emergency. Rather, the fire commanders used prior experience to rapidly make decisions. The fire commanders' decision strategies were to recognise and diagnose the situation, and react based on past experience. If there was sufficient time, they ensured that the option was time and cost effective, and practical. Only if the option was not feasible would the commanders explore the next option. This model differs from other models, as it does not assume that the fire commanders explored all of the options.

Expert decision-making seems to be consistent with the RPD model (Klein, 1997). Klein reanalysed five studies of expert decision-making across different domains, and found that experts used a RPD strategy for between 42 and 80% of decisions. Furthermore, experts seem to be able to rapidly decide on a plausible course of action. Klein, Wolf, Militello, and Zsambok (1995) presented chess experts with chess scenarios. While only 16% of legal moves were judged as acceptable by chess grand masters, 66% of the first moves by the chess experts were deemed acceptable, suggesting that these experts were able to generate a plausible first option. Another assumption of the RPD is that since experts use a pattern matching strategy, time pressure should not affect performance. Calderwood, Klein, and Crandall (1988) found that expert chess players performed well both under time pressure and at normal playing speed. In contrast, intermediate chess players made more errors during rapid play compared to normal play.
Whereas the RPD model suggests that experts should excel at diagnosing the situation, this does not relate to self-perceptions of decision-making skills. While more experienced pilots believed that they were better than the average GA pilot at recognition of the problem, and generation and implementation of the solution, experience did not relate to self-perceptions of ability to diagnose the situation (Goh & Wiegmann, 2002b). While this seems contrary to Klein’s RPD model, it may be that experienced pilots are skilful at diagnosing the situation, but for whatever reason they are unaware of this ability, perhaps because the diagnosis occurs rapidly.

The RPD model does not explain expert decision-making when faced with a novel event. O’Hare’s (1992) Artful Decision Maker model, adapted from Janis and Mann (1977), describes expert decision-making under these circumstances. Similar to the RPD model, if the problem encountered is familiar, situational awareness will allow the pilot to detect and diagnose the problem. Based on the diagnosis, the pilot will select the appropriate response, which will be based on previous experience. If the problem is unexpected or unknown, the decision strategy will depend on the immediacy of the threat and subsequent time pressure. If there is immediate threat, the decision-maker will take the first course of action which comes to mind. If the threat is not immediate, the decision-maker can assess alternative courses of action. This involves determining and assessing the risk involved in each alternative. Based on this assessment, the appropriate course of action will be undertaken.

Whereas the Decision Process Model (Orasanu & Fischer, 1997), RPD, (Klein, 1993, 1997b), and the Artful Decision Maker Model (O’Hare, 1992), describe expert pilots’ decision-making process, Jensen et al.’s (1997) model describes the characteristics which an expert pilot possesses. Jensen et al. (1997) used semi-structured and structured interviews of experienced pilots to assess the characteristics of expert pilots, and identified four key components. Expert pilots had aviation experience, ability and motivation to attend to the task of flying, dynamic problem solving skills, and excellent risk management. Expert pilots had aviation experience which was more plentiful, varied, meaningful, relevant, and recent, compared to non-expert pilots. Risk management involved the development of personal minimums, which is the maximum level of risk which the pilot feels that he or she can handle, and awareness of potential hazards and how they can affect the safety of the flight. Expert pilots were also better at dynamically solving ill-defined problems. This involved understanding the problem and making sound
decisions, while considering the risk involved in the alternatives. Finally, expert pilots were better able to maintain attentional control, by focusing on the task at hand, and leaving personal problems outside the cockpit.

This model was assessed by Jensen et al. (1997) using a simulated flight. Pilots experienced a temporary decrease in power mid-flight, and were to diagnose the problem, and decide on a course of action. The pilots’ decision-making strategies were assessed using verbal protocol analysis. Experienced pilots were better able to cope with dynamic information, such as the changing flight environment. These pilots displayed more extensive aviation knowledge and situation awareness. The pilots spent less time diagnosing the problem, quickly gaining an understanding of the situation and assessing the viability of the options. The inexperienced pilots made decisions using analytical strategies, while experienced pilots made decisions intuitively, acting more spontaneously based on perceptual judgements.

The decision-making process is complex with many factors influencing the ability of pilots to make effective decisions. This is illustrated by the many accidents that are attributed to poor decision-making. For example, the National Transportation Safety Board (NTSB, 1996) investigated a tragic accident involving a pilot who took off in severe weather conditions. The three occupants of the plane: the pilot-in-command, pilot-in-training (a 7-year old girl), and the girl’s father, were fatally injured in the crash. The pilot-in-training was attempting the record for the cross-country flight conducted by the youngest pilot, leading to much media attention and an ambitious itinerary to complete the flight before the girl turned eight. Both the pilot-in-command and the training pilot felt fatigued, as neither had enough sleep in the days preceding the accident. Furthermore, the airport where the accident occurred was at an altitude that the pilot-in-command was not used to. The pilot-in-command departed in heavy rain and hail, thunderstorms a mile away from the airport, a wind sheer of 30 knots, and low visibility. The aircraft was travelling at a low air speed. Subsequently, the aircraft stalled and collided with the terrain. The NTSB concluded that poor decision-making was implicated in the crash. The tiredness of the pilot-in-command and the media pressure contributed to the decision to take off in adverse weather at an altitude that the pilot-in-command was not experienced in. Such accidents illustrate that there is never just one reason why pilots make poor decisions.
Weather Related Decision Making

Adverse weather is one of the many hazards associated with GA. According to the AOPA (1999, 2003, 2006), each year approximately 3% of GA accidents are due to flying into adverse weather. While this number is low, weather-related accidents are seven to nine times more likely to be fatal than crashes that do not occur in adverse weather (Bensyl, Moran, & Conway, 2001; Li & Baker, 1999). This high fatality rate may be due to a number of factors, including bad weather delaying rescue attempts (Li & Baker, 1999).

A pilot can only legally fly in adverse weather with a current instrument rating and while piloting a suitably equipped aircraft. Non-instrument rated pilots can only fly in relatively good weather, or Visual Meteorological Conditions (VMC) according to the Visual Flight Rules (VFR). Conversely, when the plane enters adverse weather, or Instrument Meteorological Conditions (IMC), the pilot must follow the Instrument Flight Rules (IFR; Transportation Safety Board of Canada (TSBC), 1990).

For a plane to be flown VFR, the pilot must be able to guide and control the plane visually (Nagel, 1988). VFR is defined by visibility and cloud ceiling (Fenwick, 1982). Visibility is the horizontal distance at which the pilot can readily perceive objects, and is measured in metres (Fenwick, 1982). For example, if the visibility was 10km, the pilot would be able to perceive objects, such as other planes, that are 10km away. Visibility is affected by cloud, precipitation, smoke, dust, and haze (Fenwick, 1982). Meanwhile, cloud ceiling is defined as the distance between the cloud base and the ground and is measured in feet (Fenwick, 1982).

The legal requirement for flying VFR depends on whether the flight is conducted within a controlled airspace, the height that the plane is travelling, and the type of operation (personal, instructional, agricultural, or helicopter) (Civil Aviation Authority of New Zealand (CAA), 2006; Fenwick, 1982). Table 1.1 presents the legal requirements of VFR flight according to the CAA (2006).
Table 1.1.

*CAA Legal Requirements for VFR Flight (CAA, 2006, p 55).*

<table>
<thead>
<tr>
<th>Class of airspace</th>
<th>Distance from the cloud</th>
<th>Flight visibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Clear of the cloud</td>
<td>8 km at or above 10000 ft AMSL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 km below 10000 ft AMSL</td>
</tr>
<tr>
<td>C, D, and E</td>
<td>2 km horizontal, 1000 ft</td>
<td>Vertically outside control zone, 500 ft vertically within control zone.</td>
</tr>
<tr>
<td>F and G</td>
<td>Above 3000 ft AMSL or 1000 ft above terrain, whichever is higher</td>
<td>2 km horizontally 1000 ft vertically</td>
</tr>
<tr>
<td>F and G</td>
<td>At or below 3000 ft above the terrain, whichever is higher</td>
<td>Clear of cloud and in sight of the surface 5 km</td>
</tr>
</tbody>
</table>

Weather-related accidents most commonly occur while flying VFR-into-IMC. This occurs when a pilot intends to fly VFR, and either takes off in IMC, or encounters IMC mid-flight. In 1998, 72% of GA weather-related crashes in the United States of America were caused by the pilot flying VFR-into-IMC (AOPA, 1999). While on a VFR flight, a pilot facing deteriorating weather must divert to another airport, or turn back to the departure airport, before he or she enters clouds (AOPA, 2003). While only approximately 2.5% of accidents occur when pilots fly VFR-into-IMC (Goh & Wiegmann, 2001a), between 11 and 23% of all fatal crashes involve VFR-into-IMC (Goh & Wiegmann, 2001a, TSBC, 1990). The fatality rate of all GA accidents ranges from 7.5 to 19%, whereas between 50 and 89% of VFR-into-IMC accidents are fatal (AOPA, 2003, Batt & O’Hare, 2005; Goh & Wiegmann, 2001a, 2002a, TSBC, 1990).

Clearly, flying VFR-into-IMC is very dangerous. However, an alarming number of pilots admit to doing this. According to O’Hare and Chalmers (1999), 27.4% of New Zealand pilots admitted to flying VFR-into-IMC at least once, and 4% had done so four or more times. Similar statistics have been obtained in the United States (e.g., Hunter, 1995). The aforementioned statistics suggest that these pilots are cheating death. It is not only inexperienced pilots who fly VFR-into-IMC. The TSBC (1990) found that there was no
difference in experience between the pilots who were involved in VFR-into-IMC accidents compared to those involved in other accidents.

VFR-into-IMC accidents are most often caused by pilots flying into objects, such as mountains or other aircraft, or terrain (AOPA, 2003). Low visibility and flight into clouds can lead to flying into objects or spatial disorientation. Spatial disorientation occurs when the pilot incorrectly estimates the altitude or the movement of the aircraft with respect to the terrain (Collins & Dollar, 1996; Nagel, 1988; Taneja, 2002). Illusions of motion can occur when the vestibular system indicates that the aircraft is moving, yet the visual system suggests otherwise (Nagel, 1988). This occurs when the pilot does not have any visual references and must rely on the vestibular system. Since the inner ear cannot detect a constant turn rate, a few seconds into a constant turn the inner ear will detect that the aircraft has stopped turning. When the pilot levels the wings, the vestibular system will indicate that the aircraft is turning in the opposite direction. The pilot is then likely to lose control of the aircraft (Kleimenhagen, Keones, & Szajkovics, 1993). Spatial disorientation is highly likely to occur when the pilot flies into clouds without an instrument rating. IFR rated pilots are trained to ignore their vestibular system and to rely solely on their instruments. The loss of control of the aircraft occurs on an average 178 seconds after a non-instrument rated pilot enters cloud (Kleimenhagen et al., 1993).

The Australian Transportation Safety Bureau provided an example of spatial disorientation following a pilot flying into cloud (ATSB, n.d.). The weather was VFR on take-off and deteriorated mid-flight. The pilot was in contact with Air Traffic Control (ATC), who told him or her to turn left to avoid another aircraft. During the completion of the turn, the pilot flew into cloud. Despite maintaining contact with ATC, the pilot succumbed to spatial disorientation, went into a left spiral dive, and impacted with the ground. The pilot was killed and the aircraft was destroyed.

Weather is both complex and dynamic, with complicated interactions between weather variables, and sometimes quickly changing circumstances (Wiggins & O’Hare, 1995; Wiggins, Stevens, Howard, Henley, & O’Hare, 2002). There is also a high degree of risk and uncertainty, especially in GA, where an inexperienced pilot can be flying an aircraft ill-equipped to managing bad weather (Wiggins et al., 2002). Therefore, decision-making involving weather is a very important skill for a pilot to exercise.

Skilled weather-related decision-making (WRDM) involves recognising and avoiding flying in hazardous weather (Wiggins & O’Hare, 1995). This is important both pre-flight
and in-flight. Pre-flight WRDM involves accessing weather forecasts and reports from the internet and phone (NTSB, 2005). The pilot should be prepared for any weather conditions he or she might encounter during the flight, and should be willing to cancel or delay the flight if required (NTSB, 2005). Meanwhile, in-flight decision making involves the integration of information from the outside view, instrument panel, weather forecasts, and any information from external sources, such as ATC (NSTB, 2005). If the pilot is unable to maintain VFR, he or she should divert to another airport, turn back, or make a precautionary landing.

Flying over mountainous terrain is especially hazardous in adverse weather (TSBC, 1990). Approximately half of Canadian VFR-into-IMC crashes, which occurred between 1976 and 1985, involved flying over mountainous terrain (TSBC, 1990). Mountain flying is important to New Zealand’s tourism industry. Some of the most spectacular areas of NZ are mountainous, such as the Southern Lakes District, Milford Sound, and the volcanic area of the North Island. A large amount of cross-country flying, especially tourist flight operations, is conducted in these areas. Due to the high proportion of mountain flying, and the dynamic nature of New Zealand weather, WRDM is an especially important consideration for New Zealand pilots.

An example of poor pre-flight decision-making, which led to the pilot’s death, is reported by the NTSB (2000). The pilot, who was IFR rated, filed a VFR flight plan from Albuqueque to Taos, New Mexico. The pilot contacted a friend at Taos, and a flight service station, who both informed him that VFR flight was not recommended due to poor visibility, low cloud ceiling, and heavy snow. Regardless of this information, the pilot decided to take-off. When 5 miles (8 km) from Taos airport, the pilot reported that snow was falling fast and that ice was accumulating on the wings. The aircraft impacted 1.5 miles (2.4 km) from the airport. At the time of the accident, there was a ¼ mile (402 m) visibility, 100 feet cloud ceiling, and heavy snow.

Cohn (1994) discusses a case of deliberate VFR-into-IMC, which resulted in two fatalities. The passenger, a friend of the pilot, discovered that her mother was unwell, and asked the pilot to fly her to see her mother in the morning. Pre-flight, the pilot was informed that the weather was IFR, and it was recommended that the pilot should not fly, even in an emergency. The pilot, although IFR-rated, lacked experience and did not feel confident flying IFR. The aircraft was maintained by the pilot, who was not a trained mechanic. Consequently, the aircraft was not airworthy. Regardless of the non-airworthy
aircraft and the deteriorating weather, the pilot decided to take-off. The pilot maintained contact with a manager of a flight service station, who gave him updated weather reports and observances. The weather deteriorated, and despite advice to turn back, the pilot pressed on into low visibility and cloud, turbulence, and rain. The dangerous situation was compounded with mechanical problems. The pilot could not make an emergency landing, lost control of the aircraft, and impacted with the ground. The investigation found that the vacuum pump was not installed correctly. The NTSB concluded that the pilot chose to take-off into poor weather conditions in a non-airworthy aircraft. He then continued flying, despite warnings of thunderstorms, perhaps while feeling some pressure to continue to the destination.

Pre-flight and in-flight weather-related information is acquired differently (Potter, Rockwell, & McCoy, 1989). Pre-flight decision-making involves obtaining and comprehending weather at departure and destination points, as well as en-route weather. In-flight decision-making involves updating this information through visual references and radio communication. Pilots gain WRDM skills through task-related experience, such as cross-country flying (Wiggins & O’Hare, 2003a). The decision-making process seems to differ for experienced and inexperienced pilots. Experienced pilots acquire information differently (Wiggins et al., 2002), seek more relevant information (Stokes et al., 1997; Wiggins & O’Hare, 1995; Wiggins et al., 2002), and place importance on different cues (Flathers, Giffin, & Rockwell, 1982, Wiggins & O’Hare, 2003a) when making weather-related decisions.

Measuring pilot’s worth functions is a useful way of studying how pilots use cues to make decisions. According to Flathers et al. (1982), worth functions represent pilots’ methods of combining information to make a decision. Flathers et al. examined pilots’ worth functions used when making a decision to divert to another airport. Pilots were given pen-and-paper scenarios describing a mechanical problem part way through a hypothetical flight. The pilots’ task was to decide whether to continue to the planned destination or divert to another airport. The pilots then ranked the available airports according to preference for diversion. The pilots made these ranking based on the quality of the ATC facilities, instrument approach facilities, time available, and the weather at the airport. The researchers found that the worth functions depended on the level of pilot certification. Private pilots were likely to use ATC and instrument approach facilities, as well as weather, to make their decision. Meanwhile, time was more important to other pilots. This
The study was important as it showed that pilots' worth functions could be assessed and that they depended on the experience of the pilots.

Worth functions have been useful for studying WRDM. Driskill and colleagues (Driskill, Weissmuller, Quebe, Hand, Dittmar, & Hunter, 1995) used worth functions to assess how pilots use weather-related information to make go/no-go decisions. These researchers theorised that pilots may make poor decisions, such as flying into adverse weather, because they do not appropriately assign importance to different weather-related factors.

Pilots can use a compensatory or a non-compensatory method when deciding between two different courses of action (e.g., go or no-go). Compensatory methods are a form of heuristic, whereby the decision-maker considers the costs and benefits of each option, and chooses the option associated with the most benefits and least costs. Meanwhile, a non-compensatory method of decision-making involves selecting an option that meets the criteria for every attribute (Vining & Fishwick, 1991).

Driskill et al. (1995) presented pilots with 81 pen-and-paper flight scenarios. Each scenario varied in terrain, ceiling, visibility, and precipitation. The participants assigned a comfort rating from 0 to 100 to each scenario, whereby 0 represented a flight that the pilot was least comfortable with and 100 most comfortable with. From this the authors identified the pilots' worth function. The authors found that the majority of pilots used a compensatory model. For example, the pilot would fly with a low cloud ceiling if visibility was sufficient, as the good visibility made up for the low cloud ceiling. Although efficient, compensatory decision-making may lead to poor decisions, where a pilot takes-off in low cloud-ceiling when the visibility is adequate.

A similar study was conducted by Hunter, Martinussen, and Wiggins (2003), using a larger sample of pilots from three different geographical areas: the United States, Norway, and Australia. Like Driskill et al. (1995), the pilots used a compensatory model, regardless of geographical location. Further support for pilots' use of compensatory models in decision-making is provided by Knecht, Harris, and Shappell (2004). Pilots were presented with scenarios differing in the level of visibility, cloud ceiling, and financial incentive to take-off, and made a go/no-go decision. No single variable predicted pilot decision-making. Rather, an interaction of factors, which differed across pilots, influenced the decision to take-off.
Since VFR-into-IMC crashes are largely caused by human, rather than mechanical error, research has been directed towards why pilots enter adverse weather. Possible reasons why pilots fly VFR-into-IMC can be divided into situational and pilot factors. Situational factors include decision framing, sunk costs, and social pressures. Factors inherent in the pilot may include situation assessment, hazard and risk perception, and risk tolerance. Hazard perception, risk perception, and risk tolerance will be discussed in the next chapter.

A pilot’s decision to continue into deteriorating weather can be influenced by how he or she frames the problem (Goh & Wiegmann, 2001b; Mosier-O’Neill, 1989). When faced with adverse weather mid-flight, the pilot can divert to another airport, return to the departure airport, make a precautionary landing, or continue. If the pilot views diverting or returning as leading to a certain loss (e.g., time wasted) he or she should be more likely to continue. Conversely, if the pilot views diverting or returning as leading to a possible gain (e.g., ensuring passengers’ safety), he or she should be more likely to divert or return to departure point (Goh & Wiegmann, 2001b; O’Hare & Smitheram, 1995). O’Hare and Smitheram (1995) tested this hypothesis using a simulated VFR flight. The pilots were faced with deteriorating weather three-quarters of the way through the flight and were asked to decide whether to continue. Pilots who were encouraged to view the problem in a positive frame (by emphasising the potential gains associated with diverting) were less likely to continue into adverse weather than pilots who were encouraged to view the problem in a negative frame (by emphasising the potential losses associated with diverting). The authors suggest that this has implications for pilot training. When making an in-flight decision, such as the decision to continue or divert, pilots should be trained and encouraged to forget about what they have already invested, instead concentrating on their current position. This should promote viewing the problem in a positive frame.

Attending to prior investments is a natural inclination and can lead to the sunk cost heuristic (Arkes & Blumer, 1985). This is when individuals attend to prior investments of time, money, or effort, which can lead to continuation of an original plan, even if they would be better off giving up (Arkes & Blumer, 1985). For example, very few people will leave a bad movie, because they believe that it is a waste of a movie ticket. However, since the movie ticket has already been purchased, the money already invested should not influence the viewer’s decision.
The sunk cost hypothesis predicts that pilots are more likely to continue into deteriorating weather when they encounter it late in the flight, as pilots feel that the further they have flown, the more time and money they will waste by diverting or returning (Wiegmann, Goh, & O’Hare, 2002). Similar to the decision-framing argument, this is due to the focus on prior investment. Rather than externally manipulating the emphasis on prior investment, the sunk-cost hypothesis assumes that pilots inherently think this way. Four studies have examined this hypothesis. O’Hare and Owen (2002) examined 77 GA cross-country crashes which occurred in New Zealand between 1988 and 2000. Weather-related crashes occurred further away from the destination airport than GA crashes due to mechanical failure, supporting the sunk-cost hypothesis. Mechanical failures, which are not usually due to pilot error, occurred randomly throughout the flight. Meanwhile, weather related accidents, which are generally due to pilot error, occurred later in the flight. However, two simulator studies showed little evidence for the sunk-cost hypothesis (O’Hare, Owen, Jorgensen, Wiegmann, Hunter, & Mullen, 2007; Wiegmann et al.; 2002). In both studies, pilots encountered simulated adverse weather during the flight, and the amount of time invested in the flight was manipulated. O’Hare et al. (2007) found prior investment did not influence pilot decision-making. Wiegmann et al.’s (2002) findings contradicted the sunk-cost hypothesis. Pilots who encountered adverse weather while only a short distance into the flight flew for longer than those who encountered the adverse weather later in the flight.

The sunk-cost hypothesis may be better studied by manipulating the proportion, rather than the amount, of time invested. Batt and O’Hare (2005) examined weather-related accidents and incidents within the ATSB database. The researchers compared pilots who continued VFR-into-IMC; pilots who made a precautionary landing; and pilots who took other weather-avoidance actions, such as turning back or requesting assistance to avoid the bad weather. The three groups did not differ in time en-route when the accident or incident occurred. However, there was a difference in the proportion of flight completed. The pilots who made steps to avoid the weather mostly experienced their incident or accident during the first half of the flight. Meanwhile, the pilots in the other two groups mostly experienced the incident or accident in the second half of the flight. The authors believe that half way through the flight the pilot shifts his or her focus from the departure to the destination airport. After this change of focus, the pilot is more likely to press on towards his or her new goal.
Social and organisational pressures may also influence pilots' decisions to continue flying in adverse weather. For example, pilots with passengers may feel increased pressure to reach their destination, and therefore be more likely to fly VFR-into-IMC (Goh & Wiegmann, 2001b). In O'Hare and Smitheram's (1995) study described above, the researchers also asked participants to rate which factors they considered when deciding to continue into adverse weather. Pilots rated impressing passengers as the least important factor. Although this suggests that social pressures are not important, Holbrook et al.'s (2003) study suggests otherwise. Holbrook et al. interviewed Alaskan pilots about a critical decision they had made. The pilots who wilfully continued into adverse weather often reported feeling social and organisational pressures at the time. Such pressures arose from passengers, the company, and other pilots. For example, some pilots reported that there was an implicit emphasis on completing the flight over safety considerations.

Examination of the NTSB accident database provides further support for the influence of social and organisational pressures. Goh and Wiegmann (2001b) analysed all VFR-into-IMC accidents that occurred in the United States between 1990 and 1997. These researchers measured the effect of social pressure by comparing the proportion of VFR-into-IMC accidents involving passengers to other GA crashes. The presence of passengers was more common in VFR-into-IMC than other crashes. The researchers concluded that pilots with passengers felt greater pressure to get to their destination, and therefore were more likely to fly VFR-into-IMC.

Organisational pressures contributed to a fatal Braniff International Airline crash, where all 86 passengers and crew members were killed (Nance, 1984; NTSB, 1968). The airline implemented a policy to encourage pilots to arrive at the airport on time, where passengers were paid money if the plane arrived 15 minutes after the scheduled time. This financial pressure may have contributed to the crash, where the pilots pushed through thunderstorms rather than navigating around them. At the time the flight was nearing 15 minutes late.

Factors inherent in the pilot may also increase the likelihood of a pilot continuing into adverse weather. Research has suggested that situational and hazard awareness, and risk perception and tolerance, are related to VFR-into-IMC flight.

Situation assessment is an important part of the decision-making process (e.g., Orasanu & Fischer, 1997). Pilots may fly VFR-into-IMC because of inaccurate situational assessment, that is, they do not realise that they are entering IFR conditions (Goh &
There are two possible reasons why pilots lack accurate situational assessment. Pilots without experience flying in adverse weather may find they are unable to accurately assess hazardous weather (Goh & Wiegmann; 2001a, 2001b; Layton & McCoy, 1989). Alternatively, erroneous situational assessment may occur because the transition between VFR and IFR conditions is gradual, making it difficult to distinguish between them (Goh & Wiegmann; 2001a, 2001b).

Difficulty in accurately assessing weather information may arise because GA pilots do not always understand weather information. Giffin and Rockwell (1987) measured pilots' weather-related knowledge and decision-making. A surprisingly large number of pilots surveyed had inadequate knowledge about weather and its effect on flight. However, this study was conducted two decades ago on pilots from the United States. Therefore, it is difficult to assess its relevance to New Zealand pilots who have recently been trained.

Inadequate and ambiguous weather-related information are barriers to good weather-related decision-making (Orasanu et al., 2001). Holbrook et al. (2003) interviewed Alaskan pilots regarding a weather-related critical decision. Seventy percent of the cases involved a plan continuation error, such as taking-off with known bad weather en-route. The researchers examined these situations, and found that one of the biggest precursors to plan continuation errors was the lack of weather-related information. Furthermore, since the pilots were mainly relying on forecasts, and the weather is extremely dynamic in Alaska, the weather information tended to be ambiguous.

There has been support for the situational assessment hypothesis with studies using flight simulators and accident databases. As well as testing the sunk-cost hypothesis, Wiegmann et al. (2002) investigated whether situational assessment could account for why pilots continue into adverse weather. They theorised that pilots who encountered unanticipated adverse weather early in their flight would continue because it contradicted their weather briefing obtained before take-off. The further into the flight, the more unreliable the weather briefing becomes. Therefore, pilots who encounter deteriorating weather later in the flight are more likely to disregard the weather briefing, and use their own perception to assess the weather situation. These pilots are less likely to continue into adverse weather. As described above, Wiegmann et al. (2002) found that less experienced pilots, and those who encountered adverse weather earlier in the flight, flew further into IFR conditions, thus supporting the situational awareness hypothesis. Goh and Wiegmann (2001a) also tested this and other hypotheses using a flight simulator. Pilots who continued
into simulated adverse weather were more inaccurate in their estimates of the level of visibility encountered during the flight.

The situational assessment hypothesis has also been supported by studies using accident databases. In Goh and Wiegmann’s (2001b) analysis of GA accidents, 24 percent of VFR-into-IMC accidents could be classified as ‘Inadvertent’. Thus, these accidents were attributable to the pilots entering adverse weather unintentionally. However, 76 percent of VFR-into-IMC crashes occurred even though the pilots knew they were entering IMC. Therefore, situational assessment, although important, is not the only factor.

Wiegmann et al. (2002) believe that the loss of situation awareness that these pilots experience may result from the inability to accurately and quickly assess the situation. According to the RPD model (Klein, 1993, 1997), through experience these pilots learn to quickly diagnose the situation. Therefore, experienced pilots should be less likely to fly into adverse weather than inexperienced pilots (Wiegmann et al., 2002).

However, research concerning experience and weather-related decision-making is mixed. O’Hare et al. (2007) found no relationship between experience and flying into adverse weather in a simulated flight, except for the number of hours flown cross-country in the last 90-days. Those who continued had more recent cross-country experience than those who diverted. Wiegmann et al. (2002) found a negative relationship between distance flown in adverse weather and total hours experience in VFR cross-country flight, where less experienced pilots flew further. NTSB (2005) conducted a case-control study, matching pilots who were involved in weather-related GA accidents from 2003 to 2004 with pilots who were flying in the same vicinity. The best predictor for accident involvement was prior involvement in accidents. There was no relationship between in-flight decision-making and flight experience. According to Wiggins and O’Hare (2003a), it is the quality of experience that is important in gaining WRDM skills, rather than the quantity. Goh and Wiegmann (2002b) believe that experience sometimes improves decision-making. Other times, experience may lead to overconfidence, and overestimation of their own ability to cope with the conditions, and underestimation of the risks involved. This may help explain why some experienced pilots fly in adverse conditions.

Weather-related decision-making is an essential part of pilot decision-making. The complex and dynamic nature of WRDM is highlighted by the many factors which contribute to poor WRDM, such as decision-framing, sunk-costs, social and organisational pressures, and situation assessment. Risk management also contributes to poor WRDM. In
the following chapter, I will discuss the role of hazard perception, risk perception, and risk tolerance in WRDM.
Chapter 2: Risk Management and Expertise

Risk Management

Risk management is an important skill for a pilot to exercise (AOPA, 2003) and is part of Jensen et al.'s (1997) model of expert aviators. Pilots need to be able to manage risks by recognizing hazards, understanding the risks involved, and making an appropriate decision based on this assessment (FAA, n.d.; Waring & Glendon, 1998).

While every activity contains an element of risk, some activities and situations are riskier than others (McKenna, 1988). Although there are many definitions of risk (Haight, 1986), most researchers define it as the probability of loss, given the amount of exposure to the situation or hazard (Brown & Groeger, 1986; Haight, 1988; Oppe, 1988; Sokolowska & Pohorille, 2000; Vlek & Stallen, 1988). There are two important components of risk, probability and consequences. That is, an assessment of risk involved in a situation or hazard requires consideration of the likelihood of an adverse event occurring and the severity of the outcome.

When assessing risk, individuals usually only consider the severity of the consequences, ignoring the likelihood of the event occurring (Sokolowska & Pohorille, 2000). It is easier for lay people to predict the consequences than assess the probability of an event occurring, as individuals tend not to understand probability. For example, more individuals are afraid of flying than driving, even though an automobile passenger is more likely to be involved in a fatal accident than an airline passenger. Lay people often focus on the consequences of being in an aircraft accident, and ignore the likelihood of it occurring.

Effective risk management is especially important in GA. GA is considered a high risk activity (Hunter, 2002), with many hazards and a higher fatality rate than commercial aviation (AOPA, 2003; NTSB, 2001a, 2001b). In 2001, there were 6.78 GA accidents per 100,000 flight hours in the United States (NTSB, 2001a), compared to 0.576 airline carrier accidents per 100,000 hours, excluding the September 11 terrorist attacks (NTSB, 2001b). Both the high likelihood of adverse events occurring and the severe consequences of crashing contribute to the high risk nature of GA.

There are many hazards associated with GA flight. Such hazards exist within the pilot, aircraft, environment, and external pressures (Jensen, Guilkey, & Hunter, 1998). It is very common for flight crew to deal with hazards. Klinect, Wilhelm, and Helmreich (1999)
found that 72% of flight segments contained at least one hazard. While most crews were able to manage these threats, and stop an error from occurring, 7% of threats lead to aircrew error. Veilette (n.d.) conducted a similar study and found that 93% of airline flights had at least one hazard, and 75% of flights contained at least one error. While 71% of these errors were inconsequential, 20% lead to an accident. These studies illustrate the high number of threats and errors in airline flight. While these numbers pertain to the airline industry, GA flight also involves a high number of threats and errors. The ATSB (2004) found that 87.1% of Australian GA pilots surveyed had experienced an incident, which may have affected the safety of the flight, which they were directly involved in during the last 12-months.

GA is also considered risky due to the severe consequences of crashing. GA aircraft travel between two and four times faster than automobiles, leading to an impact force of between 2 to 16 times that of crashing a car, and a higher fatality rate (AOPA, 1999). While automobile users are 10 times more likely to experience a crash per mile travelled than aircraft users, an aircraft accident is seven times more likely to be fatal than an automobile accident (AOPA, 1999).

Due to the high risk nature of GA, and the more lenient guidelines (compared to airline operation, AOPA, 2003), it is highly important that pilots practise sound risk management. Since it is largely the responsibility of the individual pilots, and the company that hires them, to effectively manage risks, it is very important to measure and effectively train risk management skills. Many accidents have occurred because the pilots have not effectively managed risks, for example, accepting more risk than they are capable of handling.

The CAA (2004) described a crash involving a microlight aircraft in North Canterbury, New Zealand. While the pilot was practicing a low-altitude forced landing with a simulated total engine failure, the pilot had difficulty recovering from the manoeuvre. The aircraft stalled, spun, and impacted with the ground. Both the pilot and his passenger were killed, and the aircraft was destroyed. The aircraft which the pilot normally flew performed differently from the microlight. Furthermore, the pilot was a reportedly thrill-seeker, frequently pushing the boundaries while flying. For example, to give him and his passenger a thrill, he once purposefully stalled an aircraft at 500 feet. At the time of the accident, the pilot was on anti-depressant medication and his toxicology report showed high levels of Citalopram®, a selective serotonin reuptake inhibitor, in his bloodstream.
This drug and the depression itself can interfere with cognitive and psychomotor skills. Relatives reported that the pilot was feeling exuberant following his recovery from depression. The CAA suggested that the side effects of the drug, coupled with the risk-taking nature of the pilot, and his elation upon recovering from depression, may have led him to attempt a dangerous manoeuvre.

Risk-taking involves the choice between at least two options, where one of the options has a chance of producing a negative consequence (Slovic, 1987). Examples of risk-taking include substance abuse, hazardous driving (Lane & Cherek, 2000), and flying into adverse weather. There are two main models that explain risk-taking behaviour: Risk Homeostasis Theory (RHT; Wilde, 1982), and Zero Risk Theory (Näätänen & Summala, 1974). A third model, Deery’s model of risk-taking will also be discussed. While these models describe risk-taking in drivers, they have been applied to risk-taking in other domains.

Risk Homeostasis Theory: RHT is both the most well known and controversial model of risk behaviour. According to Wilde (1982), individuals feel increased arousal and anxiety when faced with increased levels of risk. Individuals wish to maintain a constant and optimal level of arousal. A lower than optimal level of arousal decreases individuals’ ability to respond to danger, while a higher than optimal level of arousal depletes the nervous system resources (Simonet & Wilde, 1997). To maintain this target level of arousal, individuals constantly compare the actual and ideal level of risk. If there is any discrepancy, individuals will titrate the level of risk experienced. For example, if a driver has airbags installed in his or her car, the driver will take more risks to offset the increased safety of the airbags. According to RHT, any safety feature introduced to driving will not have any long-term effect on the accident and fatality rate.

There has been mixed support for this model. Some researchers have suggested that there is little change in the accident rate over time following the introduction of safety features in cars, such as seatbelts and airbags (Trimpop, 1996). Furthermore, reducing/increasing road safety does not always lead to a sustained increase/decrease in accidents and fatalities. In 1967, the Swedish Government changed from driving on the left to driving on the right. Immediately, there was a reduction in the number of deaths, which then stabilised (Wilde, Robertson, & Pless, 2002). Wilde believes that following the change, the drivers offset the increased risk by driving more carefully. However, other
authors have reported reductions in accident rates following safety improvements (Graham, 1982; McKenna, 1982; O’Neill & Williams, 1988; Wilde et al., 2002).

There has also been mixed support for RHT in more controlled studies. Stanton and Pinto (2000) had participants complete a simulated journey in fog conditions. Participants who drove with a visual enhancement system, which improved visibility, drove faster than the participants without visual aids. According to the RHT, the participants drove faster to offset the safety improvements provided by the visual enhancement system.

Perhaps research by Sagberg, Fosser, and Saetemo (1997) provides the strongest evidence for RHT. These researchers measured the risk-taking behaviour of taxi drivers with and without an Advanced Braking System (ABS). The taxi drivers with ABS displayed more risk-taking behaviour, such as closer following distances, than the drivers without ABS. However, a simulation-based study found that the availability of ABS did not affect drivers’ risk-taking (Glendon, Hoyes, Haigney, & Taylor, 1996).

Zero-Risk Theory: While the RHT posits that individuals wish to maintain a constant level of risk, the Zero-Risk theory (Näätänen & Summala, 1974; Summala, 1988) hypothesises that drivers’ wish to maintain zero risk. Drivers are motivated to take risks (e.g., speeding) because of the perceived benefits of doing so (e.g., shorter travelling times). Drivers experience a fear response when exposed to a hazard, but tend to escape and avoid the negative consequences of the hazard. Through experiencing near misses, drivers learn which cues are associated with accidents (e.g., if a car in front suddenly breaks) and learn to maintain a safety margin (e.g., increasing the following distance). Increased driving experience leads to conditioning of the fear and uncertainty felt while driving and increased confidence. Näätänen and Summala (1974) believe that by adapting to high risk, drivers begin to feel that there is no risk. An accident can occur when the driver perceives the risk as lower than it actually is.

Deery’s Risk Management Model: Deery’s (1999) model of why automobile drivers take unnecessary risks can be adapted for aeronautical risk management. According to this model, there are three reasons why an individual would take risks: poor risk perception, poor driving skills, and a high level of risk tolerance (sometimes termed ‘risk acceptance’). In this model, risk perception encompasses the assessment of the hazard, the subjective experience of risk, and the self-assessment of skill. For example, a pilot may be
approaching a lowered cloud base. Hazard perception would involve the recognition that the cloud ceiling was becoming dangerously low. Meanwhile, the subjective experience of risk would involve the recognition of the risk involved in flying through the cloud. The self assessment of skill involves the pilots’ recognition of his or her ability to fly in these situations. Finally, risk tolerance relates to the pilot’s willingness to take risks. A pilot would show risk tolerance if he or she understood that flying through cloud is risky, but was willing to tolerate this risk to achieve his or her goal. This model is depicted in Figure 2.1.

Figure 2.1. Deery’s model of risk management.

Hazard perception, risk perception, and risk tolerance are important aspects of risk management. There is evidence that these factors contribute to risk-taking in aviation and driving. For example, O’Hare and Wiegmann (2003) found that pilots who flew into adverse weather during a simulated flight differed in risk perception compared to those who diverted to another airport. The pilots who flew into adverse weather gave lower ratings of the risk of continuing into adverse weather than those pilots who diverted. However, the pilots who continued also rated the risk of continuing into adverse weather as higher than the risk of diverting. Yet, they still chose to fly into the adverse weather, suggesting that they were tolerant of the risks.
Hazard perception: A hazard in aviation is anything which poses a threat to the safety of the aircraft or its occupants (O'Hare, 1990). So, hazard perception relates to the ability to recognise a hazard, and the knowledge that this hazard poses a threat. For example, hazard perception in aviation might involve perceiving that the cloud base is lowering. Hazard perception is often measured as the time taken to respond to a hazard (Deery & Love, 1996). For example, hazard perception in driving has been measured by presenting participants with a video which is filmed from the driver’s perspective. Participants indicate the perceived level of risk with a lever, pressing it according to how much risk is perceived. Time taken to respond to the hazard is also measured (Currie, 1969; Deery & Love, 1996; Pelz & Krupat, 1974). Slower hazard perception has been found in people under the influence of alcohol (Deery & Love, 1996), and those with an accident record (Currie, 1969; Pelz & Krupat, 1974).

There is accumulating evidence that pilots who fly VFR-into-IMC possess poor hazard perception. That is, they are not aware of what constitutes potential hazards, or the likelihood of those hazards occurring. O’Hare (1990) assessed pilots’ hazard awareness by asking them to estimate how often accidents are attributable to different hazards. All pilots underestimated the percentage of accidents caused by pilot error. Furthermore, the pilots who chose to take-off into marginal VFR conditions during a simulated flight were less accurate in estimating when in the flight accidents were most likely to occur, than pilots who did not take-off. Goh and Wiegmann (2001a) used a similar procedure to assess pilots’ hazard awareness. Pilots were asked to estimate the likelihood of hazards occurring in a variety of settings, in situations involving themselves and others. Pilots, who diverted when faced with IMC during a simulated flight, were more accurate when estimating the likelihood of both pilot error and weather causing accidents. Together, these results suggest that pilots who fly VFR-into-IMC have poorer hazard awareness than other pilots.

Risk perception: Risk perception involves the detection of the risk associated with a situation or hazard, from within an individual and the environment (Hunter, 2002) and is an important part of decision-making. The level of risk assigned to a particular course of action affects decision-making and subsequent behaviour (Siegrist, Gutscher, & Earle, 2005).
Risk perception is usually measured by asking participants to rate the risk involved in a situation or object. Measured in this way, inexperienced (Brown & Groeger, 1988) and younger drivers (DeJoy, 1992; Groeger, & Chapman, 1996; Kanelaidis, Zervas, & Karagioules, 2000) and those with traffic violations (Pelz & Krupat, 1974) perceived less risk in driving situations than more experienced, older, and safer drivers.

Risk perception is also related to risk-taking in non-driving areas. Green, Turner, Purdie, and McClure (2003) measured risk-taking in sky-divers. Sky-divers were presented with 9 hypothetical sky-diving scenarios, which varied in the degree of risk, such as weather factors, and familiarity with the drop zone and gear. The participants rated the risk involved in each scenario and indicated whether they would jump. Both gender and risk perception predicted the sky-divers' decision-making. Males, and those who perceived less risk in the situation, were most likely to jump.

Pilots may fly into deteriorating conditions because of inaccurate risk perception. Pilots may be especially vulnerable to this because their training tends to make them overconfident in their abilities, thus leading them to underestimate risks (Goh & Wiegmann, 2001b; O'Hare, 1990).

Accident databases and questionnaires have shown that pilots who fly VFR-into-IMC tend to have inaccurate risk perception. Goh and Wiegmann (2001b, 2002a) analysed NTSB and FAA accident databases. Both studies found that more VFR-into-IMC accidents were caused by pilots' overconfidence than other GA accidents, indicating that the pilots who flew VFR-into-IMC did not recognise the risks involved in flying into adverse weather. Hunter (2002) had pilots complete a series of scenarios depicting situations occurring in GA. Pilots were asked to rate how risky each scenario appeared. Pilots also completed the Hazardous Event Scale (HES; Hunter, 1995). This measured the pilots' involvement in hazardous aviation events, for example, stalling or flying VFR-into-IMC. There was a negative correlation between risk perception and involvement in hazardous activities. That is, pilots who gave lower ratings of risks to the scenarios were involved in more hazardous events. This finding was replicated two larger studies (Hunter, 2005, 2006).

There has been mixed evidence regarding the relationship between risk perception and flying into adverse weather in a simulated flight. O'Hare (1990) had pilots assess the likelihood of them being involved in an aviation accident. Pilots on average greatly underestimated the likelihood of them being involved in an accident. Younger and less
experienced pilots underestimated the risks involved in GA, while more experienced pilots underestimated the likelihood of accident involvement. However, there was no significant relationship between pilots’ decision to take-off in marginal weather during a simulated flight, and pilots’ risk perception. In Goh and Wiegmann’s (2001a) study, pilots completed a similar risk perception questionnaire. Pilots who continued into simulated adverse weather tended to underestimate the risk associated with GA, compared to pilots who diverted. However, this relationship did not reach statistical significance. Molesworth, Wiggins, and O’Hare (2006) also found little relationship between risk perception and risk-taking behaviour during a low altitude flight. There was no relationship between the minimum altitude which the flight was conducted and the assessment of risk involved in doing so. O’Hare and Wiegmann (2003) did find a relationship between flying into adverse weather and risk perception. Pilots, who continued into adverse weather, gave lower ratings of the perceived risk, compared to other pilots.

**Risk tolerance:** While risk perception is essentially a cognitive activity, it may be argued that risk tolerance is part of an individual’s personality. Risk tolerance relates to the amount of risk an individual is willing to accept in a given situation (Hunter, 2002). One reason why people take risks, such as pilots flying VFR-into-IMC, may be because their level of acceptable risk is set too high. Risk tolerance has been measured in a variety of ways: involvement in risk-taking behaviour, gap acceptance, risky choice selection, self-report, and personal minimums.

Researchers who measure risk tolerance through involvement in risky activities (e.g. Platenius & Wilde, 1989; Sicard, Tailllemite, Jouve, & Blin, 2003; Turner & McClure, 2004), assume that participation in high risk behaviours, such as speeding and sky diving, indicate that the participants have a high risk threshold. Therefore, the more high risk behaviours the participant is involved in, the more risk tolerant he or she is. Equating risk tolerance with risk-taking is problematic. First, this confounds risk perception and risk tolerance. As discussed earlier, people may take risks due to poor risk perception rather than risk tolerance. For example, drivers may speed because they do not recognise that speeding is risky, or they may believe that their driving skills counteract the danger. Furthermore, participants of high-risk activities, such as sky-diving, usually only do so after careful consideration of the risk involved. Therefore, such a person will be incorrectly labelled risk-tolerant. Finally, risk-taking is domain specific (Hanoch, Johnson, & Wilke,
2006). For example, smokers may take health risks, but not financial or physical risks. Therefore, while there may be some relationship between risk-taking and risk tolerance, it is not a sufficient measure.

Gap acceptance, which has been used to measure risk tolerance in driving and aviation, measures the maximum level of risk a participant is willing to accept in a given situation (Horswill & Coster, 2001). Gap acceptance has largely been used in the context of a driver merging with the flow of traffic (e.g., Bottom & Ashworth, 1978; Brookhuis, de Waard, & Samyn, 2004; Cooper & Zheng, 2002; Gibbs, 1968; Horswill & Coster, 2001; Horswill & McKenna, 1999; Hurst, Perchonok, & Seguin, 1968; Wolf, Algom, & Lewin, 1988). In these studies, the participants take the position of a driver at an intersection, and indicate the minimum gap between cars required for merging. This measure of risk tolerance of drivers has been found to relate to risk-taking behaviour, such as speeding, and motor vehicle accidents (Horswill & Coster, 2001). Furthermore, gap acceptance has been used to measure the effect of the drug ecstasy on a driver’s risk-taking (Brookhuis et al., 2004). This study compared gap acceptance while merging in a simulated driving situation before and after taking ecstasy. Drivers took more risks when under the influence of the drug, compared to when they were sober.

Hunter (2002) used this method to assess risk tolerance in aircraft pilots; creating three exercises to measure the level of risk a pilot was willing to accept while flying. For example, in one risk tolerance exercise, the pilot’s task was to fly from one airport to another. However, two thunderstorms blocked the shortest route between the airports. For each trial, a gap of varying width between the two thunderstorms was presented and the distance between the plane and the thunderstorms was measured. The more risk tolerant a pilot, the closer that he or she would fly to the thunderstorm. Hunter found no relationship between risk tolerance and involvement in hazardous aeronautical activities, using Hunter’s (1995) HES. As part of a larger study, Hunter (2005) replicated these results. Again, his risk tolerance measure did not correlate with involvement in hazardous aeronautical events.

Similar methodology can be used to measure risk tolerance in non-transport situations. Connelly and Isler (1996) examined children’s risk tolerance for crossing roads. Children stood on a curb, watched an approaching car, and indicated the point at which it was no longer safe to cross the road. DiLillo and Tremblay (2001) measured mothers’ tolerance of risk of injury using playground equipment. Mothers were shown a number of
pictures depicting children using playground equipment. Participants were asked to indicate the maximum acceptable risk in each situation. For example, one of the scenarios depicted a child climbing a tree to rescue a kite. In this case, the mothers were asked the maximum height that they would allow their children to climb. This was assumed to reflect the mothers’ tolerance for risk of injury.

Risk-taking involves the choice between two options, where one option involves risk of an adverse event (Slovic, 1987). Researchers have measured risk tolerance by presenting participants with alternative scenarios, with an incentive for choosing the risky option (e.g., money). If the participants choose the risky option, they are said to be risk tolerant. Thomson, Önkal, Avcioğlu, and Goodwin (2004) presented helicopter pilots with hypothetical scenarios, with a risky and safe option. For example, one scenario depicted a VFR flight with forecasted IMC. The participants choose to remain, or risk flying home in the adverse weather. There was a relationship between risk tolerance and experience, with the more risk tolerant pilots being less experienced than the more risk averse participants. However, this again confounds risk perception and tolerance. The participants may have chosen the risky option, not because they are risk tolerant, but because they did not recognise that it was risky.

A similar paradigm, used in aviation settings (e.g., Knecht et al., 2004) was developed by Lejuez, Read, Kahler, Richards, Ramsey, Stuart, Strong, and Brown (2002). Participants completed a computer task, where they pumped a balloon figure. Participants earned money with each pump of the balloon. If the balloon popped, the participant lost all of the money earned so far. Analogous to a real balloon, the likelihood of the computer balloon figure popping increased as more air was added. So, at each point, the participants had two options: to pump the balloon, which increased potential earnings while increasing the risk of losing money; or to take the money earned so far. The number of pumps was indicative of risk tolerance. Knecht et al. (2004) found that this measure did not predict pre-flight decision making.

Self-report measures have been more successful in relating risk tolerance to flying into adverse weather. O'Hare (1990) asked pilots to assess their own risk-taking propensity. He found that pilots who decided to take-off in marginal VFR conditions during the simulated flight rated themselves as more willing to take risks than those who did not continue with the flight, suggesting that these pilots were more tolerant of risks. Platenius and Wilde (1989) found a similar relationship. Risk tolerance was measured by
having pilots answer questions regarding how often they advocated and participated in risky activities. Pilots who were involved in aviation accidents had higher ratings of risk tolerance. Wiggins, Connan, and Morris (1996) found that self-perceptions of risk tolerance were related to perceived likelihood of safely completing a flight. Pilots were presented with scenarios, where they had to make an in-flight decision (e.g., the weather deteriorated and the pilots decided whether to continue, divert, or turn-back). The more willing the pilot was to take risks, the more likely it was that he or she was to think that it was possible to continue the flight, and the less advisable it was to turn around. Such studies show a relationship between self-perception of risk tolerance and ADM.

Personal Minimums is a self-report method specific to aviation. Jensen and colleagues (Jensen et al., 1998; Kirkbride, Jensen, Chubb, & Hunter, 1996) believe that through the development of personal minimums, pilots can improve their risk management, and subsequently their decision-making skills. This technique involves the pilots developing their own set of criteria which must be met before taking-off on a flight. Such criteria should be developed for all four categories of threat the pilot faces: threats within the pilot, the aircraft, the environment, and external pressures. For example, a pilot might only take-off if he or she has had at least seven hours sleep. Personal minimums are especially important for private pilots, who are not under the influence of company guidelines (Jensen et al., 1998). Driskill et al. (1995) and Hunter et al. (2003) found that pilots used compensatory methods of decision-making, which may lead to poor decision-making. Since pilots are encouraged to only take-off if all of their criteria are met, personal minimums training encourages pilots to use non-compensatory methods of decision-making (Hunter et al., 2003). Personal minimum programs have been widely accepted by pilots (Jensen et al., 1998). Pilots find the program both easy to understand and effective.

Since the development of personal minimum training, researchers have assessed the relationship between personal minimums and weather-related decision-making. O'Hare et al. (2007) assessed personal minimums for visibility and cloud ceiling on local and cross-country flights. Pilots were asked what was the lowest visibility and cloud ceiling they would accept before taking off on a flight. There was no relationship between flying into adverse weather in a simulated flight and personal minimums. However, Hunter (2001) found that personal minimums were related to accident involvement in GA pilots. Pilots who had been involved in an accident prior to the survey had more lenient personal minimums compared to pilots who had not been involved in an accident. Furthermore,
personal minimums predicted accident involvement after the survey. The stricter the pilots’ personal minimums, the less likely it was that they would be involved in an accident during the 5.5 year post-survey period.

Risk perception and risk tolerance may be involved in decision-making. Many factors have been hypothesised to influence the perception and tolerance of risk. According to Sokolowska and Pohorille (2000), perceived risk is a combination of the consequences of risk, the probability of loss, and the amount of gain. Meanwhile, risk tolerance is a trade-off between perceived risk and benefit.

Starr (1969) highlighted two factors which influence the tolerance of risk: the level of voluntariness, and the perceived benefit. Individuals accept 1000 times less risk when the activity is involuntary (such as nuclear power) compared to when the activity is voluntary (e.g., smoking). The public are also more willing to accept risk when there is some perceived benefit from taking the risk. Slovic and colleagues (Slovic, 1987; Slovic, Fischhoff, & Lichtenstein, 1982, 2000) argued that the publics’ perception and tolerance of risk is related to voluntariness, dread, knowledge, controllability, benefits, the seriousness of each death, and the number of deaths caused by the activity or hazard. Siegrist et al. (2005) found that trust and confidence influence the perception of risks.

In this section, I have discussed the importance of risk management in ADM, and in particular, WRDM. In the next section I will discuss the measurement of expert decision-making, and will describe a technique for measuring expert judgement. This technique will be adapted to measure expertise in perceiving aeronautical risks.

Expertise

Risk management is an important component of expert pilot decision-making and is an essential part of what differentiates a good from an exceptional pilot (Jensen et al., 1997). Risk perception is a key component of risk management, and is a cognitive skill (Hunter, 2002). Some pilots may fly into adverse weather because they lack the skills to recognise that the weather is deteriorating beyond their capabilities and that they should therefore turn around or divert to another airport (Wiggins and O’Hare, 2003b).

Both Jensen et al’s (1997) and Wiggins and O’Hare’s (2003b) perspectives suggest that risk perception is something which expert pilots exercise. Therefore, measuring risk perception may be approached by asking how good pilots are at perceiving aeronautical risks. Risk perception has largely been measured by presenting pilots with flight scenarios
and having them rate how risky the scenarios appear (e.g., Hunter, 2002; O'Hare, 1990). However, while this measures how much risk a pilot perceives there to be in a situation, it does not effectively measure how good pilots are at perceiving risk. This is because there is no ‘gold standard’, or no clearly correct answer to compare the judgement with (Weiss & Shanteau, 2003).

This difficulty in identifying a correct answer is common to other areas of judgement research (Weiss & Shanteau, 2003). For example, it is difficult to assess the ability of a doctor who is judging the likelihood that a cancer patient will survive, as the accuracy of the judgement can not be assessed at that time. In fact, a lack of a ‘gold standard’ is almost a requirement of expert judgement; why would we need experts if the answer is obvious (Gigerenzer & Goldstein, 1996)?

This is especially apparent in pilot judgement research. Factors that pilots use to make judgements (e.g., go/no-go decisions), and the weights applied to each factor vary across different pilots (Knecht et al., 2004). This lack of a consistent decision process makes it difficult to predict pilot decision-making and subsequently identify a ‘gold standard’. This is further highlighted in studies which show that pilots generally use a compensatory method of decision-making (Driskill et al., 1995; Hunter et al., 2003). That is, the effect that one cue has on judgement, depends on the level of another cue. For example, a pilot may decide to take-off in poor visibility as long as the cloud ceiling was at an acceptable level. Because the judgement depends on the pilot’s own combination of cues, it is difficult to predict pilot decision-making.

For situations where there is no ‘gold standard’, researchers have often used Subject Matter Experts (SMEs) (Weiss & Shanteau, 2003). A SME would be someone who was a designated expert in the field and his or her response would subsequently become the ‘gold standard’. For example, ADM researchers have used flight instructors to define expertise (Hunter, 2003; Jensen et al., 1997). However, this method is also problematic. First, this involves circular logic as the researcher has to define expertise before measuring expertise. Furthermore, having a flight instructor rating does not necessarily guarantee that a pilot is an expert in decision making (Wiggins & Henley, 1997). Flight instructors in most countries tend to be young and relatively inexperienced.

Other common ways of assessing expertise in aviation include level of certification (Adams & Ericsson, 2000) and experience (Adams & Ericsson, 2000; Wiggins, et al., 1996). Equating expertise with certification is useful in areas where individuals receive
accreditation in recognition of skill (Shanteau, Weiss, Thomas, & Pounds, 2002). For example, in aviation an expert could be defined as someone with a Commercial Pilots Licence (CPL), flight instructor’s licence, or instrument rating. However, one problem is that once a level of certification is gained, the individual is certified for life (Shanteau et al., 2002). A CPL rated pilot, who has not flown for a while, may not be an expert decision-maker. Therefore, the level of certification may not be the most appropriate method of defining expertise.

Often the number of hours experience is used to differentiate experts from novices. Usually, an individual cannot be an expert unless he or she has had adequate experience (Shanteau et al., 2002). This is true in aviation. According to Wiggins and O’Hare (1995), experience is vital to develop ADM, especially in the context of making weather-related decisions. WRDM cannot easily be developed through training.

Experience in aviation is often measured by the total number of hours flying (e.g., Adams & Ericsson, 2000; Stokes et al., 1997; Wiggins & O’Hare, 1995, 2003a). However, it is the quality of experience that is important for expertise (Charness & Schultetus, 1999; Seifert, Patalano, Hammond, & Converse, 1997). For example, a pilot who has many hours experience flying circuits may not be as expert in decision-making as someone with the same number of hours but in cross-country experience. Some researchers have defined expertise using hours flown cross-country (Wiggins et al., 1996; Wiggins et al., 2002). This is based on the assumption that flying cross-country is more important in developing ADM than flying local circuits. Therefore, task-specific experience should be used if equating expertise with experience.

While experience is essential in developing expertise, it is not sufficient (Shanteau et al., 2002). A lack of relationship between experience and performance has been found in many areas such as grain judges (Trumbo, Adams, Milner, & Schipper, 1962) and clinical psychologists (Goldberg, 1968). That is, a pilot may accrue many hours flying experience, without becoming an expert pilot.

Real-life situations that decision-makers face are not one-dimensional. Rather, decision-makers are often required to combine information when making a decision (Phelps & Shanteau, 1978; Slovic, 1969). For example, clinical psychologists have to combine information about the client, such as their upbringing, symptoms, lifestyle, culture, and health, to make a diagnosis or prescribe a treatment. Because situations involve a large amount of information, researchers have assumed that experts use more
information than novices. This is described by Shanteau (1992) as the information-use-hypothesis. This states that expert decision-makers should consider all information when making a decision. According to this theory, the more expert a judge, the more information he or she should use when making a decision. However, this is not always the case (Shanteau, 1992). For example, Phelps and Shanteau (1978) found that livestock judges only used a small number of cues when judging the quality of pigs.

The lack of relationship between expertise in decision-making and the number of cues used may be because decision-making is configural. The effect of one cue may be dependent on the level of other cues (Slovic, 1969). Since researchers assess the importance of a cue on decisions based on main effects, the effect of interactions between cues is not always measured. If the interaction was important, any main effect would be reduced, therefore reducing the number of cues significantly affecting judgements. The concept of configural decision-making is similar to Driskill and colleagues' compensatory model. At least for pre-flight decision-making, such as go/no-go decisions, the number of cues may not be indicative of expertise, since pilots process information in a non-linear manner.

Expert decision-makers may not use all available information because of the difficulty of combining a large number of factors. This is supported by Slovic (1969), who assessed the decision-making of two stock brokers. The participants were presented with descriptions of hypothetical companies and were asked to rate the likelihood that the share market price would increase in the next 6 to 12 months. One of the participants helped create the scenarios by indicating which factors were important in making this assessment. Slovic found that even though this participant recognised that all of these factors were relevant, he was still only able to use a small proportion of factors to make the decision. Experts deal with a large amount of information by chunking (Chase and Simon, 1973; Einhorn, 1974). That is, if an expert judge assumes that cues often go together (e.g., when flying, rain is often associated with low visibility), then they will cluster these cues together. In this way, judges can greatly reduce cue processing. Chase and Simon (1973) assessed the memory of three chess players for chess positions. The expert chess player was able to remember more and larger chunks of information than the less experienced players.

This dissociation between the number of cues used and experience may occur because it is the way in which experts use the information, rather than the quantity of cues,
which is important (Ettenson, Shanteau, & Krogstad, 1987; Shanteau, 1992). Ettenson et al. (1987) found that experts are more likely than novices to use relevant cues when making a decision. Accounting students (novices) and professional auditors (experts) were presented with hypothetical scenarios describing a firm and were asked to judge the materiality of the information. Each scenario was also presented twice. From this, the researchers measured how consistent each participant was. The results found that experts and novices use information differently when making judgements. There was no difference between the auditors and the students in the number of cues that the participants used to make a decision. However, the auditors were more consistent in their use of cues than the students. Furthermore, the cues that the auditors used were more relevant to the materiality of the information than the cues that the students used. Chase and Simon (1973) also found that the recall of chess positions depended on the relevance of the scenarios. The expert chess player was only able to recall more chunks of information than the other players when the chess pieces were displayed in a realistic manner. There was no difference in memory for chess positions when the pieces were randomly placed on the chess board.

Expert pilots also use information differently compared to novice pilots. Stokes et al. (1997) presented pilots with flight scenarios. Pilots indicated the cues which were important for identifying an in-flight problem. Highly experienced pilots tended to identify more relevant cues as being important than the novice pilots. Furthermore, experienced flight instructors were better able, and faster, at selecting relevant cues when deciding whether to send a student on a solo flight (Wiggins & Henley, 1997).

Not only are experts more likely to use relevant information than novices, but teaching novices to ignore irrelevant information can improve decision-making. Gaeth and Shanteau (1984) had soil judge students assess level of sand, silt, and clay in the soil by feel. Any other substances in the soil, such as moisture, were irrelevant for the task. The students were heavily influenced by irrelevant materials in the soil. The students were then given decision-making training, designed to reduce the effect of irrelevant information on the assessments. This training successfully reduced the influence of irrelevant information and increased the accuracy of the assessment. This suggests that ignoring irrelevant information leads to improved decision-making, further highlighting the importance of the type of cue used on decision-making. Such research suggests that it is the type of information used that is important, rather than the amount of information, when determining whether someone is an expert judge (Shanteau, 1992).
The above discussion highlights that expertise is difficult to measure and define. Rather than stating that experts are those who perform at the highest level, Einhorn (1974) aimed to find some objective criteria for measuring expert judgement. Following this, Einhorn (1974) believed that intra-judge reliability and consensus among experts are necessary conditions for expertise. However, according to Einhorn, these were not sufficient for expertise. That is, an individual may be reliable and agree with his or her peers, but still not be classed as an expert. This is analogous to the relationship between the reliability and validity of a measure. While reliability is a necessary condition for validity, a reliable measure is not necessarily valid (Cronbach, 1970).

Test-retest reliability is one way of measuring intra-judge reliability. Judges are asked to rate a set of scenarios twice within a time period. The scores are then correlated; a high correlation suggests a strong relationship between the two sets of judgements, indicating that the judges are reliable. This is a very important characteristic for expert judgement. Lay people often rely on experts to make decisions, such as judgements made by doctors, pilots, meteorologists, and radiologists; therefore, it is important that these judges are reliable (Ashton, 2000). A medical professional who changes a diagnosis after a second consultation could cause emotional strain or a delay in medical treatment (Ashton, 2000). Ashton (2000) conducted a meta analysis on studies measuring test-retest reliability in judges. The studies assessed judgement in a variety of contexts; such as decisions made by medical professionals, clinical psychologists, meteorologists, and auditors. Across all of the studies, Ashton found an average test-retest reliability of 0.78. Therefore, the first judgement made by the participants accounted for 61% of the variance in the participants’ second judgement.

Einhorn’s second criterion for expertise is inter-judge consensus. That is, experts should agree with one another. However, research has shown that in some areas of judgement, there is little agreement between experts (e.g., Einhorn, 1974; Hoffman, Slovic, & Rorer, 1968). For example, in Hoffman et al.’s (1968) study, radiologists were presented with hypothetical patients and rated the severity of stomach ulcers. The consensus between radiologists ranged from .11 to .83. There are a number of reasons why experts may not agree with one another. Consensus may not be reached because of a difference in the way that individual judges weight information (Einhorn, 1974). Einhorn (1974) presented three pathologists with biopsies who were asked to rate the severity of the disease. The participants rated different cues as being important. This difference in perception of the
importance of cues may be because of a difference in training and experience. Consensus may not be reached due to the task characteristics. In many areas of expert judgements the situations are dynamic and involve more than one correct answer. Inter-judge reliability is lowest in those who work in dynamic settings, particularly involving humans, such as psychiatrists (Shanteau, 2001).

An important facet of expertise is the ability to differentiate between similar, but not identical, stimuli or situations (Weiss & Shanteau, 2003). For example, an expert doctor should respond differently to patients who present different symptoms. Discrimination can be measured by the variation in responses. Differences between novices and experts have been found in the ability to discriminate between stimuli with expert and novice snooker players (Abernethy, Neal, & Koning, 1994), American football followers (Werner & Thies, 2000), chess players (Holding, 1979), and between risk assessors and lay people (Slovic, Fischhoff, & Lichtenstein, 1985).

This difficulty in measuring expertise led Weiss and Shanteau (2003) to develop an empirical measure of expertise: the Cochran – Weiss – Shanteau (CWS) scale. Using Cochran’s (1943) theory, Weiss and Shanteau (2003) assumed that the two important components of expertise were discrimination and consistency. That is, an expert should be able to discriminate between two relevant stimuli, and do so consistently. Weiss and Shanteau (2003) presented an example of four doctors who were given vignettes of patients. The doctors’ task was to evaluate the likelihood of the patients developing heart disease. Each vignette was presented twice. An expert doctor should give different probability ratings to different patients. That is, the doctor should be highly discriminating. This same doctor should also give very similar ratings to vignettes of the same patient. That is, the doctor should be consistent.

Following this, the CWS calculates expertise using a ratio where:

\[
\text{CWS} = \frac{\text{Discrimination}}{\text{Inconsistency}}
\]

Discrimination is calculated as the variance among responses to different stimuli. Inconsistency is the variance among responses to the same stimuli. The CWS is calculated so the higher the ratio, the more consistent the judge.
The CWS has successfully been used to discriminate between expert and novice doctors, auditors, livestock judges, personnel selectors (Shanteau et al., 2002; Weiss & Shanteau, 2003), and air traffic controllers (Thomas & Pounds, 2002; Thomas, Willem, Shanteau, Raacke, & Friel, 2001, 2002). For example, Thomas et al. (2002) created air traffic control scenarios whereby complexity (number of aircraft in the sector) of the scenarios was manipulated. The aircraft in the simulation had different routes of varying distances. The measured variable was the distance each aircraft flew. Discrimination was measured as the extent to which aircraft with different routes flew different distances. Each scenario was replicated, and inconsistency was measured as the extent to which the same aircraft flew different distances. As a comparison, SMEs rated the performance of the air traffic controllers. The CWS scores differed across the different levels of complexity. However, the SMEs’ ratings did not, suggesting that the CWS score was a more sensitive measure than the SMEs.

One of the consistent findings of research on expert judgement is that when making a decision, experts are more likely than novices to use relevant cues. To further assess this, Weiss and Shanteau (2003) re-analysed Ettenson et al.’s (1987) results using the CWS methodology. Since Ettenson et al. had separated the cues according to relevance, Weiss and Shanteau were able to calculate a CWS score for relevant and irrelevant information. This measured the effect of relevance of the cues on the participants’ expertise in decision-making.

The mean square values were calculated for each cue, which represented the effect of the cue on the judgement made. A high mean square indicated that the cue contributed to the judgement of materiality. Using the mean square values, a separate CWS score was calculated for relevant and irrelevant cues, for each participant. As there were two replications, there were two mean square scores for each factor.

Discrimination was calculated as the variation of the mean squared values for different cues. This was measured by averaging the mean squared values for each cue across replications, and then taking the variance of these values. This score represented the level of discrimination; the higher the score, the more variation in the mean square values. A high discrimination score signified that some factors were significant, while some were not, whereas a low score represents less variation. A low score could mean that all factors are highly significant, no factors are significant, or something in the middle of these two extremes.
Inconsistency was calculated as the variation in mean square values for the same cue. These variations were then averaged to get the inconsistency score, for relevant and irrelevant cues. The higher the inconsistency score, the more inconsistent participants' were in their use of factors. That is, a high score indicated that participants' were influenced by some factors in one replication, but not influenced by these factors in another replication. Meanwhile, a low score would indicate consistent use of factors.

Weiss and Shanteau (2003) found that the difference in CWS scores between experts and novices reduced as the relevance of the cue declined. With relevant cues, the experts were more consistently discriminating than the novices. Meanwhile, with the irrelevant cues, there was little difference in CWS scores for experts and novices.

In the next chapter, I will describe several studies, where I will use the CWS methodology to measure expertise in aeronautical risk perception. In the first study, I will measure expertise in risk perception in participants who are naïve (undergraduate psychology students) and experienced (trained pilots) in aeronautical decision-making. In the second study, I will measure expertise in aeronautical weather-related decision-making in expert (qualified pilots), novice (student pilots) and naïve (geography students) participants. Finally, in the third study, I will measure risk perception in aeronautical weather-related decision-making in expert (qualified pilots) and naïve (geography students) participants, where I will control for memory ability.
Chapter 3: Expertise in Aeronautical Risk Perception

The aim of the following three studies is to use the CWS procedure to measure expertise in aeronautical risk perception. Participants, who differ in their level of ADM experience, will be presented with flight scenarios. Each scenario will be presented twice, and the participants will be asked to rate the risk involved in each scenario on a scale from 0 (low risk) to 100 (high risk). From these ratings I will calculate a CWS score, measuring expertise in aeronautical risk perception. If expertise in aeronautical risk perception is related to ADM, there should be a relationship between experience in ADM and the CWS scores.

Study 1

The aim of study one is to use the CWS procedure to measure expertise in aeronautical risk perception. Ten pilots will be presented with 32 flight scenarios, each presented twice, depicting a VFR GA flight. Using Jensen et al.'s (1998) PAVE model as a framework, each scenario will present information about the pilot, aircraft, environment (weather and terrain), and external factors (e.g., tomorrow's weather forecast). As a comparison, the scenarios will be presented to 20 undergraduate psychology students with no piloting experience. Participants will rate the risk involved in each scenario on a scale from 0 (low risk) to 100 (high risk).

From these ratings I will calculate a CWS score, measuring expertise in aeronautical risk perception. If risk perception is related to pilot decision-making, the CWS scores should be related to experience in ADM. The pilots should be better at this task than the non-pilots, as the pilots have experience in making aeronautical decisions. Therefore, the mean CWS score of the pilots should be higher than the mean CWS score of the non-pilots.

When making a decision, experts are more likely than novices to use relevant information (Gaeh & Shanteau, 1984; Stokes et al., 1997; Wiggins & Henley, 1997). ADM is a complex process (Wickens, 1999), and pilots have to consider many different cues when making a decision (e.g., go/no-go decision). Some cues will be relevant to the decision, while other cues will not be relevant. To assess the effect of the relevance of the cue on decision-making, the flight scenarios will include relevant and irrelevant cues.
These cues were chosen through consultation with past research, and differed in their relevance to the safety of the flight. Following Weiss and Shanteau (2003), a separate CWS score was calculated for relevant and irrelevant cues. Since experts attend more to the relevant cues than novices (Stokes et al., 1997; Wiggins & Henley, 1997), I expect that with the relevant cues, the pilots will have a higher CWS score than the non pilots, but that there will be little difference between the pilots and non-pilots’ scores with the irrelevant cues.

Method

Participants
The 10 pilots were males aged between 20 and 64 ($M = 45.44$, $SD = 14.60$). These pilots had an average of 9856 total flight hours ($range = 250 - 20000$, $SD = 7581.20$) and had spent an average of 7114 hours as pilot-in-command ($range = 125 - 18000$, $SD = 6876.86$). The pilots had an average of 6591.63 hours as pilot-in-command on cross country flights ($range = 65 - 15000$, $SD = 6557.36$). Over the previous 90 days, the pilots had accrued an average of 52.66 hours as pilot-in-command ($range = 0 - 250$, $SD = 87.84$) and an average of 90.22 hours on cross-country flights ($range = 0 - 250$, $SD = 101.62$). A Private Pilot Licence (PPL), Commercial Pilot Licence (CPL), and Air Transport Pilot Licence (ATPL) was the highest level of certification held by 20, 10, and 70 percent of the pilots, respectively. The pilots had held this licence for an average of 16.9 years ($range = 1 - 42$, $SD = 13.39$). An instrument and flight instructor’s rating was held by 80 and 70 percent of the pilots, respectively. These pilots were recruited through the local aero club and through contacts of the experimenter.

There were 4 male and 16 female non-pilots who were between 17 and 31 years of age ($M = 20.6$, $SD = 3.085$). These participants were psychology undergraduates from the University of Otago and received course credit in return for participating. None of the non-pilots had ever piloted a plane. However, five participants had flown a computer-based flight simulator. Five of the participants had family members or close friends who were pilots. All participants gave their informed consent.
Materials

Flight Scenarios: The present experiment used 32 pen-and-paper flight scenarios to assess the participants' expertise in risk perception (see appendix A). The flight scenarios depicted a GA flight, flown VFR, by a pilot with a CPL. Each scenario was repeated, giving a total of 64 scenarios for the participants to complete.

Each scenario comprised 14 factors, with two levels of each factor. A full factorial design would need 16,384 scenarios. However, the present experiment required only 32 scenarios. A full factorial design with 32 scenarios would have only 5 factors. In keeping with the complex nature of ADM (Wickens, 1999), 14 factors were included in the scenarios, using a 1/512 fractional factorial design. The creation of fractional factorial designs is discussed by Montgomery (2000). The 32 scenarios were created by generating a full factorial design with 5 factors. Each level of the factor was represented as either a negative or positive sign. All five factors were signified by a letter from A to E. Therefore, each scenario was made up of a unique combination of levels of the 5 factors.

According to Montgomery (2000), with fractional designs, the higher order interactions are confounded, or aliased, with main effects. Main effects can only be estimated, assuming that the higher order interactions are not meaningful. These higher order interactions cannot be calculated. The 9 new factors (factors F to O) were aliased with three-way interactions of the 5 factors. For example, the 6th factor, F, was aliased with the interaction of A, B, and C. The 9 new factors were created by multiplying the values of the factors associated with the alias structure. For example, factor F was created by multiplying A * B * C. For scenario 1, factors A, B, and C were represented by -1, +1, and +1, respectively. Therefore, for scenario 1, factor F was -1 * +1 * +1, or -1. Factors A, B, C, and F, were represented by gender of the pilot, type of passengers, terrain, and whether the pilot filed a flight plan, respectively. Therefore, in the first scenario, these four factors were represented by a female pilot, flying two company employees from Dunedin to Christchurch, who did not file a flight plan. The other 8 factors were created in a similar fashion. Each factor was represented by the following aliases: G = ABD, H = ABE, J = ACD, K = ACE, L = ADE, M = BCD, N = BCE, O = BDE. This process was repeated for all 32 scenarios. An example of a scenario is presented below in Figure 3.1.
One Thursday, a female pilot prepared to take her two company employees for a familiarisation flight from Dunedin to Christchurch. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 5 minutes before take off. She did not file a flight plan. The pilot had few hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of more than 200 rpm. The en-route weather was forecast as good visibility and high cloud ceiling. The forecast for tomorrow was good.

Figure 3.1. An example of a scenario.

Some of the factors increased or decreased the likelihood of having an accident or the likelihood of surviving an accident, and therefore were termed relevant. Meanwhile, some factors did not influence this, and were termed irrelevant. Table 3.1 presents factors, levels of the factors, and the relevance of the factors. Each factor was termed relevant or irrelevant based on past research.

Table 3.1.
Factors, Levels of Each Factor, and Relevance of the Factors Used in the Scenarios.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>Female</td>
<td>Irrelevant</td>
</tr>
<tr>
<td>Passengers</td>
<td>Company employees</td>
<td>Tourists</td>
<td>Relevant</td>
</tr>
<tr>
<td>Terrain</td>
<td>Flat</td>
<td>Mountainous</td>
<td>Relevant</td>
</tr>
<tr>
<td>Complexity of aircraft</td>
<td>Cessna 172N</td>
<td>Piper Seneca</td>
<td>Relevant</td>
</tr>
<tr>
<td>Year aircraft was built</td>
<td>1996</td>
<td>1984</td>
<td>Irrelevant</td>
</tr>
<tr>
<td>Pre-flight check</td>
<td>2 hours</td>
<td>5 minutes</td>
<td>Relevant</td>
</tr>
<tr>
<td>Flight Plan</td>
<td>Filed</td>
<td>Not Filed</td>
<td>Relevant</td>
</tr>
<tr>
<td>Experience (hours)</td>
<td>Many</td>
<td>Few</td>
<td>Relevant</td>
</tr>
<tr>
<td>Sleep</td>
<td>8 hours</td>
<td>4 hours</td>
<td>Relevant</td>
</tr>
<tr>
<td>Health</td>
<td>Good Health</td>
<td>Head Cold</td>
<td>Relevant</td>
</tr>
<tr>
<td>Magneto Drops</td>
<td>Low</td>
<td>High</td>
<td>Relevant</td>
</tr>
<tr>
<td>Visibility</td>
<td>Good</td>
<td>Poor</td>
<td>Relevant</td>
</tr>
<tr>
<td>Cloud Cover</td>
<td>High</td>
<td>Low</td>
<td>Relevant</td>
</tr>
<tr>
<td>Weather tomorrow</td>
<td>Good</td>
<td>Poor</td>
<td>Relevant</td>
</tr>
</tbody>
</table>
Gender: The pilots in the scenarios were described as male or female. Due to the low proportion of female pilots, it is difficult to assess the effect of gender on pilot safety (O’Hare & Chalmers, 1999). In case-control studies, where the characteristics of accident-involved pilots are compared to nonaccident-involved pilots, females typically account for only one to four percent of the sample (Li, Baker, Grabowski, Qiang, McCarthy, & Rebok, 2003; Li & Baker, 1999; NTSB, 2005). Vail and Eckman (1986) found that males were more likely than females to be involved in overall accidents and fatal accidents. However, Vail and Eckman’s study did not control for exposure to hazardous events (Li, 1994). Males have higher total and recent hours than females (O’Hare & Chalmers, 1999), which may account for the higher accident and fatality rates. This is supported by other research. Li and Baker (1999) found no relationship between gender and fatality rates. McFadden (1999) found that after controlling for age, experience, and risk exposure, gender was not related to incident involvement. Therefore, there seems to be little evidence for a relationship between gender and safety.

The year in which the aircraft was built: The aircraft was described as being built in 1984 or 1996. New Zealand aircraft must be inspected yearly or after every 100-hours of operation (CAA, 2003). Therefore, as long as the aircraft is up to date on these checks, the year that the aircraft was built should be irrelevant.

Experience in type: The pilots were described as having many or few hours experience in the depicted aircraft model. One predictor of accident and incident involvement is the number of hours experience in the type of aircraft (AOPA, 2006; Batt & O’Hare, 2005; Collins & Dollar, 1996). For example, AOPA (2006) reported that 43% of fatal and non-fatal accidents involved pilots with less than 100 hours in the type of aircraft. The accident and fatality rate decreased after gaining 100 hours.

Sleep: The pilots were described as having eight or four hours sleep the night before. Good sleep management is recognised by the CAA (2006) and Murray (1997) as determining whether pilots are fit to fly.

Health: The pilots were described as being healthy or suffering a head cold. Health is also important for pilot safety. Even a cold can be dangerous. Hunter (2003) found that SMEs recommended that a pilot with a cold should not fly.

Complexity of the aircraft: The aircraft depicted in the scenarios were either a Cessna 172N (single engine with a fixed undercarriage, and a fixed-speed propeller) or a Piper Seneca (twin-engine with a retractable undercarriage, and constant speed propellers). The
more complex the aircraft, the greater chance of a fatality, possibly because of the higher speeds upon impact (AOPA, 1999, 2003, 2006; Li & Baker, 1999; O’Hare, Chalmers, & Scuffham, 2003).

**Pre-Flight Check:** The pilot in the scenarios arrived at the airport to do his or her pre-flight check either five minutes or two hours before take-off. Mechanical problems accounted for 18% of all GA accidents in the United States, during 2002. Many of these problems could have been picked up on through thorough pre-flight checks (AOPA, 2003).

**Magneto Drop:** A mechanical problem was represented by the magneto drop, which was described as less than 50 rpm or greater than 200 rpm. Magneto drop is the difference in the RPM between the two ignition systems. A large difference represents a possible problem. For both the Cessna 172 and the Piper Seneca, the difference in RPM between the two magnetos should be less than 50 rpm.

**Terrain:** There were two routes represented in the flight scenarios, which differed according to the type of terrain. The Dunedin to Milford Sound flight was mountainous, while the Dunedin to Christchurch flight was largely over flat terrain. There is more risk involved in flying over mountainous terrain compared to non-mountainous terrain (TSBC, 1990). An accident over mountainous terrain and water is more likely to result in a fatality than flight over flat terrain (O’Hare et al., 2003).

**Visibility and cloud ceiling:** The visibility was described as either good or poor, and the cloud ceiling was either high or low. As discussed in the previous chapter, weather is one of the pilots’ most important considerations. Weather-related accidents, especially those due to spatial disorientation, are most likely to occur in poor visibility and low cloud ceiling (Collins & Dollar, 1996; NTSB, 2005). Furthermore, crashes involving poor visibility and low cloud ceiling are seven times more likely to be fatal than accidents which occur in fine weather (TSBC, 1990).

**Passengers:** As a manipulation of social pressure, the flight was described as either a familiarisation or a tourist flight operation flight. If the passengers were tourists, the pilots may feel more pressure to fly because they are getting paid. Meanwhile, since the familiarisation flight is costing the company money, the pilots would feel less pressure to fly. Therefore, there is more risk of ‘get-there-it-is’ with the tourist flight operation flight.

**Weather tomorrow:** As another manipulation of pressure to complete the flight, the weather tomorrow was either forecast as good or bad. If tomorrow’s forecast was for bad weather, the pilot should feel more pressure to continue to his or her destination.
**Flight plan:** The participants were described as not filing a flight plan, or filing a VFR plan. It is highly recommended that pilots file a flight plan before taking off. Flight plans include information regarding the destination and departure points, whether VFR or IFR, and estimated time. Such information will help rescuers following an accident.

The flight scenarios included space for the participants to indicate the level of risk perceived in the situation. This rating was made on a scale of 0 to 100, where 0 and 100 indicated that the flight was perceived as least and most risky, respectively.

Participants were also given a sheet containing extra information (see appendix B). This included a picture and brief description of the aircraft depicted in the scenario. This sheet also informed the participants that the flight was conducted under VFR, and that the pilot depicted in the scenarios had his or her commercial licence. The non-pilot participants were also given a paragraph describing a typical GA operation. This was to give the participant an indication of what normally happens when a pilot arrives at the airport prior to their flight.

*Demographic and Aeronautical Practices Questionnaire:* The pilots were given a 5-page questionnaire measuring demographic information and past and present aeronautical practices (see appendix C). The questionnaire asked for information regarding the pilots’ age, gender, and flying experience. The pilots were required to indicate the total number of hours flown, as pilot-in-command, pilot-in-command on cross country flights, and the number of hours flown cross-country and as pilot-in-command over the previous 90 days. Pilots were also asked to state the highest level of certification held, the length of time the certification had been held, and whether they had ever held an instrument or flight instructor rating. Finally, pilots were asked if bad weather had ever forced them to land at an airport other than their intended destination. The pilots also rated the percentage of time they performed a number of behaviours in a cross country flight (usual aeronautical behaviours) and the number of times during the previous 24-months they had been involved in a potentially hazardous situation on a scale of 0 to 4 or more (Hunter’s (1995) HES). Finally, the pilots indicated the lowest level of visibility and cloud cover they would accept before taking off on a cross country and local flight (personal minimums).

The pilots’ involvement in hazardous events was measured using Hunter’s (1995) Hazardous Event Scale (HES). Hunter (1995) devised this questionnaire as a proxy measure for accident involvement. The number of previous incidents which the pilots have
been involved in predicts the likelihood of being involved in an accident in the future (Hunter, 2001). This scale was found to be related to risk perception (Hunter, 2002, 2005, 2006), whereby the more incidents experienced, the less risk was perceived in hypothetical scenarios.

Demographic Information Questionnaire: The non-pilots were given a one-page questionnaire measuring demographic information, interest, and experience in aviation (see appendix D). The participants indicated their age and gender, experience in piloting planes, and whether they had any family members or close friends who were pilots. Participants were also asked to rate their interest in aviation on a scale of 1 to 7, where 1 and 7 referred to least and most interested, respectively.

Procedure
Participants were tested individually or in small groups. Participants read the information sheet and signed the consent form. The participants were then handed the pen-and-paper flight scenarios and were given the following instructions by the experimenter:

Thank you for agreeing to take part in this study. You will be presented with a number of scenarios depicting general aviation flights. Your task will be to read each scenario and rate how risky the scenario appears on a scale of 0 to 100. 0 will indicate that you think that there is a very low level of risk and 100 will mean that you think that it is the riskiest situation imaginable. There is no right or wrong answer, and your responses will be treated in the strictest of confidence. As there are a large number of scenarios, please take a break whenever you feel tired.

Additionally, the non-pilots were given a definition of a General Aviation flight. They were also informed that the scenarios were designed for pilots, and as such, there may be technical terms that they do not understand. They were told to ignore such terms. This was to ensure that emphasis was not placed on any particular factor. All participants were then given the extra information sheet.

The participants were given all of the scenarios together. The 32 scenarios were presented in a randomised order. This order was then repeated to give the 64 scenarios. The order of presentation was constant between participants. After reading each scenario, the participants wrote the risk rating in the space provided.

After completing the 64 scenarios, the participants were required to complete the relevant demographic questionnaire. The pilots were instructed to fill out the questionnaire as accurately as possible. The questionnaire, like every other aspect of the experiment, was
anonymous. Due to logistical problems, one non-pilot participant did not complete the demographic questionnaire. Similarly, due to the lack of recent VFR flying one pilot did not complete the usual aeronautical practices, past aeronautical practices, and personal minimums questionnaire. At the end of the experiment the participants were debriefed and thanked for their participation.

Results

Data Inspection

Following Weiss and Shanteau (n.d.), the CWS scores were transformed using a square root transformation. Both raw and transformed data are presented, represented by the subscript “r” and “t”, respectively. All inferential tests are calculated using the transformed data. The distribution of the residuals was assessed. If the residuals were not normally distributed, the variables were transformed using a square root transformation and outliers more than 3SDs away from the mean were removed. If this did not render the data normally distributed, a non-parametric test was used, with outliers included. All statistical tests were evaluated against an alpha level of .05.

Effect sizes for significant and marginally significant results were calculated using Cohen’s (1988) $d$ and partial eta squared ($\eta_p^2$). Cohen’s $d$ is calculated as the difference between the two means, divided by the pooled standard deviation. According to Cohen (1988), effect sizes of 0.2, 0.5, and 0.8 are regarded as small, medium, and large effects, respectively. Cohen’s $d$ also provides an estimate of the percentile standing of the experimental group, and the percent of non-overlap between the two distributions. Cohen’s $d$ was calculated when comparing two groups. Partial eta squared was calculated when an ANOVA was used. Partial eta squared represents the level of variance of the dependent variable predicted by the independent variable. According to Cohen (1988), .01 (1% of variance), .06 (6% of variance), and .14 (14% of variance) represent a small, medium, and large effect size, respectively.
CWS Analysis

Each participant gave two risk ratings for each scenario. These risk ratings were used to evaluate expertise in risk perception. This was measured by calculating discrimination and inconsistency.

Discrimination represents the amount of variation in ratings for different scenarios. Discrimination was calculated by averaging the ratings for each scenario, calculating the variance of the averaged ratings, and then doubling the variance. The higher the score, the more the participant discriminated between different scenarios. Inconsistency represents the variation in ratings for the same scenario. This was determined by calculating the variance of the ratings for the same scenario. These variances were then averaged to get an overall inconsistency score. The higher the score, the more inconsistent the participants were at judging aeronautical risks. The CWS score was calculated by dividing discrimination by inconsistency. Therefore, the more a participant discriminated and the more consistent he or she was, the higher the CWS score. The higher the CWS, the more expert the participant was at judging aeronautical risk.

The CWS scores ranged from 2.05 to 19.42 ($M = 7.06, SD = 3.99$). The CWS scores were averaged across pilots and non-pilots. The Levene’s Test for equality of variances revealed that the pilots’ CWS scores ($M = 9.45, SD = 5.15$) varied more than the non-pilots ($M = 5.87, SD = 2.70$), $F(1,29) = 6.608, p = .016$. Following transformation, the variances were equal, $F(1,28) = 2.72, p = .110$. The mean CWS score for the pilots ($M_r = 9.45, SD_r = 5.15, M_l = 2.96, SD_l = .88$) was significantly higher than the non-pilots ($M_r = 5.87, SD_r = 2.70, M_l = 2.36, SD_l = .58$), $F(1, 28) = 5.038, p = .033, \eta^2_p = .152, d = 0.81$. The pilots’ mean CWS score was in the 79th percentile of the non-pilots scores, and there was 47.4% non-overlap between the two distributions. Figure 3.2 shows the mean raw CWS scores for pilots and non-pilots.
Discrimination ranged from 104.79 to 1877.37 (M = 525.39, SD = 359.85), while inconsistency ranged from 13.81 to 227.73 (M = 88.15, SD = 57.24). Discrimination and inconsistency of pilots and non-pilots were compared. The pilots and non-pilots did not significantly differ in the level of discrimination (pilots: M = 459.07, SD = 217.79; non-pilots: M = 287.63, SD = 279.07), $F(1,27) = .074$, $p = .788$. There was no significant difference between pilots and non-pilots in the level of consistency between the same scenarios (pilots: $M = 83.25$, $SD = 67.58$, non-pilots: $M = 90.60$, $SD = 53.09$), $F(1,28) = .106$, $p = .747$.

**Relevance of Cues**

The cues included in the scenarios were termed either relevant or irrelevant, according to whether the cues affected the safety of the flight. The extent to which participants used relevant and irrelevant cues was measured by calculating a CWS score for the two types of cues. Following Weiss and Shanteau (2003), the mean square value for each cue was calculated for each replication. For example, for each participant, the mean squared values were calculated for visibility, for both replications. Therefore, for each
participant there were mean squared values for each cue and for each replication. CWS scores were calculated for both relevant and irrelevant cues.

Discrimination was calculated as the variation of the mean squared values for different cues. This was measured by averaging the mean squared values for each cue across replications, and then taking the variance of these values. This variance was then doubled to get the discrimination score. The higher the value, the more likely that the participant used some cues more than others when making risk judgments.

Inconsistency was calculated as the variation in mean square values for the same cue. These variations were then averaged to get the inconsistency score. The lower the value, the more likely the participants used the same cues in both replications when making risk perception judgments. CWS was calculated by dividing discrimination by inconsistency. The higher the CWS score, the more the participants discriminated between the cues and the more consistently the participants used the same cues for both replications when making risk perception judgements.

There was a main effect of relevance of cue. That is, on average, participants had a higher CWS score for the relevant cues ($M_r = 15.97, SD_r = 26.12, M_i = 3.28, SD_i = 2.32$) than for the irrelevant cues ($M_r = 1.36, SD_r = 3.03, M_i = .84, SD_i = .82$), $Z(1,29) = -4.57, p < .001, d = 1.40$. This indicates that participants were more discriminating and/or consistent when the cues were relevant to the task than when the cues were irrelevant. The mean CWS score for relevant cues was in the 91.9th percentile of the scores for irrelevant cues, with 79.4% non-overlap between the two distributions. Figure 3.3 shows the raw mean CWS score when the cues were relevant and irrelevant.
The CWS scores for relevant and irrelevant cues were compared for pilots and non-pilots. The residuals for the CWS scores were not normally distributed, but there is no non-parametric equivalent of repeated measures test. However, there was no significant interaction between the type of participant and the type of cue, using parametric tests. That is, the relationship between the CWS score and the relevance of the cue did not differ between pilots and non-pilots, $F(1,28) = .157, p = .695$.

The effect of the relevance of the cues was also assessed by looking at the mean square values. Each participant's mean square values were averaged across all of the cues, for relevant and irrelevant cues. A repeated measures ANOVA was calculated, with relevance of the cue (relevant and irrelevant) as the within-subjects variable, and the experience level (pilot and non-pilot) as the between-subjects variable. This was conducted separately for the two replications.

With replication one, there was a main effect of relevance, $F(1,28) = 74.137, p < .001$. The participants' responses were more highly influenced by the relevant cues ($M = 592.01, SD = 422.5$), than the irrelevant cues ($M = 119.57, SD = 117.62$), $\eta_p^2 = .568, d = 1.52$. The
average mean square value for the relevant cues was at the 93.3\textsuperscript{nd} percentile for the irrelevant cues with 70.7\% non-overlap in the distributions. There was no main effect of experience level, and no interaction between the relevance of the cue and experience level.

With replication two, there was a main effect of relevance, $F(1,27) = 96.757, p <.001, \eta^2_p = .580, d = 1.68$. The participant’s responses were more highly influenced by the relevant cues ($M = 645.18, SD = 476.18$) than the irrelevant cues ($M = 75.09, SD = 70.17$). The average mean square values for the relevant cues was at the 94.5\textsuperscript{th} percentile of the mean square values for the irrelevant cues, and with 73.1\% overlap in the distribution. There was no main effect of experience level, and no interaction between the relevance of the cue and experience level.

Cues Used by Pilots and Non-Pilots.

Due to the large number of cues present in each scenario, the participants should have selected cues to attend to when making risk judgements. It was of interest to examine the cues used by pilots and non-pilots when making these judgments. For each cue, the number of pilots and non-pilots who used the cue to make risk judgments was tallied. This was repeated for both replications. This was converted into a percentage of pilots and non-pilots who used each cue. This showed which cues were most important for pilots and non-pilots when making risk judgements. Figures 3.4 and 3.5 show the percentage of pilots and non-pilots who used each cue, for replication 1 and 2, respectively.
Figures 3.4 and 3.5 show a similar pattern. Both pilots and non-pilots used a wide range of cues when making a risk-rating. Pilots attended to the complexity of the aircraft, experience, sleep, health, magneto drop, visibility, and cloud ceiling. Meanwhile, non-
pilots attended to the pre-flight checks, whether the pilot filed a flight plan, experience, sleep, health, and visibility.

**Number of Cues**

The number of cues used by participants to make risk perception judgements was calculated by totalling the number of cues that were significant for each participant. That is, a cue would be tallied if significantly different risk ratings were given, depending on which level of the factor was presented. The number of cues used by the participants was calculated for both replications. An average number of cues used was then calculated. The participants attended to between one and eight cues for the first replication \( (M = 4.13, SD = 1.871) \) and between one and nine cues for the second replication \( (M = 4.33, SD = 1.826) \).

The pilots and non-pilots did not significantly differ in the number cues used overall, or for the first \( (\text{pilots: } M = 3.90, SD = 2.03; \text{non-pilots: } M = 4.25, SD = 1.83) \) or second replication \( (\text{pilots: } M = 3.60, SD = 1.65; \text{non-pilots: } M = 4.70, SD = 1.84) \). The relationship between the number of cues used and expertise in risk perception was calculated. The CWS score was only related to the number of factors used during the first replication, \( r(29) = .389, p = .033 \). That is, the more cues used during the first replication of the scenarios, the higher the CWS score. There was no relationship between the CWS and the cues used during the second presentation or with the average number of cues used.

**Average Risk Rating**

Each scenario was presented twice in the same order; therefore, it may be interesting to examine participants’ responses across replications. For each replication, an average rating was calculated. This was repeated for each participant. A repeated measures ANOVA was conducted to compare the average risk rating for pilots and non-pilots for each replication. The average risk rating did not differ significantly between the groups of participants \( (\text{pilots: } M = 63.04, SD = 13.67; \text{non-pilots: } M = 57.24, SD = 11.31), F(1,28) = 1.38, p = .251 \). The average risk rating given for the first presentation of the scenarios \( (M = 59.42, SD = 12.28) \) did not significantly differ between the average risk rating given for the second presentation of the scenarios \( (M = 58.92, SD = 13.76), t(29) = .617, p = .542 \). The difference in ratings between the first and second presentation did not differ significantly between pilots and non-pilots, \( F(1,28) = .007, p = .934 \).
Hazard Event Scale

Pilots indicated how often they had experienced 12 potentially hazardous aeronautical situations in the past 24-months, on a scale from 0 to 4 or more. This was tallied across all 12 situations for each pilot to indicate how often he participated in hazardous situations. On average, pilots had participated in 5.55 hazardous activities (range = 0 – 18, SD = 6.02). The frequency of engaging in potentially hazardous activities did not significantly correlate with the CWS. There were marginally significant correlations between the number of hazardous events involved in, and the average risk rating given for the first (r(9) = .647, p = .06) and second replications (r(9) = .611, p = .081). This positive correlation suggests that the more potentially hazardous aviation incidents that the participants had been involved in, the greater the risk that participants perceived in the scenarios. This relationship existed while controlling for the number of hours cross-country over the previous 90-days (rating one, r(df= 5) = .687, p = .088; rating two: r(df= 5) = .669, p = .100), and the total number of hours (rating one, r(df= 6) = .664, p = .072, rating two: r(df= 6) = .608, p = .110.

Usual Aeronautical Practices

Pilots rated the percentage of time they participated in certain behaviours. Out of the 9 items, 7 were phrased so that agreement with the statement indicated safe behaviour. The other 2 items were reverse scored. The responses were averaged across all items. On average, pilots indicated that they were involved in safe aeronautical behaviours 78.27% of the time (SD = 17.56). The involvement in safe practices did not correlate significantly with the CWS score.

Personal Minimums

Pilots were asked to indicate the minimum visibility and cloud ceiling they would accept before taking off on a local or cross-country flight. Over half of the pilots would not take off if visibility was less than 5km on a local flight and 10km on a cross-country flight. Meanwhile, over half of the pilots would not take off if cloud ceiling was less than 1500 feet on a local flight and 1000 feet on a cross-country flight. There was no correlation between personal minimums and the CWS score.
Pilot Experience

The flying experience of pilots was analysed to determine if experience and age were related to expertise in risk perception. The CWS was not significantly related to age, level of certification, or holding an instrument or instructor rating. Similarly, there was no relationship between the CWS and any measure of aeronautical experience.

Non-pilot demographics

The non-pilots’ experience and interest in aviation were analysed to examine whether these factors were related to expertise in risk perception. The CWS score was not significantly related to the age of the participant, interest in aviation, or gender. However, the participants who had flown a simulator had a lower mean CWS score ($M = 2.75, SD = .22$) compared to the participants who had never flown a flight simulator ($M = 6.55, SD = 2.51$), $t(14.35) = -5.96, p < .001, d = 2.13$. The mean CWS score for those who have not flown a flight simulator is in the 97.7th percentile of the scores for those who have flown a simulator, and with an 81.1% non-overlap in the distribution.

Discussion

The CWS procedure was used to measure expertise in aeronautical risk perception. The higher the CWS score, the more able the participants were to consistently discriminate between different flight scenarios. Since the pilots and the psychology students differed in experience in ADM, I expected the pilots to be better at this task than the non-pilots. The mean CWS scores were higher for the pilots than the non-pilots. This supported the hypothesis, and suggested that the pilots were more expert at perceiving aeronautical risks than the psychology students.

The pilots varied more in their ability to perform this task than the non-pilots. While most of the non-pilots were poor at perceiving aeronautical risks, the pilots’ abilities ranged from poor to good. This supported past researchers (e.g., Hunter, 2002) who found that some pilots have poor risk perception, and this accounts for some variation in the decisions made. For example, Hunter (2002) found that pilots’ risk judgements were related to involvement in hazardous events.

Experience is a necessary, but insufficient requirement for expertise in aeronautical decision-making. Subsequently, there should be a relationship between experience and
CWS scores. However, there was no relationship between expertise and any form of experience. This may be due to the type of pilots used. An ATPL was held by 70% of the pilots, and these pilots had flown the greatest number of hours, with a mean of 8766 hours cross-country. In comparison, the three non-ATPL pilots had flown only a mean of 66 hours cross-country. The pilots with an ATPL had flown the majority of these hours for airlines. The type of experience predictive of expertise in a GA setting should be hours flown in a GA aircraft (Wiggins & O'Hare, 2003a). Both groups of pilots may have a similar number of hours experience under these conditions, which would explain the apparent lack of relationship between experience and expertise in these results.

The participants used an average of four cues to make their risk ratings, and the number of cues used did not differ between pilots and non-pilots. This supports past research (e.g., Ettenson et al., 1987; Phelps & Shanteau, 1978; Slovic, 1967), which found little difference between experts and novices in the number of cues used to make a decision. This lack of relationship may result from the compensatory or configural nature of pilot decision-making (Driskill et al., 1995; Slovic, 1967). If pilots make decisions using a compensatory method, main effects will not sufficiently capture decision-making. Unfortunately, a 1/512 fractional factorial design does not allow for the estimation of interactions, as two-way interactions are confounded with other two-way interactions.

Rather than attending to a large number of cues, experts should only attend to the relevant cues (Chase & Simon, 1973; Ettenson et al., 1982; Shanteau, 1992; Stokes et al., 1997; Wiggins & Henley, 1997). Subsequently, it was predicted that the difference in CWS scores between experts and novices should decrease as the relevance of the cues decreases. All participants had a higher CWS score when assessing the relevant cues, compared to the irrelevant cues. There was no interaction between the relevance of the cue and the experience level. This was found when using both Weiss and Shanteau’s (2003) procedure and mean square values. That is, the participants were more consistent and discriminating when the cues were relevant compared to irrelevant, and all participants used relevant cues more than irrelevant cues. Therefore, the hypothesis was not supported.

There are two possible explanations for this result. First, maybe the irrelevant cues were so obviously irrelevant that they were ignored by all of the participants. Second, only the main effects were measured for each cue. If pilot decision-making is compensatory or configural, then measuring only the main effects does not sufficiently capture pilot decision-making. For example, it is possible that the participants considered the interaction
of visibility and cloud cover, or sleep deprivation and health, when making the risk judgments. Therefore, simply measuring the main effects may not measure the influence of each cue on the participants' risk judgments.

The type of cues attended to by the non-pilots illustrates the problem of measuring main effects with a fractional factorial design. For approximately 60% of non-pilots, whether a flight plan was filed influenced the risk rating, and 20-30% of pilots were influenced by the magneto drop. However, it is not likely that the non-pilots considered either of these factors when rating the risk in the situation, due to the lack of technical knowledge. Through the use of a fractional factorial design, a large number of cues were included in the scenarios. The large number of cues included in the scenarios compensated for the inability to measure main effects in a fractional factorial design. The main goal of the study was for the participants to rate the risk involved in complex scenarios, not to assess which cues the pilots and non-pilots deemed important for making risk judgments.

During the present study, participants rated the risk involved in aeronautical situations. Following this, the pilots completed Hunter's (1995) Hazardous Event Scale. Hunter (2002, 2005, 2006) found a relationship between involvement in hazardous events and risk perception, with the more hazardous events experienced, the less risk perceived in the hypothetical situations. Hunter used this finding as support for his measure of risk perception. However, the present study contradicted Hunter's three studies. There was a trend for the participants who were involved in more hazardous events to produce higher ratings of risk. This relationship existed after both total and recent experiences were controlled for.

This study has been an encouraging first step in measuring expertise in risk perception. The CWS has been able to distinguish between novices (non-pilots) and experts (pilots) in risk perception. However, there may be a number of limitations with the study.

There was a gender difference between the two groups of participants. The pilots were all males, and the non-pilots were 80% female. This occurred because of the gender imbalance in both the pilot population and the 100-level psychology pool. However, as there was no difference between the CWS scores between the male and female non-pilots, it is conceivable that gender is not related to expertise in risk perception.

In this study, I presented participants with written scenarios. Each scenario comprised a number of different factors. For example, the scenario may have depicted an...
inexperienced pilot flying in high cloud ceiling and poor visibility. In reality, pilots do not get the information in this way. Furthermore, presenting the information in written format encourages the reader to process the scenario piece by piece, rather than as a whole. According to Adams and Ericsson (2000), experts perceive holistically, rather than sequentially. For example, when a pilot makes go/no-go decision, he or she considers all of the information together. For example, a pilot would look at the interaction of factors concerning the environment, pilot, and aircraft, when deciding to go on a particular flight. Therefore, presenting the information as written scenarios is not consistent with what pilots do in real life.

Additionally, the scenarios had to be simplified, as psychology students were used as controls. For example, visibility and cloud ceiling were described as good/poor and high/low, respectively. Again, this lacks realism, as in the real world pilots receive this information as kilometres of visibility and feet above ground level, respectively.

A second study, focussing on WRDM, will be conducted to rectify these problems. Weather-related decisions will be examined because of the importance of WRDM in aviation. Roughly two percent of all United States GA accidents, between 1990 and 1997, occurred while the pilot was flying into adverse weather. While this number is low, approximately 80% of these accidents were fatal, compared to 19% for all GA accidents (Goh & Wiegmann, 2002a).

I will use the CWS to measure expertise in aeronautical, weather-related risk perception. Pilots will be presented with 16 flight scenarios, each presented twice, depicting a VFR GA flight. Each scenario will comprise 4 pieces of information: a map of the route, and the weather at departure, en-route, and at the destination. The presentation of information will be consistent with how pilots perceive the information in real life. That is, the weather at departure and the map will be presented as a visual graphic. Meanwhile, the en-route weather and weather at the destination will be a written forecast. The participants will be asked to rate the risk in each scenario from 0 (least risk) to 100 (most risk).

Three groups of participants, with differing in experience in making weather-related decisions, will take part in the study. The experienced group will be trained pilots with at least their PPL. This group of pilots will be experienced in making weather-related decisions, such as go/no-go decisions. The pilots’ performance on the CWS task will be compared to that of student pilots and geography students. The student pilots, who will not have gained their PPL yet, will have experience in flying and will be familiar with
weather-related terminology, yet will not have any experience making weather-related decisions, which are usually made by a flight instructor. The geography students will be university students or recent graduates who have taken a climatology paper. These participants will be able to understand weather-related terminology, yet not have any experience making weather-related decisions, and will lack flight experience.

I will compare the CWS scores of the trained pilots, student pilots, and geography students. I predict that the CWS scores will reflect the participants' experience in making weather-related decisions. That is, the trained pilots will be better at making weather-related risk judgements than the student pilots and the geography students.

Study 2
Method

Participants

Qualified pilots, student pilots, and geography students were recruited for this study. The 10 male and one female qualified pilots were aged between 18 and 38 years (M = 23.64, SD = 5.887). These pilots had an average 603.45 total hours flying time (range = 165 - 2000, SD = 561.00) with between 40 and 1,850 hours as pilot-in-command (M = 471.18, SD = 546.70). The pilots had completed an average of 55.76 hours as pilot-in-command over the previous 90 days (range = 3.35 - 200, SD = 55.64). Between 20 and 1,000 hours were spent as pilot-in-command on a cross-country flight (M = 210.47, SD = 292.20), with a mean of 12.28 of these hours in the previous 90 days (range = 0 - 50, SD = 13.91). A PPL or CPL was the highest level of certification held by two and nine pilots, respectively. This certification was held an average of 2.15 years (range = 0.17 - 6, SD = 1.81). Out of the 11 pilots, six held both an instrument and flight instructors rating. One pilot had a flight instructors rating but not an instrument rating. Meanwhile, one pilot had an instrument rating but not a flight instructors rating.

The 10 male and 1 female student pilots were between 18 and 38 years of age (M = 23.55, SD = 6.67). These pilots had an average of 36 hours total flying time (range = 5 to 55, SD = 14.69) and a mean of 7.03 hours were spent as pilot-in-command (range = 0 to 13, SD = 4.49). The student pilots had spent an average of 3.85 hours flying as pilot-in-command over the previous 90 days (range = 0 to 8, SD = 2.96). Between zero and eight hours were spent as pilot-in-command on cross-country flights (M = 3.85, SD = 2.96), with
between zero and five hours in the previous 90 days ($M = 1.28$, $SD = 1.86$). Both groups of pilots were recruited through flying businesses in Queenstown, and flying schools and aero clubs in Dunedin and Invercargill.

There were three male and eight female Geography students, from the University of Otago. These students ranged in age from 21 to 28 years ($M = 22.45$, $SD = 2.25$). Out of the 11 participants, two were PhD students, eight were 4th year honours students, and one was a Masters graduate. All participants had completed a 100-level Geography course, which included a climatology component. All but two of the participants had completed a 300-level climatology paper. One of the participants who had not completed this advanced paper had tutored a climatology paper. The Masters graduate and one of the PhD students’ theses were on climatology. Therefore, all of the participants had some level of understanding of weather. Two of the students had flown a plane, although this was only an introductory flight, and three had flown a computer-based flight simulator. None of the students had a family member that was a pilot. These pilots were recruited through friends of the experimenter and an e-mail sent around the University of Otago’s Geography department. All participants gave their informed consent.

**Apparatus**

An Acer Travel-Mate 270 laptop computer was used to present the flight scenarios to the participants. The scenarios were run using the Inquisit 1.30 program by Millisecond software.

**Materials**

*Flight Scenarios:* The participants were presented with 16 flight scenarios (see appendix E). Each flight scenario depicted a cross-country flight flown VFR in a Cessna 172. Each scenario was presented twice.

Each scenario comprised 14 factors, with 2 levels of each factor representing high and low risk. A full-factorial design with 14 factors would require 16,384 scenarios ($2^{14}$). To keep the number of scenarios at a reasonable quantity, a $1/1024$ fractional factorial design was used, requiring 16 scenarios. The information within the scenarios needed to be consistent with the information that the pilots use to make weather-related decisions in real
life. Therefore, while only 14 factors were varied, each scenario included other information, which remained constant throughout the scenarios.

As in study one, the combination of levels used in each scenario was chosen using a fractional factorial design. A full factorial design with 16 scenarios would allow for four factors with two levels of each factor (Montgomery, 2000). The 16 scenarios were created by generating a full factorial design with four factors. Each level was represented by a positive or negative sign. All 4 factors were represented by a letter, A through to D. Therefore, each of the four scenarios was made up of a unique combination of the levels of the factors. The 10 new factors (E - O) were aliased with either a two-way, three-way, or four-way interaction of the first four factors. Each factor was represented by the following aliases: E = A*B*C, F = A*B*D, G = A*C*D, H = B*C*D, J = A*B*C*D, K = A*B, L = A*C, M = A*D, N = B*C, and O = B*D. Figure 3.6 shows an example of a scenario.

Figure 3.6. An example of a scenario used in study two.
Each flight was dated the 22\textsuperscript{nd} of April, depicting an autumn flight, and the weather information was valid for 2pm. The participants were provided with weather-related information similar to what they would seek while planning a real flight. Subsequently, each scenario conveyed four pieces of information: the route, weather at departure, en-route weather, and weather at arrival.

The route was either from Taieri to Timaru; a coastal, non-mountainous flight; or from Taieri to Queenstown, a non-coastal, mountainous flight. These two flights were chosen to represent different terrain, and subsequently, different levels of risk. The mountainous flight represented a higher level of risk than the non-mountainous flight.

The weather at departure was presented as a picture of what the weather looked like at Taieri. Taieri Aerodrome is a small aerodrome, 15 minutes south of Dunedin, and is where the Otago Aero Club operates from. The weather at departure was shown as a screen shot from Microsoft Flight Simulator, 2004 and represented what a pilot would see when looking out of the cockpit window. Table 3.2 presents the information about the weather at Taieri.

Table 3.2.
Factors, Levels, and the Simulator Settings for the Weather at Taieri

<table>
<thead>
<tr>
<th>Factor</th>
<th>Levels</th>
<th>Simulator setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation</td>
<td>No Precipitation</td>
<td>No precipitation</td>
</tr>
<tr>
<td></td>
<td>Rain</td>
<td>Rain</td>
</tr>
<tr>
<td>Cloud Ceiling</td>
<td>Medium</td>
<td>4000 - 6000 feet</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>900 - 1000 feet</td>
</tr>
<tr>
<td>Cloud Cover</td>
<td>Scattered</td>
<td>4/8</td>
</tr>
<tr>
<td></td>
<td>Broken</td>
<td>7/8</td>
</tr>
<tr>
<td>Wind Strength</td>
<td>Strong</td>
<td>Strong Wind</td>
</tr>
<tr>
<td></td>
<td>Weak</td>
<td>Weak Wind</td>
</tr>
<tr>
<td>Wind direction</td>
<td>Non cross-wind</td>
<td>Non cross-wind</td>
</tr>
<tr>
<td></td>
<td>Cross-wind</td>
<td>Cross-wind</td>
</tr>
</tbody>
</table>

Microsoft Flight Simulator, 2004 enables the user to customise the weather and the location of the airport. A windsock represented the wind strength and direction. The wind strength was shown by the angle that the windsock made with the ground. An acute angle represented a weak wind, while a right angle represented a strong wind. The direction of the windsock represented the direction of the wind. A cross-wind was shown by the windsock pointing across the runway. Meanwhile, a non-cross-wind was represented by
the windsock pointing towards the observer. The picture also contained information that remained constant throughout each scenario. The visibility was set at 48km, which was an acceptable level. Stratus clouds were used, as this type of cloud can exist at low and medium levels, and are associated with precipitation (http://ww20JO.atmos.uiuc.edu/).

This picture was created by using the flight simulator settings. For example, scenario one depicted rain, high scattered cloud, and a weak cross-wind. To create this picture, the location “Taieri airport” was chosen, and the settings were changed to reflect this weather. This picture was then saved by taking a screen shot and saving it in Microsoft Paint, version 5.1. The flight simulator did not have a wind-sock near the runway for Taieri. Therefore, a wind-sock at the appropriate wind strength and direction was created at Dunedin airport, and a screen shot was taken. Then, the wind-sock was copied onto another document and increased in size. This was then copied and pasted into the picture of Taieri Aerodrome. I then used Microsoft Paint to match the colour of the sky or clouds, so it was not apparent that the windsock was photo-shopped. This process was repeated for all scenarios, so there were 16 pictures representing weather at Taieri.

The en-route weather was presented as an area forecast for the Plains when travelling to Timaru, and for Clyde when travelling to Queenstown. Both the Plains and Clyde represent one of the areas which the pilot flies over when flying this route. En-route information is usually presented in code. Figure 3.7 shows an example of an en-route forecast available to pilots.
Figure 3.7. An example of an en-route weather forecast. From http://www.ifis.airways.co.nz/

However, since non-pilot participants needed to understand this information, in the scenarios, the en-route weather was presented in plain English, rather than in code. An example of this format is presented below in Figure 3.8.

Area forecast for Plains
Valid for 22nd of April, 10am to 11pm
FORECAST WINDS ALOFT
3000' 10 knots at 315°
VISIBILITY
20 km reducing to 5000m in showers
CLOUD
Scattered at 900' and broken at 8000'

Table 3.3 presents the information included in the en-route weather forecast.
Table 3.3.
Factors and Levels of En-route Weather Presented in Scenarios

<table>
<thead>
<tr>
<th>Factors</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precipitation/Visibility</td>
<td>30 km visibility</td>
</tr>
<tr>
<td></td>
<td>20 km visibility, reducing to 5000m in showers</td>
</tr>
<tr>
<td>Cloud Ceiling</td>
<td>4000 feet</td>
</tr>
<tr>
<td></td>
<td>900 feet</td>
</tr>
<tr>
<td>Cloud Coverage</td>
<td>Scattered</td>
</tr>
<tr>
<td></td>
<td>Broken</td>
</tr>
</tbody>
</table>

The en-route weather forecast also included information usually provided in area forecasts. At 3000 feet, the winds were 10 knots at 315°. The cloud ceiling was broken at 8000 feet. This information was kept constant across each scenario.

Information about the weather at arrival is usually presented as a Meteorological Aerodrome Report (METAR) or a Terminal Aerodrome Forecast (TAF). A METAR presents a recent weather report at a specific aerodrome, and is provided hourly (Ryan, 1984). Meanwhile, a TAF presents a forecast for an 8km radius around the aerodrome. Usually, pilots would use both to gain an overall impression of the weather and what will happen in the next few hours. However, I did not have enough room to present both a METAR and a TAF. Since METARs present reported weather, and are more reliable when the flight is taken straight away, I chose to present a METAR. This choice was validated by participants who commented on the usefulness of METARs in the task. Figure 3.9 depicts how a METAR would be presented to a pilot.

DUNEDIN (NZDN):
METAR NZDN 080300Z 24004KT 4000 FEW080 BKN120 12/07 Q1003=

Figure 3.9. Example of a METAR. From http://www.ifis.airways.co.nz/

Again, this information was presented in plain English, so non-pilots could understand it. An example of a METAR presented in the scenarios is depicted in Figure 3.10.
**METAR NZTU** (current weather at Timaru)
Valid for 22nd of April, 2pm.

WIND
- 20 knots at 200°

VISIBILITY
- 25 km reducing to 5000m in showers

CLOUD
- Broken at 850' and broken at 8000'

TEMPERATURE
- 16°

DEW POINT
- 9°

QNH (Barometer)
- 1002

*Figure 3.10. An example of a METAR at the destination airport*

Information that was included in the METAR is presented in Table 3.4.

*Table 3.4. Factors and levels for weather information at the destination airport*

<table>
<thead>
<tr>
<th>Factors</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud Ceiling</td>
<td>5000 feet</td>
</tr>
<tr>
<td></td>
<td>850 feet</td>
</tr>
<tr>
<td>Cloud Cover</td>
<td>Scattered</td>
</tr>
<tr>
<td></td>
<td>Broken</td>
</tr>
<tr>
<td>Precipitation/visibility</td>
<td>40km visibility</td>
</tr>
<tr>
<td></td>
<td>25km visibility, reducing to 5000m in showers</td>
</tr>
<tr>
<td>Wind strength</td>
<td>2 knots</td>
</tr>
<tr>
<td></td>
<td>20 knots</td>
</tr>
<tr>
<td>Wind direction</td>
<td>Non-cross-wind</td>
</tr>
<tr>
<td></td>
<td>Cross-wind</td>
</tr>
</tbody>
</table>

There are two runways at both Timaru airport and Queenstown Airport; a main runway and a cross-runway. The main runway at Timaru airport is orientated 020°/200°, while the cross-runway is at a right angle to this, at 110°/290°. Meanwhile, the main runway at Queenstown airport is orientated 050°/230° and the cross-runway is orientated 140°/320°.

It is ideal for pilots to take-off and land into the wind. The most dangerous situation is to land or take-off into a strong cross-wind. If there was a cross-wind, a pilot would use an available cross-runway. So, I wanted one of the conditions to be a cross-wind, or at least a wind that would be difficult to land in. It would be ideal to make the cross-wind at a 90° angle from the main runway. However, the pilot could just land at the cross-runway.
Therefore, the difficult wind (cross-wind) was a wind that crosses both runways on an angle. At Timaru airport, the cross-wind and non-cross-winds were represented at an angle of 335° and 200°, respectively. Meanwhile, at Queenstown airport, the cross-wind and non-cross-winds were represented at an angle of 5° and 140°, respectively.

Again, other information is usually included in a METAR, such as air temperature, dew-point, and barometric pressure. Furthermore, more than one level of cloud ceiling is usually presented. Therefore, information was included that was not varied: broken cloud at 8000 feet, temperature and dew point were 16° and 9°, respectively, and barometric pressure was 1002. All of these variables were chosen to represent realistic weather for mid-autumn, and weather that could exist, given the weather in each scenario.

The realism of the scenarios was assessed by an experienced flight instructor, who agreed that all scenarios represented realistic weather situations.

Demographic and Aeronautical Practices Questionnaire: The pilots completed the same questionnaire as described in study one. However, because student pilots were also used, the option for the highest level of certification currently held was expanded to include "none".

Demographic Questionnaire: The geography students completed the same questionnaire that the non-pilots completed in study one.

Aeronautical Maps: Participants were given two colour aeronautical maps depicting both routes. These maps measured 41 by 29.5 cm, and included topographical information such as the height of the terrain. The maps were at the scale of 1:500,000.

Procedure

All participants were tested individually. Participants read the information sheet and signed the consent form. The participants were then sat in front of the computer and were verbally given the following instructions:

Thank you for agreeing to take part in this study. You will be presented with a number of scenarios depicting VFR general aviation flights. You are to assume that you are flying a Cessna 172. You are to imagine that you are planning a cross-country flight and will be given weather-information needed to plan the flight. For each scenario I will give you four pieces of information; 1. the route, which will either be from Taieri to
Timaru or from Taieri to Queenstown. 2. the weather at departure, which will be a picture of what the weather looks like at Taieri Airport, 3. en-route weather, which will be an area forecast for either Plains or Clyde, depending on the route, and 4. the weather at destination, which will be a METAR for either Timaru or Queenstown. For the weather at destination, I did not have enough room for both the TAF and the METAR, so assume that the TAF is roughly the same as the METAR. Your task will be to read each scenario and rate how risky the scenario appears on a scale of 0 to 100. 0 will indicate that you think that there is a very low level of risk and 100 will mean that you think that it is the riskiest situation imaginable. There is no right or wrong answer, and your responses will be treated in the strictest of confidence. This task is not timed, so take as long as you need.

The participants were also given written instructions on the computer:

You will be shown a number of scenarios depicting a VFR General Aviation flight. For each scenario you will be given 3 pieces of information: 1. Weather at departure, 2. En-route weather, and 3. Weather at your destination. Please consider each flight and rate how risky each flight appears on a scale of 0 to 100, where 0 = least risky and 100 = most risky. Press the return key to record your response and continue to the next page.

In addition to this, participants were given information about the pilot. The pilots were asked to assume that they were the pilot, while the geography students imagined that they were pilots. The geography students were also given definitions of General Aviation, VFR flight, TAFs, and METARs. If asked, the pilots were given the co-ordinates of the airports, although only one inquired about this. All participants were given the maps and were shown the departure and destination points.

The scenarios were presented one at a time with all of the information about each scenario presented on one screen. The 16 scenarios were presented in the same order for all of the participants; this order was then repeated for the final 16 scenarios. After reading and considering each scenario, the participants entered the risk rating using the numerical keyboard.

After completing the scenarios, the participants filled in the relevant demographic questionnaire. Following this, the pilots were thanked, debriefed, and given chocolate as thanks for participation.
Results

Data Inspection

Following Weiss and Shanteau (n.d.), the CWS scores were transformed using a square root transformation. Both raw and transformed data are presented, represented by the subscript “r” and “t”, respectively. All inferential tests are calculated using the transformed data. The distribution of the residuals was assessed. If the residuals were not normally distributed, the variables were transformed using a square root transformation and outliers more than 3SDs away from the mean were removed. If this did not render the data normally distributed, a non-parametric test was used, with outliers included. All statistical tests were evaluated against an alpha level of .05.

Effect sizes for significant and marginally significant results were calculated using Cohen’s (1988) $d$ and partial eta squared ($\eta^2$). Cohen’s $d$ is calculated as the difference between the two means, divided by the pooled standard deviation. According to Cohen (1988), effect sizes of 0.2, 0.5, and 0.8 are regarded as small, medium, and large effects, respectively. Cohen’s $d$ also provides an estimate of the percentile standing of the experimental group, and the percent of non-overlap between the two distributions. Cohen’s $d$ was calculated when comparing two groups. Partial eta squared was calculated when an ANOVA was used. This represents the level of variance of the dependent variable predicted by the independent variable. According to Cohen (1988), .01 (1% of variance), .06 (6% of variance), and .14 (14% of variance) represent a small, medium, and large effect size, respectively.

CWS Analysis

Each participant gave two risk ratings for each scenario. Discrimination, inconsistency, and the CWS score were calculated in the same manner as in study one. The CWS scores across all participants ranged from 1.02 to 9.80 ($M = 3.80$, $SD = 1.98$). The CWS scores were adjusted using a square root transformation and averaged across the three groups of participants. There was no significant difference in the CWS score between the three groups of participants, $F(2,30) = 2.35$, $p = .113$ (Qualified Pilots: $M_r = 2.98$, $SD_r = 1.41$; $M_t = 1.68$, $SD_t = .40$; Student Pilots: $M_r = 3.66$; $SD_r = 1.88$; $M_t = 1.85$, $SD_t = .52$; Geography Students: $M_r = 4.77$, $SD_r = 2.28$; $M_t = 2.13$; $SD_t = .52$).

Discrimination values ranged from 44.79 to 1973.99 ($M = 595.90$, $SD = 429.78$). The level of discrimination was adjusted using a square root transformation, averaged, and
compared between the qualified pilots, student pilots, and the geography students. There was a significant difference between the groups of participants, $F(2,30) = 3.71, p = .036$, $\eta^2_p = .198$. Post Hoc tests with a Sidak correction were performed and showed that the qualified pilots ($M_r = 826.41, SD_r = 580.93, M_t = 27.28, SD_t = 9.50$) showed more discrimination than the geography students ($M_r = 384.38, SD_r = 284.96, M_t = 18.10, SD_t = 7.90$), mean difference, $t = 9.1815, SE_t = 3.63, p = .032, d = 1.05$. The average qualified pilots’ level of discrimination was in the 84th percentile of the mean of the geography students, with 55.4% non-overlap between the distributions. The raw mean discrimination values for each group are depicted in Figure 3.11.

Figure 3.11. Raw mean discrimination scores across the three groups of participants

Inconsistency values ranged from 9.38 to 718.75 ($M = 217.33, SD = 195.02$). The level of inconsistency was adjusted using a square root transformation, averaged, and compared across the three groups. There was a significant group effect on inconsistency, $F(2,30) = 4.12, p = .026, \eta^2_p = .216$. Post Hoc tests with a Sidak correction were conducted and it was found that there was a significant difference between the qualified pilots and the geography students. The qualified pilots were more inconsistent than the geography students (pilots, $M_r = 307.38, SD_r = 192.19, M_t = 16.66, SD_t = 5.72$; geography students, $M_t$
= 123.92, $SD_t = 166.11$, $M_t = 9.45$, $SD_t = 6.17$), mean difference, $t = 7.21$, $SE = 2.53$, $p = .023$, $d = 1.21$. The qualified pilots were in the 88th percentile of the geography student’s scores, and with 62.2% non-overlap in the scores. Figure 3.12 shows the raw mean inconsistency score across the three groups.

![Figure 3.12. Raw mean level of inconsistency across the three groups of participants.](image)

**Number of Cues Used to Make a Decision**

The number of cues used by participants to make risk perception judgements was calculated by totalling the number of cues that were significant for each participant. That is, a cue would be tallied if significantly different risk ratings were given, depending on which level of the factor was presented. One geography student, two trained pilots, and one student pilot used at least one cue to make their risk rating during the first replication. Meanwhile, two geography students, two trained pilots and one student pilot used at least one cue to make their risk rating during the second replication. Participants were divided into two groups according to whether they used any cues in either of the replications. There was a marginally significant difference in the CWS score between the two groups of participants, $F(1,30) = 3.79$, $p = .061$, $\eta_p^2 = .112$, $d = 0.75$. The participants who had at least used one cue to make a decision had a marginally higher CWS score ($M_t = 5.24$, $SD_t = 2.76$, $M_t = 2.21$, $SD_t = .63$) than the participants who did not use any cues to make a decision ($M_t = 3.45$, $SD_t = 1.59$, $M_t = 1.80$, $SD_t = .45$). The participants who used cues had
CWS scores in the 76th percentile of the CWS of the participants who did not use cues, with 43% non-overlap.

**Average Risk Rating**

The average risk rating given by qualified pilots, student pilots, and geography students were compared. A repeated measures ANOVA was calculated with a Sidak correction. The average risk rating did not differ across the groups of participants (qualified pilots, $M = 44.87$, $SE = 5.51$; student pilots, $M = 51.73$, $SE = 5.51$; geography students, $M = 40.62$, $SE = 5.51$), $F(2,30) = 1.03$, $p = .367$. The participants' responses did not differ between the two replications, $F(1,30) = 3.173$, $p = .085$, $\eta^2 = .096$, although there was a non-significant trend for the ratings to be higher for the second replication ($M = 47.07$, $SD = 18.37$) compared to the first ($M = 44.42$, $SD = 19.19$). There was no group*replication interaction.

**Hazardous Event Scale**

The involvement in hazardous events was calculated in the same manner as study one. The 11 qualified pilots had been involved in an average of 7.73 hazardous events in the past 24-months ($range = 2-15$, $SD = 4.13$). The frequency of involvement in hazardous events did not significantly correlate with the CWS score, discrimination, or inconsistency. There was no relationship between the number of hazardous events the pilots were involved in and the average risk ratings given.

Of particular interest was the number of times the pilots had flown VFR-into-IMC. Out of the 15 pilots with cross-country experience, 6 had flown VFR-into-IMC at least once in the previous 24-months. One pilot had flown VFR-into-IMC at least four times. The pilots were divided into two groups according to whether they had flown VFR into IMC in the previous 24-months. There was no significant difference between these two groups of participants in their CWS scores.

**Usual Aeronautical Practices**

The qualified pilots rated the percentage of time they participated in certain behaviours. Usual aeronautical practices were calculated in the same manner as study one. This was averaged across all items for each participant to indicate how often the pilot was
involved in safe practices. On average, pilots indicated that they were involved in safe aeronautical behaviours 53.48% of the time (range = 31.11 – 75.56%, SD = 13.32). The involvement in safe practices did not significantly correlate with the CWS score.

**Personal Minimums**

Pilots were asked to indicate the minimum visibility and cloud ceiling they would accept before taking off on a local or cross-country flights. Over half of the qualified pilots would not take off if visibility was less than 5km on a local flight and 7km on a cross-country flight. Meanwhile, over half of the qualified pilots would not take off if cloud ceiling was less than 1,000 feet on a local flight and 1,500 feet on a cross-country flight. There was no correlation between personal minimums and the CWS score, discrimination, and inconsistency.

The student pilot who had only completed 5 hours did not complete the personal minimums section of the questionnaire. Over half of the student pilots would not take off if visibility was less than 5km on a local flight and 10km on a cross-country flight. Over half of the student pilots would not take off if the cloud ceiling was less than 2,000 feet for a local flight and 4,000 feet for a cross-country flight. There was no relationship between personal minimums and the CWS score, discrimination, and inconsistency. Personal minimums were compared for qualified pilots and student pilots, using a Mann-Whitney U test. There was a significant difference between the qualified pilots (local flight: \( M = 1,363.64, SD = 452.27 \), cross-country: \( M = 2,000, SD = 1,072.38 \)) and the student pilots (local flight: \( M = 2,300.00, SD = 918.94 \), cross-country: \( M = 3,700, SD = 1,059.35 \)) in the minimum level of cloud ceiling that they would accept for a local flight and a cross country flight (Local flight: \( Z(2.20) = -2.52, p = .013, d = 1.29 \); Cross-country flight: \( Z(2.20) = 2.95, p = .003, d = 1.59 \)). In both cases, the student pilots had stricter personal minimums than the qualified pilots. The mean student pilots’ cloud ceiling personal minimums for cross-country flights was in the 88th percentile of the trained pilots’ personal minimums, with 62.2% of non-overlap between the two distributions. The mean student pilots’ cloud ceiling personal minimums for local flights was in the 93.3rd percentile of the trained pilots’ personal minimums, with 70.7% non-overlap between the two distributions.
Forced Landing

Since this questionnaire concerned weather-related risk perception, it was of interest to compare the CWS scores of the pilots who had and had not made a forced landing due to bad weather. Of the 15 pilots with cross-country flying experience, 8 pilots had made a forced landing due to bad weather. The pilots with cross-country experience were divided into those who had and had not experienced a forced landing. One pilot with a CWS score more than 3 SD higher than the mean, was excluded from the analysis. Those who had made forced landings due to bad weather had a higher CWS score than the pilots who had not (had made a forced landing: $M_t = 3.45$, $SD_t = 1.38$, $M_i = 1.83$, $SD_i = .36$; had not made a forced landing, $M_r = 1.49$, $SD_r = .28$, $M_i = 1.22$, $SD_i = .12$), $F(1,14) = 8.00$, $p = .014$, $\eta^2_p = .56$, $d = 2.27$. The mean CWS score of the pilots who made a forced landing is in the 97.7th percentile of the CWS scores of the pilots who had not made a forced landing, with 81.1% non-overlap.

Pilot Experience

The flying experience of pilots was analysed to determine if experience and age were related to expertise in risk perception. The CWS was not significantly related to age, level of certification, or holding an instrument or instructor rating. Similarly, there was no relationship between the CWS and any measure of aeronautical experience, other than hours flown cross-country in the previous 90-days, $r(22) = -.404$, $p = .063$. This marginally significant result suggests that the more hours flown cross-country during the previous 90-days, the lower the CWS score.

Geography Student Demographics

The geography students' experience and interest in aviation were analysed to examine whether these factors were related to expertise in risk perception. The CWS score was not related to the age of the participant or their interest in aviation. Similarly, the CWS score did not significantly differ according to gender, experience flying flight simulators, or whether the participant had a family member or a friend who was a pilot.

Discussion

The CWS procedure, as applied to weather-related decision-making, measured the level of expertise in judging aeronautical weather-related risks. The higher the CWS score,
the more consistently the participants discriminated between different weather-related scenarios. The qualified pilots have more experience making weather-related decisions than the student pilots and the geography students. Therefore, I expected that the qualified pilots would be best at perceiving weather-related aeronautical risks. However, the CWS scores did not significantly differ between the three groups of participants. While the qualified pilots were most discriminating, the geography students were significantly more consistent. There are three possible explanations for this result: the task may not measure aeronautical risk perception, it measures risk perception but risk perception may not be related to experience in WRDM, or the task may measure memory ability. Each of these will be discussed in turn.

It is possible that this task is not related to risk perception. There was no relationship between the CWS scores and involvement in hazardous events. Risk perception should be related to involvement in hazardous events (Hunter, 2002, 2005, 2006). Hunter had participants rate the risk involved in aeronautical scenarios, and found that these ratings were related to involvement in hazardous events, where the more risk perceived in the hypothetical scenarios, the fewer hazardous events the participants were involved in. The lack of a relationship between the CWS score and involvement in hazardous events found in the present study may have occurred because the CWS score does not measure risk perception. However, this lack of relationship may have occurred because of the way the HES and CWS scores were measured. The HES measured involvement in hazardous aeronautical incidents in general, such as nearly running out of fuel and making poor decisions. Meanwhile, the CWS measured risk perception in weather-related scenarios. According to Ajzen and Fishbein (1977), it is difficult to find a relationship between an attitude and a behaviour (e.g., risk judgments and involvement in incidents) when they are measured at different levels of specificity. Therefore, it is of interest to measure the relationship between expertise in weather-related risk perception and involvement in weather-related hazardous events.

One of the questions addressed in the HES is whether the pilot had flown VFR-into-IMC over the previous 24-months. Out of the 15 pilots with cross-country flying experience, six had flown VFR-into-IMC, and one had at least four times, supporting past findings that a substantial proportion of pilots fly VFR-into-IMC (Hunter, 1995; O'Hare & Chalmers, 1999). I expected that the pilots who flew VFR-into-IMC would have a lower CWS score than the pilots who had not flown VFR-into-IMC over the past 24-months.
This over-reliance on memory may explain why the geography students were more consistent, but no: more discriminating, as discrimination relies less on memory.

One way to decrease the task’s reliance on memory is to prevent the transfer of information from Short Term Memory (STM) to Long Term Memory (LTM). There are many theories and models of STM, but all describe STM as having a limited storage capacity and a limited duration (Baddeley, 1999). Baddeley and Hitch (1974) developed a model to describe how information is transferred from STM to LTM. They proposed that working memory is used to hold and relate multiple pieces of information, essential for tasks such as reading and calculating mathematics equations. The model is presented below in Figure 3.12.

![Figure 3.12. Baddeley and Hitch's (1974) Working Memory Model (from Baddeley, 1999).](image)

Baddeley and Hitch proposed that there are three components of working memory: the visuo-spatial sketchpad, phonological loop, and central executive. The role of the central executive is to control the phonological loop and sketchpad and relate the information to LTM. The role of the visuo-spatial sketchpad is to hold and manipulate visual information, and is used in tasks such as mental rotation. The role of the phonological loop is to hold and manipulate verbal information. Information is transferred to LTM using rehearsal. Baddeley (1999) describes three lines of evidence to support the existence of the phonological loop.
Baddeley (1966) found that participants had more difficulty recalling words that sounded similar (e.g., man & can) than words that sounded dissimilar (e.g., pit & few). In contrast, there was no effect of semantic similarity on the recall of words. Participants recalled words that were semantically similar just as easily as words which were semantically dissimilar. This suggests that individuals’ transfer information from working memory to LTM via rehearsal.

Presenting an irrelevant spoken word interfered with the participants’ ability to recall a visually presented word (Salame & Baddeley, 1982). Participants were presented with meaningful (e.g., COW) or nonsense words (e.g., CAW), or silence, while being presented with words to recall. Despite being told to ignore the irrelevant information, the participants recalled fewer words when the competing words were presented, compared to when there was no distraction. This suggests that the distraction is occurring at the phonological level, rather than the semantic level.

Finally, there is a link between word length and memory span. Baddeley, Thomson, and Buchanan (1975) found that participants had better recall for shorter compared to longer words. This suggests that the participants sounded each word while rehearsing, as longer words take more time to rehearse.

The existence of the phonological loop suggests that preventing phonological rehearsal should interfere with recall. If the participants are asked to remember items which can be verbalised (e.g., letters), while simultaneously engaged in a competing verbal task (e.g., counting), phonological rehearsal of the to-be-remembered items will be suppressed (Baddeley, 1999). This suppression of phonological rehearsal will reduce the participants’ ability to recall information. Peterson and Peterson (1959) presented participants with stimuli comprising three letters and three numbers (e.g., ABC 309). The participants counted backwards in multiples of three for a period of time (3 – 18 seconds), after which they recalled the three letters. The participants had difficulty recalling the letters after counting backwards for 18 seconds. Muter (1980) followed a similar procedure, except participants were rarely asked to recall the letters. Forgetting occurred after two seconds of distraction. Both of these studies suggest that preventing phonological rehearsal (through a counting task) can reduce participants’ ability to recall information.

One of the criticisms of the first study was that the participants were largely airline pilots, and that this may account for the lack relationship between the CWS score and experience. This was rectified in the second study, as none of the qualified pilots had an
ATPL. However, there was no relationship between their flight experience and expertise in aeronautical weather-related risk perception. Since task-related experience should be related to developing expertise (Wiggins et al., 1996; Wiggins et al., 2002), there should be a positive relationship between recent cross-country experience and the CWS score, as pilots gain most weather-related experience flying cross-country. However, there was a marginally significant negative relationship between the CWS score and recent cross-country experience. The pilots with the most recent hours flying cross-country tended to be worse at this task. This suggests that the lack of relationship between flight experience and expertise in risk perception cannot be attributed to the type of pilot used in the first study.

The participants in study two used even fewer cues when making the risk judgments, than the participants did in study one. The three groups of participants in study two did not differ in the number of cues used. This supports both study one and past research (e.g., Ettenson et al., 1987; Phelps & Shanteau, 1978; Slovic, 1967), and suggests that experts do not use more cues than novices when making judgments. However, like study one, the present study used a fractional factorial design, making it difficult to measure the effect of cues on the participants’ judgements.

While Hunter (2002, 2005, 2006) found a relationship between risk perception and involvement in hazardous events using the HES, there was little relationship between risk ratings and scores on the HES in the present study. As mentioned previously, this may be due to the level of measurement, with the CWS measuring weather-related risk perception and the HES measuring general aviation incidents. However, this does not explain why the results of the first study did not support Hunter’s research.

Risk tolerance relates to the participants’ willingness to accept risk. Therefore, personal minimums are one way of measuring risk tolerance. The pilots indicated the lowest level of visibility and cloud ceiling they would accept before going on a local and cross-country flight. The student pilots were stricter in their personal minimums for cloud ceiling than the qualified pilots. This was expected, as the qualified pilots have more experience, and can cope with lower cloud ceilings compared to the student pilots.

There was no relationship between expertise in weather-related aeronautical risk perception and flight experience. To improve this task, I have designed an experiment to prevent the participants from rehearsing this information. Participants will be presented with the same 16 flight scenarios used in study two and will be asked to rate the risk in the situation on a scale of 0 to 100. After the presentation of each scenario, the participants
will be presented with three-digit numbers (e.g., 891) and will be asked to count backwards aloud in multiples of three for 15-seconds. This will prevent the participants from rehearsing the information contained in the scenario, and subsequently transferring the information to LTM. Therefore, participants will have to intuitively rate the risk involved in the scenario, without relying on their memory of rating from the previous presentation.

A third study was conducted, to assess whether the difference in the CWS between the geography students and the pilots was due to the reliance on working memory. The scenarios were identical to study two, with the addition of a blocking task (counting backwards) between each scenario. Due to difficulties in obtaining geography students, this study was conducted as a pilot study.

**Study 3**

**Method**

**Participants**

Qualified pilots and geography students were recruited for the pilot study. The four male pilots were aged between 19 and 35 years ($M = 26, SD = 8.19$). These pilots had an average of 1,066 total hours flying time ($range = 140 - 3,800, SD = 1,822.62$) with between 60 and 3,550 hours as pilot-in-command ($M = 941.25, SD = 1,739.20$). These pilots had completed an average of 48.25 hours as pilot-in-command over the previous 90 days ($range = 12 - 70, SD = 25.12$). Between 50 and 3,200 hours were spent as pilot-in-command on a cross-country flight ($M = 843.25, SD = 1,571.19$), with a mean of 47.75 of these hours in the previous 90 days ($range = 15 - 70, SD = 23.39$). Three pilots held a PPL and one held a CPL. This certification was held an average of 1.79 years ($range = 0.33 - 6, SD = 2.81$). None of the pilots held an instrument or instructor rating.

There were two male and one female Geography students from the University of Otago. These students ranged in age from 22 to 35 ($M = 29, SD = 6.56$). Two were PhD candidates, and the other was an honours student. All participants had completed a 300-level climatology paper. One student had flown a plane, although this was only an introductory flight, and two had flown a computer-based flight simulator. None of the students had a family member that was a pilot. These pilots were recruited through friends of the experimenter and an e-mail sent around the University of Otago Geography department. All participants gave their informed consent.
Apparatus

An Acer Travel-Mate 270 laptop computer was used to present the flight scenarios to the participants. The scenarios were run using Inquisit 1.3 program by Millisecond software.

Materials

Flight scenarios: Each participant was presented with the same 16 flight scenarios that were used in the second study, with the addition of a blocking task. This task was similar to Peterson and Peterson (1959) and Muter (1980). In between each scenario, the participants were presented with a randomly selected 3-digit number and were instructed to count backwards in multiples of three. The number was presented in the centre of the screen in size 72 font.

Demographic and Aeronautical Practices Questionnaire: The pilots completed the same questionnaire as described in Study One.

Demographic Questionnaire: The geography students completed the same questionnaire that the non-pilots completed in Study One.

Aeronautical Maps: The routes depicted in the scenarios used in the present experiment were identical to those used in Study Two, Taieri to Queenstown, and Taieri to Timaru. Participants were given the same aeronautical maps that the participants received in Study Two.

Procedure

All participants were tested individually. Participants read the information sheet and signed the consent form. The participants were then sat in front of the computer and were given the same verbal and written instructions that participants received in Study Two. Participants were also given written instructions regarding the blocking task:
When a 3-digit number appears on the screen, say the number aloud, and count backwards in multiples of three as quickly and accurately as you can. For example, if the number is “391”, you should say “three ninety-one, three eighty-eight, three eighty-five, etc”, until something else appears on the screen.

They were also given the following verbal instructions:

After each flight scenario you will be presented with a 3-digit number. Your task is to say this number aloud and count backwards aloud in multiples of three. Do this until the next scenario appears.

In addition to this, participants were given information about the pilot. The pilots were asked to assume that they were the pilot in the scenario, while the geography students imagined that they were pilots. The geography students were also given definitions of GA, VFR flight, TAFs, and METARs. If asked, the pilots were given the co-ordinates of the airports. All participants were given the maps and were shown the departure and destination points.

The scenarios were presented one at a time, with all of the information presented on one page. The 16 scenarios were presented in the same order as Study Two, and this order was consistent across participants. This order was then repeated for the final 16 scenarios. After reading and considering each scenario, the participants entered their risk rating on the numerical keyboard. After entering this information, participants were presented with a 3-digit number. The participants counted backwards from this number in multiples of three. This was continued for 15 seconds, after which the next scenario appeared.

After the completion of the scenarios, the participants’ completed the relevant demographic questionnaire. Following this, the pilots were thanked, debriefed, and given chocolate as thanks for participation.

Results

There were 3 geography students and 4 trained pilots. The CWS scores for the qualified pilots and geography students are presented in Table 3.5.
Table 3.5.

Raw CWS Scores for the Pilots and Geography Students.

<table>
<thead>
<tr>
<th></th>
<th>Pilots</th>
<th>Geography Students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.17</td>
<td>3.36</td>
</tr>
<tr>
<td></td>
<td>7.57</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td>9.78</td>
<td>3.79</td>
</tr>
<tr>
<td></td>
<td>3.14</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>5.41</td>
<td>3.77</td>
</tr>
<tr>
<td>SD</td>
<td>3.96</td>
<td>0.39</td>
</tr>
</tbody>
</table>

The CWS scores ranged from 1.18 to 9.78 ($M = 4.91$, $SD = 2.94$). The average CWS score for the pilots was higher than the CWS scores of the geography students. Due to the small sample size, inferential statistics were not calculated. Cohen’s $d$ was 0.59, indicating that the mean pilot CWS score was in the 69th percentile of the geography students’ scores, with 33% non-overlap in the distribution. The power of the test was low, at 20.6%. With the beta error level set at 0.2 (power level at 0.8), a sample size of 36 for both samples would be needed for this to be a significant difference.

The qualified pilots in the present study had a higher mean CWS score compared to the qualified pilots in Study Two. However, this did not reach statistical significance, $t(3.41) = -1.04$, $p = .366$. Cohen’s $d$ was 0.82, indicating that the mean CWS score for the pilots in the present study was in the 79th percentile of the CWS scores for the pilots in study two, with 47.4% non-overlap between the two distributions. The power of the test...
was 32.8%. With a beta level set at 20% (80% power), a sample size of 19 for both groups would be needed for this to be a significant difference.

The geography students in the present study had a lower mean CWS score compared to the geography students used in study two. However, this did not reach statistical significance, $t(12) = .603, p = .558$. Cohen’s $d$ was 0.61, indicating that the mean CWS score for the geography students in study two was in the 73rd percentile of the CWS scores for the geography students in the present study, with 38.2% non-overlap between the two distributions. The power of the test was very low, at 11.4%. With a beta level set at 20% (80%) power, a sample size of 33 for both groups would be needed for this to be a significant difference.

**General Discussion**

Pilots may make poor decisions, such as flying into adverse weather because they lack awareness of the risks involved in doing so. In Studies One and Two, I used the CWS procedure to measure expertise in aeronautical risk perception in general (Study One) and weather-related risk perception (Study Two). Performance in this task should be related to experience in ADM. To test this, pilots’ performance in the task was compared to undergraduates (Study One); geography students, and student pilots (Study Two). Since the comparison participants lack experience in ADM, they should show inferior performance compared to the pilots.

The results of the first study supported the hypothesis, as the pilots were better at the task than the non-pilots. However, the results of the second study did not support the hypothesis. There was no relationship between expertise in weather-related aeronautical risk perception and WRDM experience. There are two possible explanations for the disparate results. It is possible that risk perception in general (Study One) is related to experience in ADM, but weather-related risk perception (Study Two) is not related to experience in WRDM.

Alternatively, the disparate results may be explained by the type of control participants used. The participants in the first study were undergraduate students, while the participants in the second study were post-graduate students. As mentioned previously, the CWS procedure may rely on memory performance, and this may account for the lack of relationship between the three groups of pilots in the second study. This was assessed in the third study, where geography students and qualified pilots completed the CWS task.
from Study Two. However, a blocking task was included, designed to reduce the effect of memory on performance. Since very few geography students and pilots completed the study, no firm conclusions can be made. However, the pilots tended to have improved in the new version of the CWS, while the geography students’ performance tended to worsen. When the blocking task was included, the pilots tended to be better at making weather-related aeronautical risk perception judgements compared to the geography students.

In these studies, I sought to validate the CWS as a measure of expertise in aeronautical risk perception, by comparing expert and novice performance. Another method of validation is to compare performance on this task to aeronautical decision-making. As part of my fifth study, qualified pilots will complete the same CWS procedure as described in the third study, as well as a simulated flight into adverse weather. If flying into adverse weather is related to risk perception, the pilots who fly into adverse weather should show inferior performance on the CWS task compared to the pilots who do not fly into adverse weather.

To further assess the role of working memory on performance on the CWS task, the performance on the CWS task of participants in my second and fifth studies will be compared. If the inferior performance of the pilots in the second study was due to the role of memory on the task, the participants in Study Five should perform better than the participants in Study Two. The blocking task should improve performance as it will encourage pilots to make an intuitive judgement. That is, the pilots should make a risk-rating without consideration of what was rated in the previous scenarios. Also, I will assess the relationship between flight experience and performance on the CWS task. In Study Two, I found little relationship between flight experience and the CWS scores. If by including the blocking task, the CWS task becomes a more valid measure of expertise in aeronautical risk perception, there should be a greater relationship between flight experience and CWS scores, compared to the second study. I should also find a relationship between expertise in risk perception and involvement in hazardous aeronautical events, using Hunter’s (1995) HES.

The three studies described in this chapter examined expertise in aeronautical risk perception. Together with risk perception, risk tolerance is a key component of risk management (Deery, 1999). In the next chapter, I will describe a study where I will measure risk tolerance in GA pilots, using Lopes’s (1987) theory of risk tolerance.
Chapter 4: Measuring Risk Tolerance

Individuals may take risks because their acceptable level of risk is set too high. That is, they are too risk tolerant (Hunter, 2002). For example, a pilot who recognises that the weather is deteriorating yet continues flying shows a higher degree of risk tolerance than a pilot who diverts (O'Hare & Wiegmann, 2003).

There are always gains and losses associated with taking risks. According to Lopes (1987), an individual’s propensity to take risks depends on whether he or she is motivated by the opportunity for gain or the threat of loss. For example, an individual may be motivated to skydive or bungee jump because of the opportunity to experience an adrenaline rush. Meanwhile, another individual may choose not to skydive or bungee jump because of the threat of being seriously injured. The former individual is more risk tolerant than the latter.

This theory is supported by Hanoch et al. (2006) who measured risk taking in people known for risk-taking behaviour (e.g., sky divers & gamblers) and risk avoidance (e.g., health-seekers, such as gym members). The researchers measured domain specific risk-taking (e.g., whether the participants smoked) and perceived benefits and risks for each of the activities. They found that risk taking was domain specific. For example, the smokers reported smoking but not any other risk-taking behaviour. The participant’s risk-taking behaviour was related to perceived benefits. For example, the smokers reported higher perceived benefits for smoking than non-smokers. The perceived benefits of behaviour were better predictors of risk-taking behaviour than the perceived risk involved. This illustrates that risk takers are motivated by the benefits or the opportunity to be gained from taking risks.

The aim of the present study will be to use Lopes’s model to measure risk tolerance in GA pilots. This will be achieved by using a policy-capturing methodology to assess whether pilots are influenced more by opportunities or threats when making risk-taking decisions. In policy-capturing research, participants are presented with scenarios and are asked to make a judgment or an evaluation. Cues are systematically varied within each scenario. A separate regression equation is calculated for each participant and the within-subject regression coefficients indicate the influence of each cue on the response and the direction of the relationship (Tommasi, Williams, & Nordstrom, 1998).

Policy capturing has successively been used to understand how judges make decisions in many different areas, including Human Resources (e.g., Hemingway & Conte,
2003; Klein, Berman, & Dickson, 2000; Tommasi et al., 1998) and ADM (Driskill et al., 1995; Flathers et al., 1982; Hunter et al., 2003; Knecht et al., 2004). For example, Hemingway and Conte (2003) used the policy capturing methodology to assess employees' opinion of the fairness of lay-off policies. Participants were asked to imagine that their company was preparing to lay-off employees. They were presented with 100 pen-and-paper scenarios, with 12 factors being varied in each scenario, with two or three levels of each factor. Participants rated the scenarios according to how fair they perceived each lay-off. Hemingway and Conte calculated a regression equation for each participant, and the within-subject regression coefficients indicated the influence of each factor on the perception of fairness. Such research is important in understanding how certain factors influence how people feel and behave.

The present study was conducted in two phases. To ensure that the levels of each factor did represent varying degrees of opportunity and threat, four flight instructors were presented with different situations and asked to rank the level of opportunity or threat in each situation from one (high) to six (low). The flight instructors' ratings were used to develop scenarios for the second phase. During the second phase the level of opportunity and threat were varied, and qualified pilots rated the likelihood of going on each flight.

**Phase One: Scenario Development**

**Method**

**Participants**

The participants were four flight instructors from a local flight training school. The one female and three male pilots were aged between 21 and 31 ($M = 26.5$, $SD = 4.80$). The pilots had a mean of 927.5 total hours flying time ($range = 340 – 2100$, $SD = 805.29$) with between 210 and 1900 hours as pilot in command ($M = 762.5$, $SD = 782.11$). In the last 90 days the pilots had completed an average of 62.5 hours pilot-in-command ($range = 40 – 110$, $SD = 32.02$). Between 50 and 1500 hours were spent as pilot-in-command on cross-country flights ($M = 460$, $SD = 696.23$), with a mean of 41.25 hours cross-country in the last 90 days ($range = 10 – 100$, $SD = 41.31$). Three of the pilots held a CPL, while one pilot held an ATPL. This certification was held an average of 1.57 year ($range = .25 – 3$ years, $SD = 1.20$).
Stimuli

The participants were presented with an 11-page booklet (see appendix F). Participants were presented with lists of 6 situations and were asked to rank them according to the level of opportunity or threat represented. These rankings were used to select levels of opportunity and threat used in phase two of the study. The participants were presented with eight lists, four lists regarding opportunity and four regarding threats.

There are many different benefits that pilots can gain from flying. These include social approval, excitement, income, and career advancement. This framework was used to create the four lists, so that each list represented a different way of operationalising the opportunity gained from flying.

Social approval was represented by 6 different scenarios which differed in the amount of social status gained: gaining hours towards a CPL, taking a film director (Peter Jackson) on a flight to look at film locations, taking an afternoon off work to go on a flight, taking a friend on a flight, transporting a critically injured adult, and transporting an adult with non-life threatening injuries. These items were chosen to differ in the potential amount of social approval. For example, a situation representing a low level of potential social approval was as follows:

It is a Friday afternoon and you decide to take some time off work and go on a flight.

Meanwhile, a situation representing a high level of potential social approval was as follows:

You are a medical rescue pilot and get a phone call to transport a critically injured adult male.

Excitement was operationalised by the level of variety offered by the flights. The participants were presented with 6 different flights: circuits at home airfield, local flight around home town, local flight around another town, cross-country flight that the pilot had been on many times, cross-country flight to a destination that the pilot had driven to, and a cross-country flight to a destination that the pilot had never been to. Therefore, the most opportunity for excitement would be a flight that they had never been on, to a location that they had never been to. To keep the level of threat consistent between all scenarios, the participants were told that another pilot was present on the flight, and that the other pilot
had flown all of the routes before. For example, a situation representing a low level of potential excitement was as follows:

Circuits at your home airfield.

Meanwhile, a situation representing a high level of potential excitement was as follows:

A cross-country flight. You have never flown this route before, although your passenger has. You have never been to this destination.

Income was operationalised by the number of passengers. Pilots were asked to imagine that they were an owner of a tourist flight operating company. In this set of scenarios participants were presented with different numbers of passengers (1, 2, 3, 5, 6, and 7 passengers) which represented the amount of money they could earn. The participants were also told how much money each participant paid ($250) and the operating costs ($400), which enabled the calculation of profit. Therefore, one and seven passengers represented a loss of $150 and a gain of $1,350, respectively.

Career advancement was represented by the type of aircraft that would be flown. In this set of scenarios the participants were to imagine that they usually fly a Cessna 172 and there were given the chance to fly 6 different types of aircraft; a Cessna 172, Piper Arrow, Piper Seneca, Cessna 152, Cessna 182, and Piper Warrior. Each of these aircraft were characterised by the number of engines (single or twin), type of propeller (fixed or constant-speed), and type of landing gear (fixed or retractable). Flying more advanced aircraft would represent more opportunity for career advancement. Advanced aircraft represent a greater threat to flight safety (e.g., AOPA, 2006). Therefore, the pilots may perceive this as a manipulation of threat, rather than opportunity. To keep the level of threat constant, pilots were told that a flight instructor qualified for that type was present on each of the flights. For example, a situation representing a low level of potential career advancement was as follows:

Cessna 152. This single-engine aircraft has a fixed pitch propeller and a fixed undercarriage.

Meanwhile, a situation representing a high level of potential career advancement was as follows:
Piper Seneca. This twin-engine aircraft has constant-speed propellers and a retractable landing gear.

According to Jensen (1996), threats in aviation can be grouped into four categories; pilot, aircraft, environment, and external pressures. This framework was used to operationalise threat.

Threats from within the pilot were represented by the amount of alcohol the pilot had consumed in the last couple of days; the amount of sleep received the previous night, health, and the existence of personal problems. For example, the situation representing the highest threat was as follows:

The pilot had been drinking heavily the night before and only had 5 hours sleep the previous night. She was on medication for a cold and had personal problems.

Meanwhile, the situation representing the lowest level of threat was as follows:

The pilot had not drunk any alcohol for the last week and had 8 hours sleep the previous night. She felt healthy and had no personal problems.

For the other four situations, the levels of threat were systematically changed.

Threats from within the aircraft were represented by the length since the last annual or 100-hour inspection and the degree of the magneto drop. Magneto drop is the difference in the RPM between the two ignition systems. A large difference represents a potential problem. The situation representing the highest threat was:

The aircraft had its last inspection 11 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of more than 200 rpm.

The situation representing the lowest threat was as follows:

The aircraft had its last inspection 2 months ago (annual or 100-hour). The run-up check showed a magneto drop of 50 rpm.

For the other four situations, the levels of threat were systematically changed so that if the situation became more threatening, it became longer since the last inspection and there was a larger magneto drop.
Threats within the environment were operationalised by the en-route weather. Participants were given an en-route weather forecast which depicted the level of visibility, precipitation, cloud height and cover, and the wind strength at 3,000 feet. The situation with the highest level of threat was as follows:

Isolated thunderstorms and rain with 1000m visibility. The cloud is overcast at 800 feet AMSL. The winds aloft at 3,000 feet are 35 knots.

The situation with the lowest level of threat was:

80 km visibility with no precipitation. The skies are clear and the winds aloft at 3,000 feet are 5 knots.

For the other four situations, the level of threat was systematically changed so if the situation became more threatening, the visibility became poorer, the cloud cover greater, the cloud ceiling lower, and the winds stronger. In three out of the six situations there was precipitation (thunderstorms or showers).

External pressure was represented by the time the pilots had to do prepare for the flight (5, 15, 30, 45, 60, or 90 minutes). This factor was chosen as time pressure is an important threat and the less time the pilot has to do a pre-flight check, the greater the likelihood that the pilot has missed something awry with the aircraft (AOPA, 2003). Furthermore, the pilot is not likely to be able to complete a proper flight plan.

Procedure

Participants were tested as a group, although there was no collaboration. Participants read the information sheet, and signed the consent form. The participants were then given the booklet and asked to read the instruction page. This read:

On the next four pages you will be presented with a number of different situations and will be asked to rank the situations from 1 (high threat or opportunity) to 6 (low threat or opportunity). Use each number from 1-6 only once.

The participants were then given an example of how to rank items. The participants were then given a definition of opportunity:

On the next four pages you will be presented with sets of situations, each situation varying in the amount of
opportunity. Opportunity refers to the upside, or the possibilities of gain, of a particular situation. Any situation may vary in the likelihood of gain, and gains may include such things as excitement, interest, and career advancement, admiration of others, financial gain, and self-satisfaction.

The participants were then presented with the four lists and asked to rank the opportunity in each situation from one (most opportunity) to six (least opportunity). The participants were encouraged to ask any questions and leave comments about the lists. The pilots were not informed what type of opportunity each list represented.

After ranking the four lists the participants were given a definition of threat:

On the next 4 pages you will be presented with sets of situations, each situation varying in the amount of threat involved. You should rank the amount of threat from most to least threat. Threat refers to the downside, or the possibility of loss, of a particular situation. Any particular situation might contain a variety of threats, such as financial loss, danger, injury or property damage, loss of self-esteem, and disapproval of others. While all situations contain some probability of loss, some have a higher likelihood of loss, and therefore threat, than others.

The participants were presented with the four lists and asked to rank the threat in each situation, from one (high threat) to six (low threat). Again, the participants were encouraged to ask questions and leave any comments.

The participants then completed the demographic information questionnaire. Following this, the participants were thanked, debriefed, and given chocolate as thanks for their participation.

Results

Opportunity

The following four tables present the rankings for each facet of opportunity, where one and six represent high and low opportunity, respectively. The highest, median, lowest, and mode rankings for the opportunity for social approval are shown in Table 4.1.
### Table 4.1.

*Rankings for the Amount of Social Approval (Social)*

<table>
<thead>
<tr>
<th>Item</th>
<th>Highest Rank</th>
<th>Median Rank</th>
<th>Lowest Rank</th>
<th>Mode Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPL Training</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Peter Jackson</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>After work</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Friend</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Critically Injured Patient</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Non-critically Injured Patient</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

High, medium, and low opportunity were best represented by transporting a critically injured passenger (mode rank = 1), taking a friend on a flight (mode rank = 4), and taking a Friday afternoon off work to go on a flight (mode rank = 6), respectively. Taking a friend on a flight was chosen rather than taking Peter Jackson, as the Peter Jackson scenario contained references to money. Furthermore, it would be conceivable for some people to value this higher than a medical rescue. The highest, median, lowest, and mode rankings for the opportunity for excitement are shown in Table 4.2.

### Table 4.2.

*Rankings for the Level of Excitement (Excitement)*

<table>
<thead>
<tr>
<th>Item</th>
<th>Highest Rank</th>
<th>Median Rank</th>
<th>Lowest Rank</th>
<th>Mode Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuits</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Local Flight Home</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Local Flight Other</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Cross-country Usual</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Cross-country Driven</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>New Cross-country</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

High opportunity was best represented by a cross-country flight that the pilot had never been on, to a destination that the pilot had never been to (mode rank = 1). Medium
opportunity was best represented by a cross-country flight that the pilot had never been on, to a destination that the pilot had driven to (mode rank = 2). Low opportunity was best represented by a cross-country flight that the pilot had been on many times (mode rank = 4). These scenarios were chosen over the circuits and local flights because each flight was a cross-country flight, in an attempt to keep the threat constant. All flight instructors agreed that the new cross-country flight represented more opportunity than the cross-country flight where the pilot had driven to the location, which was ranked higher than the cross-country flight that the pilot had been on many times before. The highest, median, lowest, and mode rankings for the opportunity to gain income are shown in Table 4.3.

Table 4.3.

<table>
<thead>
<tr>
<th>Item</th>
<th>Highest Rank</th>
<th>Median Rank</th>
<th>Lowest Rank</th>
<th>Mode Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 passenger</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>2 passengers</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>3 passengers</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5 passengers</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>6 passengers</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7 passengers</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Five passengers (mode rank = 3) were chosen to best represent high opportunity as it is the most passengers that a Cessna 206 can safely carry, and the business would earn $850. A Cessna 206 has seven seats, however, it is industry practice to only carry six people (five passengers and the pilot). Two passengers (mode rank = 5) were chosen to best represent low opportunity as no business would take-off with one passenger as they would lose money (-$150). Two passengers were chosen to represent a situation where the business would make a small profit ($100) but not a loss. Three passengers were chosen to best represent medium opportunity (mode rank = 4), where the business would earn $350. The highest, median, lowest, and mode rankings for the opportunity for career advancement are presented in Table 4.4.
Table 4.4.

*Rankings for the Type of Aircraft (Career)*

<table>
<thead>
<tr>
<th>Item</th>
<th>Highest Rank</th>
<th>Median Rank</th>
<th>Lowest Rank</th>
<th>Mode Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cessna 172</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Piper Arrow</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Piper Seneca</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Cessna 152</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Cessna 182</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Piper Warrior</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

The three aircraft chosen were all Pipers, to ensure that all scenarios included aircraft that the hypothetical pilots were not used to flying, and only differed according to the performance of the aircraft. High opportunity was best represented by a Piper Seneca, a twin engine aircraft with a constant speed propeller and a retractable landing gear (mode rank = 1). Medium opportunity was best represented by a Piper Arrow, a single-engine aircraft with a constant-speed propeller and a retractable landing gear (mode rank = 2). Low opportunity was best represented by a Piper Warrior, a single-engine aircraft with a fixed-speed propeller and a fixed landing gear (mode rank = 5). Out of the Pipers, all flight instructors agreed that the Piper Seneca, Piper Arrow, and Piper Warrior represented high, medium, and low opportunity, respectively.

**Threat**

The following four tables present the rankings for each facet of threat, where one and six represent high and low threat, respectively. The highest, median, lowest, and mode rankings for the threat within the pilot are presented in Table 4.5.
Table 4.5.

**Rankings for the Threat Within the Pilot (Pilot)**

<table>
<thead>
<tr>
<th>Item</th>
<th>Highest Rank</th>
<th>Median Rank</th>
<th>Lowest Rank</th>
<th>Mode Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>No alcohol in last couple of days, 8 hours sleep, medication for cold, personal problems.</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>No alcohol in the last week, 8 hours sleep, onset of cold, personal problems.</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Moderately drinking the night before, 6 hours sleep, medication for cold, personal problems.</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>No alcohol in the last week, 8 hours sleep, healthy but personal problems</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Drinking heavily the night before, 5 hours sleep, healthy, personal problems.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>No alcohol in the last week, 8 hours sleep, healthy, no personal problems.</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

The highest level of threat was best represented by the pilot having 5 hours of sleep in the last 24-hours, drinking heavily the night before, on medication for a cold, and some personal problems (mode rank = 1). The medium level of threat was best represented by the pilot having 8 hours of sleep the previous night, having not drunk any alcohol in the last couple of days, on medication for a cold, and some personal problems (mode rank = 3). The lowest level of threat was best represented by the pilot having had 8 hours of sleep the previous night, having not drunk any alcohol in the last week, and feeling healthy with no personal problems (mode rank = 6). The highest, median, lowest, and mode rankings for the threat within the aircraft are presented in Table 4.6.
### Table 4.6.

**Rankings for the Threat Within the Aircraft (Aircraft)**

<table>
<thead>
<tr>
<th>Item</th>
<th>Highest Rank</th>
<th>Median Rank</th>
<th>Lowest Rank</th>
<th>Mode Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 months, magneto drop 200 rpm</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9 months, magneto drop 70 rpm</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>9 months, magneto drop 150 rpm</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>11 months, magneto drop 150 rpm</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2 months, magneto drop 50 rpm</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>9 months, magneto drop 100 rpm</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Threat within the aircraft was represented by the length of time since the last annual or 100-hour inspection and the level of the magneto drop. In New Zealand, aircraft must be inspected after 100-hours or every year. High threat was best represented by an aircraft which had its last inspection 11-months ago and a magneto drop of 200 rpm (mode rank = 1). Medium threat was best represented by an aircraft with its last inspection 9-months ago and a magneto drop of 100 rpm (mode rank = 4). Meanwhile, low threat was best represented by an aircraft which had its inspection 2-months ago and a magneto drop of 50 rpm (mode rank = 6). The highest, median, lowest, and mode rankings for the threat of en-route weather are presented in Table 4.7.
Table 4.7.

*Rankings for the Threat of En-Route Weather (Environment)*

<table>
<thead>
<tr>
<th>Item</th>
<th>Highest Rank</th>
<th>Median Rank</th>
<th>Lowest Rank</th>
<th>Mode Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 km visibility with showers. Broken cloud at 2000 feet AMSL, 15 knot winds</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>80 km visibility, no precipitation. Clear skies, 5 knot wind</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>30 km visibility, no precipitation. Scattered cloud at 5000 feet, 10 knot wind</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10 km visibility, showers. Overcast at 1000 feet, 20 knot wind</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>40 km visibility, no precipitation. Few clouds at 10000 feet, 5 knot wind</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Isolated thunderstorm, rain. 1000 km visibility, overcast at 800 feet, 35 knot winds</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

High threat was best represented by isolated thunderstorms and rain, with 1000 metres visibility, overcast cloud at 800 feet, and 35 knot winds (mode rank = 1). Medium threat was best represented by 30 km visibility, no precipitation, scattered clouds at 5000 feet, and 10 knot winds (mode rank = 4). Low threat was best represented by 80 km visibility with no precipitation, clear skies, and a 5 knot wind (mode rank = 6). The highest, median, lowest, and mode rankings for the threat of time pressure are presented in Table 4.8.
Table 4.8.

Rankings for the Threat of Time Pressure (External)

<table>
<thead>
<tr>
<th>Item</th>
<th>Highest Rank</th>
<th>Median Rank</th>
<th>Lowest Rank</th>
<th>Mode Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 minutes</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15 minutes</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>30 minutes</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>45 minutes</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>1 hour</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>1.5 hour</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

High, medium, and low threat were best represented by 5 minutes (mode rank = 1), 30 minutes (mode rank = 3), and 1 hour (mode rank = 5). One hour was chosen rather than 1.5 hours, as the pilots indicated that a pilot who showed up 1.5 hours early would be a hindrance.

During the second phase, pilots will be presented with pen-and-paper scenarios differing in the level of opportunity and threat, and will rate the likelihood of going on the flight. If a pilot is more influenced by threat, he or she should go on a flight when the threat is low, regardless of what there is to gain. Meanwhile, if the pilot is influenced more by the opportunity, he or she should be more likely to go on a flight when the opportunity is high, regardless of what there is to lose. The influence of opportunity and threat on each pilot’s response will indicate whether the pilot is risk averse or tolerant.

Since risk management is an important aspect of pilot decision-making, I will assess the relationship between risk tolerance and past aeronautical risk-taking. Aeronautical risk-taking will be measured using Hunter’s (1995) HES, which assesses the number of aircraft incidents that the pilot had been involved in during the last 24-months. Hunter (1995) devised this questionnaire as a proxy measure for accident involvement. The number of previous incidents which the pilots have been involved in predicts the likelihood of being involved in an accident in the future (Hunter, 2001). I expect that the risk tolerant pilots will have been involved in more hazardous events than the risk averse pilots.
Phase Two: Policy Capturing

Method

Participants

The participants were 24 male and 3 female pilots, ranging in age between 21 and 54 years ($M = 29.12, SD = 9.55$). The pilots had a mean of 1,224.99 hours total flying time ($range = 125 - 11,500, SD = 2,252.75$) with between 45 and 10,000 hours as pilot-in-command ($M = 674.92, SD = 2,008.13$). During the last 90 days the pilots had completed a mean of 60.74 hours as pilot-in-command ($range = 7 - 150, SD = 42.40$). Between 22 and 7,500 hours had been spent as pilot-in-command on cross-country flights ($M = 674.92, SD = 1,492.03$) and the pilots had spent an average of 39.74 hours flying cross-country in the last 90 days ($range = 0 - 120, SD = 41.47$). A PPL and a CPL was the highest level of certification held by 5 and 22 participants, respectively. The pilots had held this certification for a mean of 4.5 years ($range = .20 - 26, SD = 5.54$). Out of the 27 pilots, 10 held an instructor and an instrument rating, 4 held an instructor, but not an instrument rating, 4 held an instrument, but not an instructor, rating, and 9 pilots held neither rating.

Stimuli

Participants were given a booklet containing 36 scenarios depicting GA flights (see appendix G). It would have been possible to have a completely crossed design, and to present 16 types of scenarios to the pilots. This would create 144 scenarios ($16*9$ scenarios per group of scenarios), which was deemed too many scenarios for the pilots to complete. Therefore, scenarios were created by combining one facet of threat with one facet of opportunity, creating four types of scenarios. The facets of threat and opportunity represented in each of the four types of scenarios are depicted in Table 4.9 below.

<table>
<thead>
<tr>
<th>Type of Opportunity</th>
<th>(is combined with)</th>
<th>Type of Threat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Approval</td>
<td>Pilot</td>
<td></td>
</tr>
<tr>
<td>Excitement</td>
<td>Aircraft</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>Weather</td>
<td></td>
</tr>
<tr>
<td>Career Advancement</td>
<td>Time pressure</td>
<td></td>
</tr>
</tbody>
</table>
Based on the rankings given by the pilots in phase one, three levels of each facet of opportunity and threat were presented in the scenarios, representing high, medium, and low opportunity and threat. Therefore, there were four sets of scenarios, with nine scenarios in each set. Each scenario contained a description of the situation, which included the level of opportunity and threat. Unlike the first phase, the pilots were asked to imagine that they were the pilots in the scenarios. Below the description was a Likert scale ranging from 1 to 6, with an anchor at each end. The anchor above 1 was “definitely no”, while the anchor above 6 was “definitely yes”. Under the Likert scale was a space to include any comments.

An example of a scenario depicting the opportunity for social approval and threat within the pilot is shown below.

| Imagine that is a Friday afternoon and you decide to take some time off work to go on a flight. |
| You had been drinking heavily the night before and had only 5 hours sleep. You are on medication for a cold and have relationship problems. |

This scenario depicts low opportunity to gain social approval and high threat within the pilot. During phase one a flight instructor commented that “personal problems” was too ambiguous, so this was changed to “relationship problems”.

An example of a scenario depicting the opportunity for excitement and threat within the pilot is shown below.

| Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have never flown this route before, although your passenger has. You have driven to this destination before. |
| You will fly a Cessna 172 which had its last inspection 2 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 50 rpm. |

This scenario depicts medium opportunity for excitement and low threat of aircraft problems. The passenger was always depicted as an experienced pilot who had flown the route before. This was to decrease the threat associated with flying a new route, so any differences between the routes were due to the opportunity, rather than the threat involved.
An example of a scenario depicting the opportunity to earn income and the threat of en-route weather is shown below.

Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs.

Five tourists come into your business; if you take these passengers you will earn $850. Assume that you will be flying a Cessna 206.

The en-route weather forecast is as follows:

80 km visibility with no precipitation. The skies are clear and winds aloft at 3000 feet are 5 knots.

This scenario represents high opportunity to gain income and low threat of en-route weather. Participants were asked to imagine that they were the owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. This route was chosen as it was flown over flat terrain. Participants were told that each passenger pays $250 for the flight and that each flight would cost $400 in operating costs. This was to highlight that there was an operating cost involved and that each successive passenger would lead to an increase in money earned. Participants were given information about the number of passengers, and then were explicitly told how much money would be made in each situation. The profit made was calculated using the following equation:

\[(\text{Number of passengers} \times \$250) - \$400\]

The participants were then told that they would be flying a Cessna 206. The same aircraft was used despite the number of passengers to control for the different dangers involved in flying different aircraft.

An example of a scenario depicting the opportunity to for career advancement and the threat of time pressure is shown below.

Imagine that you are planning on going on a cross-country flight in a Piper Seneca (twin-engine aircraft, constant speed propeller, and a retractable landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight.

You have 30 minutes to complete the pre-flight check and the flight planning.
This scenario represents high opportunity for career gain and a medium threat of time pressure. The pilots were asked to imagine that they were going on a cross-country flight. Since there is a degree of risk associated with flying high performance aircraft, especially if the pilot has not flown the aircraft before, the participants were told that a flight instructor was present on the flight. The flight instructor was qualified on the type of plane depicted in the scenario.

Procedure

Participants were tested individually or in small groups. Participants read the information sheet and signed the consent form. The participants were then given the 36-page booklet containing the scenarios and were given the following instructions:

This booklet contains 36 scenarios depicting a General Aviation flight. You are to assume that you are the pilot in the scenario. Please read each scenario and rate the likelihood that you will go on the flight on a scale from 1 to 6, 1 indicating that you would definitely not go on the flight and 6 indicating that you definitely would. There may be additional information that you need to make your decision; please only consider the information depicted in each scenario. There is a space [at the bottom of the page] where you can write any comments. Please feel free to ask me any questions during the study and to take a break if you need to.

After completing the scenarios participants were given the same demographic information questionnaire as used in Studies One and Two. Participants were also reminded that the I-IES referred to incidents that occurred in the past 24 months. This was to ensure that the experienced pilots would not have a higher number of incidents simply because they were more experienced, and consequently had greater exposure to potentially hazardous situations. The participants were then thanked, debriefed, and given chocolate as a thank you for their participation.

Results

Data Inspection

Any outliers which were more than 3 SDs from the mean were excluded from data analysis. Any data which were not normally distributed were transformed. If after the transformation, the data were not normally distributed, a non-parametric statistical test was used, including any outliers. All statistical tests were evaluated against an alpha level of .05.
Effect sizes for significant and marginally significant results were calculated using Cohen's (1988) $d$ and partial eta squared ($\eta_p^2$). Cohen's $d$ is calculated as the difference between the two means, divided by the pooled standard deviation. According to Cohen (1988), effect sizes of 0.2, 0.5, and 0.8 are regarded as small, medium, and large effects, respectively. Cohen's $d$ also provides an estimate of the percentile standing of the experimental group, and the percent of non-overlap between the two distributions. Cohen's $d$ was calculated when comparing two groups. Partial eta squared was calculated when an ANOVA was used. Partial eta squared represents the level of variance of the dependent variable predicted by the independent variable. According to Cohen (1988), .01 (1% of variance), .06 (6% of variance), and .14 (14% of variance) represent a small, medium, and large effect size, respectively.

**Multiple Regressions**

The levels of each factor (opportunity and threat) were coded as high, medium, and low. High, medium, and low values of threat and opportunity were given dummy variables of 3, 2, and 1, respectively. Each participant rated the likelihood of going on each flight, from one (low) to 6 (high). A multiple regression equation was calculated for each participant, assessing the influence of threat and opportunity on the likelihood of going on the flights. The within-subject regression coefficients are presented in Table 4.10. This shows the influence of threat and opportunity on the likelihood of going on the flight and the $R^2$. The $R^2$ measures the level of variance predicted by the model. If the $R^2$ is greater than .50 the policy has been captured (Fritzsche, Finkelstein, & Penner, 2000).
Table 4.10.
Within-Subject Regression Coefficients for Overall Threat and Opportunity

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.81**</td>
<td>.05</td>
<td>.68</td>
</tr>
<tr>
<td>2</td>
<td>-.91**</td>
<td>-.08</td>
<td>.83</td>
</tr>
<tr>
<td>3</td>
<td>-.75**</td>
<td>-.05</td>
<td>.56</td>
</tr>
<tr>
<td>4</td>
<td>-.70**</td>
<td>.00</td>
<td>.49</td>
</tr>
<tr>
<td>5</td>
<td>-.87**</td>
<td>-.02</td>
<td>.75</td>
</tr>
<tr>
<td>6</td>
<td>-.78**</td>
<td>-.11</td>
<td>.62</td>
</tr>
<tr>
<td>7</td>
<td>-.89**</td>
<td>.00</td>
<td>.64</td>
</tr>
<tr>
<td>8</td>
<td>-.67**</td>
<td>-.03</td>
<td>.46</td>
</tr>
<tr>
<td>9</td>
<td>-.67**</td>
<td>.04</td>
<td>.48</td>
</tr>
<tr>
<td>10</td>
<td>-.87**</td>
<td>-.02</td>
<td>.76</td>
</tr>
<tr>
<td>11</td>
<td>-.84**</td>
<td>.02</td>
<td>.70</td>
</tr>
<tr>
<td>12</td>
<td>-.81**</td>
<td>.10</td>
<td>.67</td>
</tr>
<tr>
<td>13</td>
<td>-.87**</td>
<td>-.02</td>
<td>.76</td>
</tr>
<tr>
<td>14</td>
<td>-.79**</td>
<td>-.06</td>
<td>.63</td>
</tr>
<tr>
<td>15</td>
<td>-.88**</td>
<td>.04</td>
<td>.77</td>
</tr>
<tr>
<td>16</td>
<td>-.86**</td>
<td>.03</td>
<td>.74</td>
</tr>
<tr>
<td>17</td>
<td>-.90**</td>
<td>.00</td>
<td>.81</td>
</tr>
<tr>
<td>18</td>
<td>-.81**</td>
<td>.02</td>
<td>.65</td>
</tr>
<tr>
<td>19</td>
<td>-.76**</td>
<td>-.03</td>
<td>.58</td>
</tr>
<tr>
<td>20</td>
<td>-.86**</td>
<td>-.03</td>
<td>.75</td>
</tr>
<tr>
<td>21</td>
<td>-.86**</td>
<td>-.02</td>
<td>.74</td>
</tr>
<tr>
<td>22</td>
<td>-.84**</td>
<td>-.06</td>
<td>.71</td>
</tr>
<tr>
<td>23</td>
<td>-.85**</td>
<td>.08</td>
<td>.73</td>
</tr>
<tr>
<td>24</td>
<td>-.85**</td>
<td>.00</td>
<td>.72</td>
</tr>
<tr>
<td>25</td>
<td>-.74**</td>
<td>.12</td>
<td>.56</td>
</tr>
<tr>
<td>26</td>
<td>-.84**</td>
<td>.01</td>
<td>.71</td>
</tr>
<tr>
<td>27</td>
<td>-.84**</td>
<td>-.05</td>
<td>.71</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.

The within-subject regression coefficients indicate the relationship between response and the independent variables. In all cases the pilots' responses were significantly influenced by the level of threat in the situation. These within-subject regression coefficients ranged from -.99 to -.67, indicating that the higher the threat in the situation, the less likely the pilot was to go on the flight. Meanwhile, in all cases the pilots' responses were not significantly influenced by opportunity. The within-subjects regression coefficients ranged from -.11 to .12, indicating that there was little relationship between the level of opportunity and the decision to go on the flight. A positive relationship would indicate that the higher the opportunity, the more likely the participant would go on the
flight. The $R^2$ values ranged from .83 to .46, with three participants with $R^2$ less than .50. These pilots were excluded from subsequent analyses.

Pilots indicated how often they had experienced 12 potentially hazardous aeronautical situations in the past 24-months, on a scale from 0 to 4 or more. Involvement in hazardous events was calculated in the same manner as studies one and two. The participants were involved in an average of 4.88 hazardous events in the last 24 months ($range = 2 - 12, SD = 2.21$). The number of hazardous events experienced did not significantly correlate with the influence of threat or opportunity on the decision to go on the flight. There was no relationship between experience and age of the participant and the influence of threat or opportunity on the decision to take-off.

Separate multiple regression equations were calculated for each participant, for each way of operationalising opportunity and threat. The within-subject regression coefficients for when the opportunity was for social approval and the threat was within the pilots are displayed in Table 4.11.
Table 4.11.
Within-Subject Regression Coefficients for Pilot-Related Threat and Opportunity for Social Approval.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.95**</td>
<td>.00</td>
<td>.89</td>
</tr>
<tr>
<td>2</td>
<td>-.94**</td>
<td>-.19</td>
<td>.92</td>
</tr>
<tr>
<td>3</td>
<td>-.91**</td>
<td>-.06</td>
<td>.84</td>
</tr>
<tr>
<td>4</td>
<td>-.99**</td>
<td>.00</td>
<td>.98</td>
</tr>
<tr>
<td>5</td>
<td>-.96**</td>
<td>.06</td>
<td>.92</td>
</tr>
<tr>
<td>6</td>
<td>-.82*</td>
<td>-.15</td>
<td>.69</td>
</tr>
<tr>
<td>7</td>
<td>-.88**</td>
<td>.18</td>
<td>.81</td>
</tr>
<tr>
<td>8</td>
<td>-.90**</td>
<td>-.21</td>
<td>.87</td>
</tr>
<tr>
<td>9</td>
<td>-.97**</td>
<td>-.07</td>
<td>.95</td>
</tr>
<tr>
<td>10</td>
<td>-.92**</td>
<td>-.07</td>
<td>.84</td>
</tr>
<tr>
<td>11</td>
<td>-.86**</td>
<td>-.06</td>
<td>.74</td>
</tr>
<tr>
<td>12</td>
<td>-.67*</td>
<td>.40</td>
<td>.60</td>
</tr>
<tr>
<td>13</td>
<td>-.96**</td>
<td>.09</td>
<td>.93</td>
</tr>
<tr>
<td>14</td>
<td>-.91**</td>
<td>.07</td>
<td>.83</td>
</tr>
<tr>
<td>15</td>
<td>-.89**</td>
<td>.12</td>
<td>.81</td>
</tr>
<tr>
<td>16</td>
<td>-.95**</td>
<td>.00</td>
<td>.89</td>
</tr>
<tr>
<td>17</td>
<td>-.95**</td>
<td>.00</td>
<td>.89</td>
</tr>
<tr>
<td>18</td>
<td>-.87**</td>
<td>0.00</td>
<td>.75</td>
</tr>
<tr>
<td>19</td>
<td>-.89**</td>
<td>.12</td>
<td>.81</td>
</tr>
<tr>
<td>20</td>
<td>-.97**</td>
<td>.07</td>
<td>.95</td>
</tr>
<tr>
<td>21</td>
<td>-.87**</td>
<td>0.00</td>
<td>.75</td>
</tr>
<tr>
<td>22</td>
<td>-.86**</td>
<td>-.07</td>
<td>.74</td>
</tr>
<tr>
<td>23</td>
<td>-.95**</td>
<td>0.00</td>
<td>.89</td>
</tr>
<tr>
<td>24</td>
<td>-.87**</td>
<td>0.00</td>
<td>.75</td>
</tr>
<tr>
<td>25</td>
<td>-.88**</td>
<td>.13</td>
<td>.79</td>
</tr>
<tr>
<td>26</td>
<td>-.87**</td>
<td>0.00</td>
<td>.75</td>
</tr>
<tr>
<td>27</td>
<td>-.87**</td>
<td>0.00</td>
<td>.75</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.

In all cases the pilots' responses were significantly influenced by the level of threat in the situation. These within-subject regression coefficients ranged from -.99 to -.67, indicating that the more pilot-related threat in the situation, the less likely that the pilot would go on the flight. Meanwhile, in all cases the pilots' responses were not significantly influenced by opportunity. The within-subject regression coefficients ranged from -.21 to .40, indicating that there was little relationship between the level of social approval and the decision to go on the flight. The R² values for all participants were greater than 0.5, indicating that the model captured the decision-making policies. Subsequently, all participants were included in the following analyses.
There was no relationship between the number of hazardous events the participants were involved in and the influence of threat and opportunity on the decision to take-off. There was no relationship between the influence of threat on the decision to take-off and any measure of experience. However, there was a positive relationship with the total number of hours ($r_s(27) = .55, p < .003$), number of hours pilot-in-command ($r_s(27) = .55, p < .003$), hours pilot-in-command on cross-country flights ($r_s(27) = .51, p < .003$), PIC in the last 90 days ($r_s(27) = .40, p = .038$), and length held ($r_s(27) = .59, p = .002$), and the influence of opportunity on the decision to take-off. This positive relationship indicates that the more experienced pilots were more influenced by the level of opportunity in the scenarios.

The within-subject regression coefficients for scenarios with opportunity for excitement and aircraft related threat are displayed in Table 4.12.
Table 4.12.
Within-Subject Regression Coefficients for Aircraft-Related Threat and Opportunity to Gain Excitement.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.96**</td>
<td>-.06</td>
<td>.92</td>
</tr>
<tr>
<td>2</td>
<td>-.91**</td>
<td>.06</td>
<td>.84</td>
</tr>
<tr>
<td>3</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>4</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>5</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>6</td>
<td>-.83*</td>
<td>-.15</td>
<td>.70</td>
</tr>
<tr>
<td>7</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>8</td>
<td>-.84**</td>
<td>.07</td>
<td>.71</td>
</tr>
<tr>
<td>9</td>
<td>-.67</td>
<td>.17</td>
<td>.47</td>
</tr>
<tr>
<td>10</td>
<td>-.95**</td>
<td>.14</td>
<td>.92</td>
</tr>
<tr>
<td>11</td>
<td>-.88**</td>
<td>.20</td>
<td>.81</td>
</tr>
<tr>
<td>12</td>
<td>-.95**</td>
<td>-.14</td>
<td>.93</td>
</tr>
<tr>
<td>13</td>
<td>-.94**</td>
<td>.00</td>
<td>.88</td>
</tr>
<tr>
<td>14</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>15</td>
<td>-.89**</td>
<td>.06</td>
<td>.79</td>
</tr>
<tr>
<td>16</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>17</td>
<td>-.96**</td>
<td>-.07</td>
<td>.92</td>
</tr>
<tr>
<td>18</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>19</td>
<td>-.90**</td>
<td>-.16</td>
<td>.84</td>
</tr>
<tr>
<td>20</td>
<td>-.91**</td>
<td>-.06</td>
<td>.84</td>
</tr>
<tr>
<td>21</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>22</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>23</td>
<td>-.92**</td>
<td>.08</td>
<td>.86</td>
</tr>
<tr>
<td>24</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>25</td>
<td>-.76**</td>
<td>-.44</td>
<td>.77</td>
</tr>
<tr>
<td>26</td>
<td>-.89**</td>
<td>-.06</td>
<td>.79</td>
</tr>
<tr>
<td>27</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.

In all but one case the pilots’ responses were significantly influenced by the level of threat in the situation. These within-subject regression coefficients ranged from -.96 to -.67, indicating the more aircraft-related threat in the situation, the less likely the pilot would go on the flight. Meanwhile, in all cases the pilots’ responses were not significantly influenced by opportunity. The within-subject regression coefficients ranged from -.44 to .20, indicating that there was little relationship between the level of excitement and the decision to go on the flight. One pilot had a $R^2$ value less than 0.5, and this pilot was excluded from subsequent analyses.
There was no relationship between the number of hazardous events experienced and the influence of threat or opportunity on the decision to take-off. There was no relationship between the influence of threat on the decision to take off and any indices of experience. There was no relationship between experience and the influence of opportunity on the decision to take off. However, there was a relationship between age and influence of opportunity on the decision to take-off, \( r, (24) = -.616, p = .001 \), indicating the younger the pilot, the more likely he or she was to take-off if the opportunity was high.

The within-subject regression coefficients for the scenarios with weather-related threat and the opportunity to earn income are shown in Table 4.13.

Table 4.13.

*Within-Subject Regression Coefficients for Weather-Related Threat and the Opportunity to Earn Income.*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.44</td>
<td>.66*</td>
<td>.62</td>
</tr>
<tr>
<td>2</td>
<td>-.95**</td>
<td>.00</td>
<td>.89</td>
</tr>
<tr>
<td>3</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>4</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>5</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>6</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>7</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>8</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>9</td>
<td>-.88**</td>
<td>.06</td>
<td>.78</td>
</tr>
<tr>
<td>10</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>11</td>
<td>-.86**</td>
<td>.00</td>
<td>.74</td>
</tr>
<tr>
<td>12</td>
<td>-.73*</td>
<td>.22</td>
<td>.58</td>
</tr>
<tr>
<td>13</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>14</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>15</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>16</td>
<td>-.80*</td>
<td>.19</td>
<td>.67</td>
</tr>
<tr>
<td>17</td>
<td>-.85**</td>
<td>.12</td>
<td>.74</td>
</tr>
<tr>
<td>18</td>
<td>-.81**</td>
<td>.34</td>
<td>.78</td>
</tr>
<tr>
<td>19</td>
<td>-.81**</td>
<td>.27</td>
<td>.74</td>
</tr>
<tr>
<td>20</td>
<td>-.87**</td>
<td>.00</td>
<td>.74</td>
</tr>
<tr>
<td>21</td>
<td>-.87**</td>
<td>.13</td>
<td>.77</td>
</tr>
<tr>
<td>22</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>23</td>
<td>-.74*</td>
<td>.37</td>
<td>.68</td>
</tr>
<tr>
<td>24</td>
<td>-.95**</td>
<td>.00</td>
<td>.89</td>
</tr>
<tr>
<td>25</td>
<td>-.59*</td>
<td>.59*</td>
<td>.70</td>
</tr>
<tr>
<td>26</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>27</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
</tbody>
</table>

*\( p < .05. **p < .01.\)
In all but one case the pilots' responses were significantly influenced by the level of threat in the situation. These within-subject regression coefficients ranged from -.95 to -.44, indicating that the pilots were less likely to go on the flight when the weather was adverse. Meanwhile, the within-subject regression coefficients ranged from .00 to .65. For two pilots, there was a significant relationship between the level of opportunity and the likelihood of going on the flight. For one participant, the within-subject regression coefficient for opportunity was .66 ($t = 2.60, p = .04$) and the within-subject regression coefficient for threat was -.44 ($t = -1.73, p = .13$). This suggests that when the amount of income earned for the flight was high, the pilot was more likely to go on the flight compared to when the income earned was low. Meanwhile, there was no relationship between the level of threat and the likelihood of going on the flight. For the second risk tolerant participant, the within-subject regression coefficient for both opportunity and threat was -.59 ($t = 2.65, p = .04$). This suggests that this pilot was influenced by both opportunity and threat equally. Following Lopes's terminology, these pilots were risk tolerant, as they were influenced by the level of opportunity in the situation.

The pilots were divided into those who were and were not risk tolerant, and the number of hazardous events which the pilots were involved in was compared. The risk tolerant pilots had been involved in significantly more hazardous events ($M = 8.00, SD = 1.41$) compared to the risk averse pilots ($M = 4.52, SD = 2.20$), $Z(27) = -2.08, p = .034, d = 1.88$. The average number of hazardous events which the risk tolerant pilots were involved in was in the 96.4th percentile of the risk averse pilots, with an 77.4% non-overlap between the two distributions. The mean number of incidents for the risk tolerant and risk averse pilots is shown in Figure 4.1.
There were no significant differences in the experience level between the two groups of pilots. However, there was a trend towards the risk tolerant pilots to have less total hours ($M = 201.50, SD = 26.16$) and hours pilot-in-command ($M = 88.5, SD = 44.55$) compared to the risk averse pilots (total hours: $M = 1302.86, SD = 2326.40$; pilot-in-command: $M = 1087.53, SD = 2072.14$), total hours: $Z(25) = -1.714, p = .091$, pilot-in-command: $Z(25) = -1.806, p = .068$.

There was no relationship between the influence of threat or opportunity on the decision to take-off and the number of hazardous events experienced or the experience of the participants.

The within-subject regression coefficients for scenarios with threat of time pressure and opportunity for career advancement are shown in Table 4.14.
## Table 4.14.

*Within-Subject Regression Coefficients for Threat of Time Pressure and the Opportunity for Career Advancement.*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.92**</td>
<td>-.12</td>
<td>.85</td>
</tr>
<tr>
<td>2</td>
<td>-.89**</td>
<td>-.22</td>
<td>.85</td>
</tr>
<tr>
<td>3</td>
<td>-.43</td>
<td>-.43</td>
<td>.38</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-.91**</td>
<td>-.20</td>
<td>.87</td>
</tr>
<tr>
<td>6</td>
<td>-.83**</td>
<td>-.33</td>
<td>.81</td>
</tr>
<tr>
<td>7</td>
<td>-.74*</td>
<td>-.22</td>
<td>.59</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-.78**</td>
<td>-.39</td>
<td>.75</td>
</tr>
<tr>
<td>10</td>
<td>-.93**</td>
<td>-.15</td>
<td>.88</td>
</tr>
<tr>
<td>11</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>12</td>
<td>-.97**</td>
<td>-.07</td>
<td>.94</td>
</tr>
<tr>
<td>13</td>
<td>-.96**</td>
<td>-.16</td>
<td>.94</td>
</tr>
<tr>
<td>14</td>
<td>-.70*</td>
<td>-.47</td>
<td>.72</td>
</tr>
<tr>
<td>15</td>
<td>-.99**</td>
<td>.00</td>
<td>.98</td>
</tr>
<tr>
<td>16</td>
<td>-.93**</td>
<td>-.07</td>
<td>.87</td>
</tr>
<tr>
<td>17</td>
<td>-.97**</td>
<td>-.08</td>
<td>.95</td>
</tr>
<tr>
<td>18</td>
<td>-.88**</td>
<td>-.32</td>
<td>.88</td>
</tr>
<tr>
<td>19</td>
<td>-.59</td>
<td>-.54</td>
<td>.64</td>
</tr>
<tr>
<td>20</td>
<td>-.88**</td>
<td>-.14</td>
<td>.80</td>
</tr>
<tr>
<td>21</td>
<td>-.91**</td>
<td>-.21</td>
<td>.87</td>
</tr>
<tr>
<td>22</td>
<td>-.88**</td>
<td>-.24</td>
<td>.83</td>
</tr>
<tr>
<td>23</td>
<td>-.89**</td>
<td>-.15</td>
<td>.82</td>
</tr>
<tr>
<td>24</td>
<td>-.92**</td>
<td>.00</td>
<td>.85</td>
</tr>
<tr>
<td>25</td>
<td>-.85**</td>
<td>-.07</td>
<td>.72</td>
</tr>
<tr>
<td>26</td>
<td>-.92**</td>
<td>.14</td>
<td>.86</td>
</tr>
<tr>
<td>27</td>
<td>-.78*</td>
<td>-.24</td>
<td>.68</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01.

Two participants were excluded, as they rated the likelihood of going all 9 flights as 6. In all but two cases the pilots’ responses were significantly influenced by the level of threat in the situation. These within-subject regression coefficients ranged from -.99 to -.43, indicating that the more time pressure, the less likely that the participant would go on the flight. Meanwhile, in all cases the pilots’ responses were not significantly influenced by the level of opportunity. The within-subject regression coefficients ranged from -.54 to .14, indicating that there was little difference between the opportunity for career advancement and the likelihood of going on the flight. One pilot had an $R^2$ of less than .50, and was excluded from the subsequent analysis.
There was no relationship between the influence of threat or opportunity on the decision to take-off and the number of hazardous events experienced or the age and experience of the participants.

**Number of Flights Agreed to Go On.**

One of the reasons I used a 6-point Likert scale was to force the participants to make a go/no-go decision. A rating of between 1 (definitely not) and 3 was coded as no-go. A rating of between 4 and 6 (definitely yes) was rated as a go. The number of flights that the participants indicated that they would go on was calculated, this was out of a maximum of 36 flights. The pilots indicated that they would go on a mean of 20.62 (57.28%) flights \( (range = 12 – 28, \ SD = 3.97) \). There was no relationship between the number of flights that the participants agreed to go on and involvement in hazardous events.

The mean likelihood that each pilot would go on the flights was also calculated, by averaging the rating given for each flight. The mean rating could range from 1, indicating that the pilot definitely would not go on any flight, to 6, indicating that the pilot definitely would go on all flights. The mean likelihood ranged from 2.75 to 4.75 \( (M = 3.88, \ SD = .53) \). There was no relationship between the mean likelihood of going on the flight and involvement in hazardous events.

**Personal Minimums**

Pilots were asked to indicate the minimum visibility and cloud ceiling they would accept before taking off on a local or cross-country flight. Over half of the pilots would not take off if visibility was less than 5km on a local flight and 10 km on a cross-country flight. Meanwhile, over half of the pilots would not take off if the cloud ceiling was less than 1,000 feet on a cross-country flight and 2,000 feet on a local flight. There was no correlation between personal minimums and the overall influence of opportunity or threat on the decision to go on the flight.

**Discussion**

Across all scenarios, all of the participants were influenced by the level of threat in the situation, and not the level of opportunity. Following Lopes's terminology, this indicates that the pilots were risk averse. The pilots were more willing to go on the flight if
the threat of loss was low, regardless of what there was to gain. This is a very positive finding, as it may indicate the success of flight training in New Zealand. For example, flight training emphasises the dangers of flying when under personal stress, under the influence of drugs and alcohol, and danger of fatigue, and all of these factors were shown to influence the likelihood that the pilots would take-off.

However, it is possible that the risk averse nature of the pilots reflected social desirability. That is, the pilots' responses may have been influenced by what they thought was an appropriate answer. This is always an issue when dealing with potentially sensitive issues, such as pilot decision-making. For example, the pilots would have been aware that they should not fly when isolated thunderstorms have been forecast, regardless of any potential gains from flying. This knowledge may have influenced the pilots' responses, while some pilots may be willing to take-off under these circumstances in real-life.

There was little relationship between the influence of threat and opportunity and the involvement in hazardous events. However, the lack of a relationship may be due to the low variation in within-subject regression coefficients for opportunity and threat. According to Karpinski and Hilton (2001), the chance of finding a significant correlation is decreased if there is little variation in one or both continuous variables.

Separate multiple regression equations were calculated to assess the effect of each facet of opportunity and threat on pilot decision-making. For most participants and across all facets of threat, there was a relationship between the level of threat and the likelihood of going on the flight. Therefore, the majority of participants were risk averse across all facets of threat.

The opportunity to earn income was the only facet of opportunity which influenced pilots' decision-making. Here two pilots showed risk tolerant tendencies, as the level of opportunity in the situation influenced their decision to take-off. These two pilots were more likely to take-off when the potential profit was larger. Therefore, the pilots were more willing to accept worsening weather if the reward was great enough.

The hypothesis was that the risk tolerant pilots will have had a history of risk-taking while flying, which was measured using Hunter's (1995) HES. When the pilots were divided into those who were and were not risk tolerant when the opportunity was income earned, the two risk tolerant pilots had been involved in more aviation incidents than the risk averse pilots. This relationship was not mediated by exposure to hazardous events. There were only two risk tolerant participants; therefore, I have to be cautious when
interpreting the evidence. However, the difference was substantial, with very little overlap between the distributions. With only 27 pilots, it is not surprising that only two were risk tolerant, especially as this is an unusual and potentially maladaptive behaviour. For example, one pilot agreed to go on the flight with the worst conditions when offered the highest financial reward. It is not surprising that very few pilots behaved in this manner.

This supports the hypothesis that risk tolerance will be related to involvement in hazardous events. The results also validate this as a measure of risk tolerance, as risk tolerance was related to risk-taking in a real-life context. This may be a better way of measuring risk tolerance than the gap acceptance methodology employed by Hunter (2002, 2005). Hunter found no relationship between risk tolerance and involvement in hazardous events. Furthermore, the results support Lopes's model of risk tolerance. Individuals who were motivated by the opportunity for gain were more likely to take risks. However, this was only found with opportunity to earn income.

High opportunity for career advancement, excitement, and social approval did not lead to an increased likelihood of going on the flight. The lack of relationship between the level of social approval and the likelihood of going on a flight supports research by O'Hare and Smitheram (1995). These researchers found that social approval did not influence the pilots' decision to continue flying into adverse weather in a simulated flight.

The lack of relationship between the degree of opportunity in the situation and the likelihood of going on the flight may be explained by pilots viewing some facets of opportunities as threats. For the scenarios with career advancement and excitement as opportunity, for many pilots there was a negative relationship (although not significant) between opportunity and the likelihood of going on the flight. When the opportunity was high, these pilots tended to be less likely to go on the flight, compared to when the opportunity was low. This suggests that the pilots treated the opportunity as a threat. For example, career advancement was represented by the type of aircraft that a hypothetical pilot was to fly. Participants imagined that they usually flew a Cessna 172 (single engine, fixed propeller, and fixed landing gear) and they were given the chance to fly 3 different types of aircraft; each characterized by the number of engines (single or twin), type of propeller (fixed or constant-speed), and type of landing gear (fixed or retractable), with the more advanced aircraft representing the most opportunity.

It is possible that the participants viewed flying more advanced aircraft as more threatening, as pilots are more likely to be involved in fatal accidents when flying a multi-
engine aircraft compared to when flying a single-engine aircraft (Li & Baker, 1999). This partially explains why many participants had negative within-subject regression coefficients for career advancement. If the pilots viewed opportunity as threat, the higher the opportunity, the less likely the participants would be to go on the flight. To make opportunity and threat orthogonal, the scenarios specified that a flight instructor qualified on each type was present on the flight. This was to ensure that the pilots would view flying in a more advanced aircraft as an opportunity, rather than a threat. However, it is possible that the pilots ignored this information.

Similarly, excitement was operationalised by the novelty of the flight, with more novel flights representing higher opportunity for excitement. However, it is again possible that the pilots viewed more novel flights as more threatening, as flying an unfamiliar route can be dangerous. This may explain why some pilots had negative within-subject regression coefficients for the opportunity for excitement. To make opportunity and threat orthogonal, the scenarios specified that the passenger was an experienced pilot, who had flown the route before. Again, it is possible that the pilots ignored this information. The interdependence of threat and opportunity, when the opportunity was career advancement and excitement, may explain the lack of influence of opportunity on pilots' decision-making.

The most obvious limitation with the present study is the low number of participants, especially as there are only two risk tolerant pilots. This severely limits the power and subsequently the likelihood of finding significant results. Furthermore, 27 pilots represent a very small proportion the pilot population, so the results may not generalise. However, as there are only approximately 4,000 fixed-wing PPL or CPL pilots with an active medical certificate in New Zealand (www.caa.govt.nz/), it is very difficult to get large samples. This study would be ideally suited to the internet, where large numbers of pilots could complete the scenarios and Hunter's (1995) HES.

This measure of risk tolerance is a unique and interesting way of studying risk tolerance, an area that has traditionally been difficult to measure. The present study has found some evidence that risk tolerance was related to risk-taking behaviour. The two risk tolerant pilots had been involved in more aviation incidents than the risk averse pilots. The next step in this research is to further validate the measure of risk tolerance, by looking at the relationship between risk tolerance and risk-taking during a simulated flight. As part of my fifth study, pilots will complete a simulated flight into adverse weather, as well as the
risk tolerance measure developed in the present study. This measure will be further validated if the pilots who continue flying into adverse weather will be more risk tolerant than those who turn back or divert.
Chapter 5: Implicit Associations and Time Pressure

 Implicit Associations

In the previous literature review chapters, I have described research suggesting that risk management is an important component of ADM. In this chapter, I will argue that how people (e.g., pilots) feel about hazards and objects (e.g., adverse weather) can influence decision-making and subsequent behaviour, and that these attitudes can exist on an implicit level.

Intuition and Risk: Attitudes and affect play an important role in decision-making, especially decision-making involving risk (Finucane, Alhakami, Slovic, & Johnson, 2000; Finucane, Peters, & Slovic, 2003). Activities which are high in risk can be high or low in benefit, whereas activities which are low in benefit are usually low in risk. Despite the often positive relationship between risk and benefit, people usually believe that the higher the benefit of an object or activity, the lower the risk (Alhakami & Slovic, 1994; Fischhoff, Slovic, Lichtenstein, Read, & Combs, 2000; McDaniels, Axelrod, & Slovic, 1995; Slovic, Kraus, Lappe, & Major, 1991). For example, people tend to believe that vaccines are high in benefit and low in risk. Finucane et al. (2000) argue that this perceived relationship is mediated by how the participant feels about the object or activity. That is, if an object is liked, he or she will see it as more beneficial and less risky than a disliked object.

Finucane et al. (2000) assessed the influence of affect, defined as the positive and negative feelings associated with an object (Finucane et al., 2003), on the perceived relationship between risk and benefit. Finucane et al. (2000) assessed whether manipulating the time available to make a risk judgement changed the relationship between perceived risk and benefit. People are likely to revert to their intuitive judgment when under time pressure, rather than making a cognitive judgement (Maule & Svenson, 1993). This is due to increased arousal, making affect more salient, as well as decreased time available to make analytical evaluations. Finucane et al. (2000) found that the negative relationship between perceived risk and benefit was greater when the participants were under time pressure. This suggests that the perceived relationship between benefit and risk is largely influenced by affect, and supports the influence of affect on risk judgments.
What aspect of affect is important? According to Slovic and colleagues (e.g., Slovic et al., 1985, 2000), dread and risk of the unknown are related to perceptions of risk. The more dread that an individual feels, and the more unfamiliar the risk, the more risk perceived in the object or hazard. Humans do not perceive all types of death as equal, dreading deaths which are painful and lengthy. It is this feeling of dread which leads to fear of nuclear power over driving, even though driving is more objectively risky than nuclear power. This further highlights the role of an affective component in the evaluation of hazards and risk.

The above research suggests that how people feel about risk can influence decision-making. People do not always make decisions based on cognitive evaluation. Rather, affect seems to play an important role. Past research into aeronautical risk management has focused on self-report measures, often assessing how much risk pilots perceive in the situation, and how much risk pilots are willing to accept. That is, the researchers gauged risk perception and risk tolerance by measuring explicit attitudes, or attitudes under conscious control.

Relying on self-report measures poses a problem, as it assumes that people are willing and able to report these attitudes (Egloff & Schmukle, 2002; Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellot, 2002). A pilot may not be able to report his or her level of risk tolerance, as he or she may lack awareness of, or be unwilling to report, high tolerance for risky situations.

There is often a dissociation between self-reported attitudes and behaviour (e.g., Festinger, 1964; LaPiere, 1934; Minard, 1952; Wicker, 1969). For example, LaPiere (1934) travelled across the United States with a Chinese couple and recorded the reactions of restaurant and hotel owners. Only one out of 251 owners refused to serve the couple. When contacted later, 92% of the owners said that they would not serve Chinese guests. Although a poorly controlled study, it illustrates how attitudes and behaviour are sometime disparate. More recent research suggests that sometimes attitudes can be a strong predictor of behaviour (Fazio, 1990). This relationship ranges from weak to strong, and depends on a large number of mediating factors. Furthermore, attitudes best predict behaviour when the same level of specificity of behaviour and attitudes are being measured (Ajzen & Fishbein, 1977).

Along with explicit attitudes, humans possess implicit attitudes. Greenwald and Banaji (1995) define implicit attitudes as "introspectively unidentified (or inaccurately
identified) traces of past experience that mediate favourable or unfavourable feeling, thought, or action toward social objects” (p 8). Therefore, unlike explicit attitude measures, implicit attitude measures assess feelings that the individual is unaware of, or does not wish to divulge, and as such, these measures bypass the problems of self-report measures.

Implicit attitudes are measured in an indirect way, by misinforming or not telling participants what is being measured (Greenwald & Banaji, 1995). There are a number of ways to measure implicit attitudes and associations, including the semantic priming procedure (Meyer & Schvaneveldt, 1971) and the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). These procedures measure reaction times, which are related to attitude accessibility (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). The stronger the association between an object and the evaluation, the quicker the participant’s reaction time should be (Fazio et al., 1986). For example, if a participant holds an attitude towards an object (e.g., candy tastes nice), he or she should respond faster when the evaluation precedes the object (e.g., words with positive connotations preceded candy) then when the opposite evaluation precedes the object (e.g., words with negative connotations preceded candy). Fazio et al. (1986) found that this facilitation only occurred when the attitude was strong, and when the participants were under time pressure.

The semantic priming procedure: During the semantic priming procedure, a prime (e.g., “nurse” or “doctor”) is presented for 200 milliseconds. After a blank screen, the target word is presented (e.g., “he” or “she”) which the participants categorise into two groups (e.g., “male” or “female”). This procedure was used by Banaji and Hardin (1996) to measure implicit gender stereotyping. If a participant holds an implicit association that doctors and nurses are usually males and females, respectively, then the participant should be faster at the classifying task when “nurse” precedes the female pronoun, then when “nurse” precedes the male pronoun. Banaji and Hardin (1996) found than an implicit association towards doctors and nurses being male and female occupations, respectively, was found independent of the participant’s explicit association.
The Implicit Association Test: The IAT measures implicit associations between two target concepts (e.g., flowers and insects) and two attributes (e.g., pleasant and unpleasant). Participants sort words or pictures (e.g., types of flowers and insects; pleasant and unpleasant words) into categories (e.g., flower/insect, pleasant/unpleasant) using one hand for each response. In this example, participants sort pictures of insects and unpleasant meaning words using one response key, and pictures of flowers and pleasant meaning words with another response key. On completion of these trials, participants sort pictures of insects and pleasant meaning words with one hand, and pictures of flowers and unpleasant meaning words with the other. The participant will show an implicit association between an attribute and a concept if they are faster at sorting when they are paired together (e.g., classifying pictures of insects and unpleasant words with one hand, and pictures of flowers and pleasant words with the other hand).

In this example, the participant shows an implicit association between flowers and pleasantness, and insects and unpleasantness. From this we can infer that the participant has more positive attitudes towards flowers than insects. Both the semantic priming procedure and the IAT measure attitudes or associations that people are unaware of, or unwilling to report. Greenwald et al. (1998) developed the first IAT to measure implicit valence towards musical instruments versus guns, and flowers versus insects. Participants reacted faster when flowers and pleasant, and insects and unpleasant, were paired together than flowers and unpleasant, and insects and pleasant, indicating that the participants had an implicit liking for flowers over insects. Similar results were found with weapons and musical instruments, with participants indicating that they liked instruments more than weapons.

The IAT has been successfully adapted to measure a wide range of attitudes and associations, including prejudice, fear, anxiety, and self-esteem. A number of these studies have assessed the validity and reliability of the IAT. A test is valid if it measures the construct it is created to measure. There are many different methods of assessing the validity of a measure, including predictive, incremental, and convergent validity. These methods of testing validity will be discussed later in the chapter. If the IAT is a valid measure of an attitude, groups that are known to hold different attitudes should show differential attitudes using the IAT. This has been shown assessing prejudice (Cunningham, Preacher, & Banaji, 2001; Dasgupta, McGhee, Greenwald, & Banaji, 2000;
Greenwald et al., 1998; Rudman, Greenwald, Mellot, & Schwartz, 1999), and fear towards phobic stimuli (Teachman, Gregg, & Woody, 2001).

The IAT procedure has been successful in measuring implicit racism. Japanese and Korean-Americans' showed an in-group bias (Greenwald et al., 1998). Japanese and Korean Americans found it more difficult to perform an IAT (shown by longer latencies) when surnames associated with their ethnic group (e.g., Kawabashi & Youn) were paired with unpleasant words (e.g., poison) compared to when surnames associated with their ethnic group were paired with pleasant words (e.g., happy). Greenwald et al. adapted this procedure to measure implicit attitudes towards African-Americans. White-Americans found it more difficult to perform an IAT when first names associated with White-Americans (e.g., Heather) were paired with unpleasant words, or names associated with African-Americans (e.g., Tashika) were paired with pleasant names, compared to the reverse. This suggested that these participants showed an in-group bias. The participants also rated their explicit attitudes towards African and White-Americans. Although there was a strong implicit in-group bias, the participants did not show an explicit in-group bias. A bias towards White-Americans was also shown by Cunningham et al. (2001) and Dasgupta et al. (2000). Rudman et al. (1999) found an in-group bias with Americans towards Russians, Christians and Jewish-Americans, and young people towards older people. In these three experiments, the participants responded quicker when their in-group was paired with words meaning pleasant, and when the out-group was paired with words meaning unpleasant, compared to the reverse.

The IAT has been successfully adapted to measure fear. Teachman et al. (2001) had individuals with spider and snake phobias complete four IATs measuring implicit valence, fear, disgust, and perceived danger. If these IATs were valid measures of implicit attitudes towards phobic stimuli, they should have been able to differentiate between people with spider and snake phobias. Participants' showed implicit dislike, fear, disgust, and perceived danger towards the feared animal, relative to the other animal. Individuals who were fearful of spiders responded faster when pictures of spiders were paired with words meaning unpleasant, fear, disgust, and danger, compared to when pictures of snakes were paired with words meaning unpleasant, fear, disgust, and danger. Similarly, participants who were fearful of snakes responded faster when pictures of snakes were paired with words meaning unpleasant, fear, disgust, and danger, compared to when pictures of spiders were paired with words meaning unpleasant, fear, disgust, and danger.
The participants in the aforementioned studies showed implicit attitudes and associations consistent with their group membership, suggesting that the IAT procedure successfully measures these attitudes and associations.

For the IAT to be a valid measure of implicit attitudes, it should also show good predictive, incremental, and convergent validity. Predictive validity would be shown if implicit attitudes predicted a related behaviour. Meanwhile, incremental validity requires that the IAT predict behaviour beyond that predicted by explicit attitude measures. The IAT procedure shows both predictive and incremental validity. Egloff and Schmukle (2002) adapted the IAT to measure implicit anxiety. Participants categorised words relating to self (e.g., my) and others (e.g., they) and words meaning anxiousness (e.g., nervous) and calmness (e.g., relaxed). Participants who showed implicit anxiety should have taken longer to respond when words relating to self were paired with words relating to calmness, and words relating to others were paired with words relating to anxiety, compared to the reverse. Participants also completed a simple discrimination task, before and after they received negative feedback concerning their performance on the discrimination task. The participants' implicit anxiety predicted their ratings of anxiety and performance on the discrimination task after they received negative feedback, beyond that accounted for by an explicit measure of trait-state anxiety.

Greenwald and Farnham (2000) found that implicit self-esteem predicted participants' reaction to poor performance. Participants completed an IAT measuring implicit self-esteem, where they classified words meaning self and others, and unpleasant and pleasant. The participants would have shown high self-esteem if they responded slower when words meaning self and unpleasant, and words meaning others and pleasant were paired together, compared to the reverse condition. Participants also completed either an easy or difficult discrimination task, and then rated their mood. High implicit self-esteem protected participants from negative feelings following the difficult discrimination task. Meanwhile, the participants with poor implicit self-esteem experienced a greater degree of negative feelings following poor performance.

McConnell and Leibold (2001) also found that the IAT has good predictive and incremental validity. While interacting with an African-American and White-American experimenter, white participants completed implicit and explicit measures of racism. Implicit racism towards African-Americans predicted behaviour towards the African-American experimenter, as the participants who showed implicit racism acted less
friendly towards the African-American experimenter compared to the White-American experimenter. Implicit racism better predicted behaviour towards the experimenters than the Modern Racism Scale, an explicit measure of racism.

Nosek, Baraji, and Greenwald (2002b) measured college students' implicit valence towards mathematics and arts. Implicit valence towards mathematics, relative to arts, was related to past mathematical performance. Although this is not a measure of predictive validity, it does show some relationship between attitude towards mathematics and mathematics performance. It may be that participants showed implicit dislike for mathematics, relative to arts, because they had performed poorly in the past.

The aforementioned studies suggest that the IAT has good predictive and incremental validity. However, Karpinski and Hilton (2001) found that the IAT did not have good predictive validity. Participants completed an IAT and an explicit measure of valence towards apples and candy bars. Participants were then given the option between taking a candy bar or an apple. Whereas the participants' explicit attitude (i.e., whether they said they liked candy bars or apples) predicted the choice of food, their implicit attitude did not.

A measure shows convergent validity if it correlates with other measures that assess the same construct (Cunningham et al., 2001). Cunningham et al. (2001) assessed this by correlating scores on the standard IAT procedure with that of two other implicit measures: response window evaluative priming and response window IAT. If these measures assess implicit attitudes, there should be a relationship between the measures. Response window evaluative priming involves presenting a stimulus for 200 milliseconds (e.g., a black or white face) before presenting words that are to be categorised (e.g., as negative or positive). The participant has a defined time to classify the words. The response window IAT includes a time limit for the participant to classify the words into categories. Cunningham et al. (2001) used these three measures to assess implicit racism towards black compared to white faces. The standard IAT demonstrated good convergent validity, as it was highly correlated with the other implicit measures.

The above studies largely suggest that the IAT has good validity. The IAT should also show good reliability, that is, it should be consistent within itself, and across time. IATs measuring implicit racism (Cunningham et al., 2001) and anxiety (Egloff and Schmukle, 2002) showed adequate internal consistency, as items within the IAT were correlated. Egloff and Schmukle (2002), and Cunningham et al. (2001) also assessed test-
retest reliability of the IAT. A measure has good test-retest reliability if the participants’ scores on one occasion correlate with scores of the same measure, on another occasion (Cunningham et al., 2001). Cunningham et al. (2001) found a test-retest reliability coefficient of .68, which represents good reliability. In contrast, Egloff and Schmukle (2002) found a reliability coefficient of .57. Although lower than recommended, it is higher than other implicit measures (Bosson, Swann, & Pennebaker, 2000; Egloff & Schmukle, 2002).

Other requirements of a measure of implicit attitudes are that it should not be affected by changes in context, participants’ motivational state, or conscious distortion. The IAT procedure was adapted by Sherman and colleagues (Sherman, Rose, Koch, Presson, & Chassin, 2003) to measure implicit attitudes towards cigarette smoking. Participants, who were smokers, completed measures of implicit and explicit valence towards cigarettes, insects, babies, and cuddly animals. The smoking-related stimuli either emphasised the sensory (e.g., a picture of a burning cigarette) or health and economic implications of smoking (e.g., a picture of the health warning on a cigarette package). Implicit attitudes did not differ depending on the emphasis of sensory or health/economic benefits of smoking, suggesting that the IAT was not sensitive to changes in context. The researchers then assessed whether implicit attitudes towards smoking were influenced by changes in motivational state, by manipulating exposure to cigarettes. Half of the participants smoked a cigarette before the study, and the other half were nicotine-deprived as they had not smoked for 4-hours before the experiment. The participants did not hold different implicit attitudes depending on whether they had a cigarette before the study, suggesting that the IAT was not sensitive to changes in motivational state. Meanwhile, explicit attitudes were influenced by both changes in context and motivational state.

Since implicit measures access information that participants are unaware of, the IAT should not be influenced by conscious distortion (Egloff & Schmukle, 2002). Egloff and Schmukle (2002) tested this by giving the participants IAT and self-report measures assessing anxiety. Half the participants were encouraged to fake both measures by imaging that they were trying to make a good impression on a possible future employer. In contrast to the participants in the control group, those in the faking group showed lower levels of explicit anxiety. However, the IAT scores were comparable across both groups.

Asendorpf, Banse, and Mücke (2002) employed a similar faking procedure to assess whether participants’ responses on an IAT, measuring implicit shyness, were influenced
by conscious distortion. Participants completed implicit and explicit measures of shyness. Prior to completing these measures, half of the participants were told to imagine that these questionnaires were part of a job interview, and that appearing confident was crucial for a successful application. The participants in the faking condition showed lower levels of implicit shyness compared to the participants in the control condition. However, the two groups of participants did not differ in implicit shyness. Both Egloff and Schmukle (2002) and Asendorpf et al. (2002) suggest that the IAT procedure is not subject to conscious distortion.

One of the major criticisms of the IAT is that it may be affected by familiarity. For example, the IAT designed to measure implicit racism towards African-Americans uses names associated with African-Americans and White-American as the target concepts. However, the IAT may be measuring differential familiarity, as names associated with White-Americans were more familiar to the White-American participants than names associated with Black-Americans (Dasgupta et al., 2000). A number of researchers have empirically investigated this idea. Rather than using stereotypically African-American and White-American names as stimuli, Dasgupta et al. (2000) measured implicit racism using pictures of African- and White-American faces. Dasgupta et al. found an implicit bias towards White-American faces despite both African-American and White-American pictures being equally unfamiliar. Rudman et al. (1999) manipulated familiarity in two studies. In their study looking at ageism, participants were asked to rate how familiar they were with the old and young names used in the IAT. Prior exposure to the target concepts had no effect on the IAT scores. Rudman et al. (1999) also looked at implicit prejudice towards Soviets by Americans. Four IATs were administered, using a combination of familiar and unfamiliar American and Soviet leaders. If prior exposure to target concepts contributed to implicit prejudice, differences in reaction time would be greatest with unfamiliar Soviet and familiar American leaders. However, this was not found, suggesting that prior exposure did not contribute to the in-group bias shown by the American participants. Further evidence against the familiarity argument comes from findings that in-group bias is shown when groups are assigned using the minimal-group paradigm (Ashburn-Nardo, Volis, & Monteith, 2001). That is, even when the participants were assigned to groups using an arbitrary classification, the participants showed in-group bias.

Another criticism of the IAT is that it is unrelated to explicit measures. Most studies comparing IAT scores to explicit measures, such as questionnaires, have found weak or
nonsignificant correlations (Cunningham et al., 2001; Dasgupta et al., 2000; Geer & Robertson, 2005; Greenwald et al., 1998; Karpinski & Hilton, 2001; Nosek, Banaji, & Greenwald, 2002a; Rudman et al., 1999; Teachman et al., 2001). This may be because implicit and explicit attitudes are measuring different constructs (Greenwald et al., 1998; Greenwald et al., 2001; Karpinski & Hilton, 2001). However, other studies have shown a strong relationship between explicit and implicit attitudes. For example, Wittenbrink, Judd, and Park (1997) found a relationship between explicit and implicit racism, using the semantic priming method as a measure of implicit racism. The sometimes weak or nonexistent relationship between implicit and explicit attitudes may be explained by the use of self-report methods. Often the explicit questionnaires concern socially sensitive subjects (e.g., racism) and participants may be unwilling to admit their true feelings. Meanwhile, since the implicit measures should not be sensitive to social desirability, the IAT should tap into these feelings. This is supported by Nosek, et al. (2002b), who measured attitudes towards mathematics. There was a relationship between implicit and explicit mathematics attitudes, an attitude which should not be socially sensitive.

Implicit associations and ADM: The above research suggests that the IAT is a valid and reliable measure of implicit attitudes. Can the IAT procedure be adapted to measure pilots' implicit attitudes towards weather, and are these implicit attitudes related to WRDM, such as flying into adverse weather? How pilots feel about risk may relate to pilot decision-making. For example, pilots may hold certain attitudes towards hazards and situations, such as adverse weather, and these attitudes may affect subsequent decision-making. Some of these attitudes may be implicit, as the pilots may not be aware that they hold such attitudes. Expert pilots make some decisions on an intuitive or implicit level. According to Adams and Ericsson (2000), since expert pilots make decisions so rapidly, it appears that it is not a conscious process. Rather, the decision seems to be made through immediate access to long term memory.

O'Hare and Wiegmann (2003) measured pilots' heart rate while flying a simulated flight into adverse weather. The pilots who continued into the adverse weather did not show as marked an increase in heart rate when confronted with adverse weather, as the pilots who diverted to another airport. Changes in cardiovascular activity are potentially driven by anxiety (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001), but also may be related to workload. Since it was not measured through self-report, heart rate can
be an implicit measure of anxiety. This suggests that the pilots who diverted to another airport showed more implicit anxiety than the pilots who chose to take the riskier option. This study shows the potential importance of implicit anxiety in making weather-related decisions.

The role of implicit anxiety and valence towards weather has also been measured with aeroplane passengers. Hughson (2002) empirically investigated whether implicit attitudes influenced aeroplane passengers’ decisions to take a flight into adverse weather. Non-pilots were presented with four flight scenarios, which depicted a GA flight conducted in different weather situations. The participants decided whether they would be a passenger on these flights, and rated how comfortable they would be doing so. The participants also completed two IATs, which measured implicit valence and fear of adverse and fine weather. For the measure of implicit valence, participants classified pictures of fine and adverse weather and words meaning pleasant and unpleasant into categories. Participants who showed implicit liking towards fine weather relative to adverse weather, should have reacted quicker when pictures of fine weather were paired with words meaning pleasant, and when pictures of adverse weather were paired with words meaning unpleasant, compared to the reverse. For the measure of implicit fear or anxiety, participants classified pictures of fine and adverse weather and words meaning anxious and calm into categories. Participants who showed implicit anxiety towards adverse weather should have reacted quicker when pictures of adverse weather were paired with words meaning anxiety, and when pictures of fine weather were paired with words meaning calm, compared to the reverse. The participants who chose to fly into bad weather during the hypothetical scenarios, tended to implicitly feel more positive, and more fearful, towards adverse weather compared to the other participants. However, these results did not reach statistical significance.

Risk perception and risk tolerance are related to WRDM (e.g., Hunter, 2002). Studies of aeronautical risk management have largely focussed on explicit processes. In my honours dissertation (Pauley, 2003) I assessed the role of implicit risk tolerance and risk perception in the decision to fly into adverse weather. Pilots were presented with 18 pen-and-paper flight scenarios. Each flight scenario depicted a different weather pattern, some of which were dangerous to fly in. Pilots indicated if they would fly in each condition and completed two IATs, measuring implicit risk tolerance and risk perception. To assess implicit risk perception, pilots classified pictures of IMC and VMC conditions
and words meaning risky and safe. If the pilot held an implicit attitude towards adverse weather being riskier than fine weather, he or she should have taken longer to classify the words and pictures when IMC was paired with safe, and VMC was paired with risky, compared to the reverse. Risk tolerance was operationalised as how exciting participants found flying in adverse weather, relative to fine weather. As discussed in chapter four, one reason why pilots may fly is to gain excitement. To assess risk tolerance, pilots classified pictures of IMC and VMC conditions and words meaning exciting and dull. If the pilots held an implicit attitude towards IMC conditions being more exciting than VMC conditions, he or she should have taken longer to classify words and pictures when IMC was paired with dull, and VMC was paired with exciting, compared to the reverse.

I found that whereas all pilots held an implicit attitude that adverse weather was more risky than fine weather, for some pilots there was a weaker association between risk and adverse weather. These pilots choose to fly in more hypothetical flight scenarios, suggesting a relationship between implicit risk perception and WRDM. That is, the pilots who chose to fly in more of the hypothetical scenarios implicitly perceived less risk involved in adverse weather. Meanwhile, the majority of participants held an attitude that VMC conditions were more exciting than IMC. However, there was no relationship between the number of flights taken and implicit risk tolerance.

The role of implicit attitudes towards weather in WRDM was further assessed by O’Hare et al. (2007). Qualified pilots completed three simulated flights into adverse weather, as well as IATs measuring implicit risk perception and risk tolerance. The IATs were identical to those used in my dissertation. Across all three flights, decision-making was not related to implicit risk perception or risk tolerance. That is, the pilots who continued into adverse weather did not hold different implicit attitudes than the participants who did not continue into adverse weather. However, there was some relationship between implicit attitudes and involvement in aeronautical incidents, using Hunter’s (1995) HES. The weaker the pilot’s association between adverse weather and risk, the more weather-related incidents, such as flying into adverse weather, the pilots had been involved in. Similarly, the stronger the association between adverse weather and excitement, the more weather-related incidents the pilots had been involved in.

In O’Hare and Wiegmann’s (2003) study, pilots who flew into adverse weather during a simulated flight did not experience as marked an increase in heart rate when the weather deteriorated, as compared to the participants who did not fly into adverse
weather. This suggests that anxiety or fear may play a role in WRDM. The role of implicit risk perception and implicit anxiety towards weather and WRDM will be further assessed in study five. Qualified pilots will complete a simulated flight into adverse weather and IATs measuring implicit risk perception and implicit anxiety towards adverse weather. Implicit risk perception will be assessed in the same manner as Pauley (2003) and O’Hare et al. (2007). Implicit anxiety will be assessed in the same manner as Hughson (2002).

Time Pressure

Operators of complex systems such as nuclear power plants, operating rooms, and aircraft, often make decisions under time constraints (Christensen, Fetters, & Green, 2005; Kerstholt, 1994; Ordóñez & Benson, 1997; Sarter & Schroeder, 2001). Pilots can face time pressure from delays, and time pressure from passengers, other pilots, and companies (Holbrook et al., 2003; McElhatton & Drew, 1993). Since decision-making often occurs under time constraints, it is important to understand how pilots and other operators cope in such situations.

Time constraints can lead to increased psychological stress, a sense of unavoidable failure, and feelings of helplessness (Ben Zur & Breznitz, 1981), which can lead to poorer individual (Bar-Eli & Tractinsky, 2000; Hoge, 1970; Holbrook et al., 2003) and team performance (Lusk, 1993; McElhatton & Drew, 1993; Neck & Moorhead, 1995; Urban, Weaver, Bowers, & Rhodenizer, 1996), changed decision strategy (Christensen et al., 2005; Johnson, Payne, & Bettman, 1993; Kerstholt, 1994; Maule, Hockey, & Bdzola, 2000; Maule & Mackie, 1990; Payne, Bettman, & Johnson, 1988; Raby & Wickens, 1994; Wright, 1974; Zakay, 1985), and increased risk avoidance (Ben Zur & Breznitz, 1981; Busemeyer, 1985; Wright, 1974).

Zakay (1993) presents a model of decision-making under time pressure, shown in Figure 5.1.
According to Zakay’s (1993) model, the effect of time pressure on performance depends on the complexity of the task. With simple tasks, time pressure does not lead to a shortage of resources and subsequently the operator can perform normally. Meanwhile, with more complex tasks, the operator or decision maker will make up for the shortage of resources by using simpler decision strategies. The shortage of resources also will lead to more errors. Together, the simpler decision strategies and performance errors will lead to suboptimal performance under time pressure.

Time pressure can worsen group performance, leading to non-optimal decisions being made. For example, Lusk (1993) examined the performance of two-person teams of aviation weather forecasters over a 37-day period. Time pressure was operationalised as the degree of environmental activity in the area. The more environmental activity, the more time pressure the forecasters would feel, and the less time they could spend...
attending to each storm. The forecasters' ability to predict the likelihood of storms was reduced when under time pressure, compared to when they were not under time pressure.

Time constraints can also lead to poorer individual performance. Bar-Eli and Tractinsky (2000) presented professional basketball coaches with videos of basketball games. The coaches rated the decisions made at the end of the game, when the players were under the most time pressure, as poorer compared to decisions earlier in the game. Not only can time pressure lead to poor decision-making and performance, but individuals may be more confident in their decisions when under time pressure compared to when they are not under time pressure. Hoge (1970) presented undergraduate psychology students with written descriptions of aircraft, described by four characteristics. The participants were given incomplete information, as they were never given information regarding all four characteristics, and selected the most ideal aircraft to perform a military operation based on the information given. Participants operated under three levels of time pressure (no time limit, 15 second limit, or an 8 second limit). In the highest time pressure condition the participants made the greatest number of errors, as they were less likely to select the best aircraft. Even though they made more errors, when the participants were under time pressure they were more confident that they had made the best decision compared to when they were under less time pressure.

Pilot performance can suffer under time pressure. Three percent of incidents reported to NASA's Aviation Safety Reporting System involved time pressure (McElhatton & Drew, 1993). Time pressure while flying can result in plan continuation errors (Holbrook et al., 2003) and 'hurry up' syndrome (McElhatton & Drew, 1993). As discussed in chapter one, plan continuation errors occur when pilots continue with a plan when faced with evidence that they should revise the plan. In Holbrook et al.'s interviews with Alaskan GA pilots, time pressure was often a contributing factor towards plan continuation errors. For example, if faced with pressure from passengers to reach the destination by a given time, the pilot may be reluctant to turn back or delay take-off if the weather deteriorated.

The 'hurry up' syndrome is when the pilot perceives that he or she must hurry up, which can lead to increased risk taking (McElhatton & Drew, 1993). For example, if a pilot tries to make up for delays caused by bad weather or mechanical problems, he or she may be inclined to take-off when they should stay put, or take a more direct route. As discussed in chapter one, time pressure was implicated in the Braniff International airline
crash, where the pilots flew through a thunderstorm to avoid a late arrival (Nance, 1984; NTSB, 1968).

Decision makers can cope with increased time pressure by changing their decision strategy. When under time constraints decision makers compensate by taking less time to make decisions (Johnson et al., 1993; Kerstholt, 1994; Maule & Mackie, 1990; Payne et al., 1988). This acceleration of decision-making is achieved by considering less information (Wright, 1974); employing heuristics (Johnston et al., 1993; Payne et al., 1988), and prioritising tasks (Raby and Wickens, 1994). These strategies can lead to improved performance while under time pressure. For example, Johnson et al. (1993) found that when under time pressure, the use of heuristics (e.g., elimination-by-aspects) led to improved decision-making. Decision-makers tended to naturally employ these heuristics when under time pressure (Johnson et al., 1993). Raby and Wickens (1994) examined the effect of time pressure on decision-making strategies during simulated landing procedures. When under time pressure, pilots prioritised their tasks, spending more time on important tasks and less time on low priority tasks. Therefore, although time pressure may be implicated in poor decision-making, many operators adapt to increasing time constraints.

One of the predictions of the RPD model (Klein, 1993, 1997b) is that expert decision-makers will use a pattern-matching strategy. Subsequently, time pressure should not affect expert performance to the same degree as novice performance. Wiggins et al. (2002) presented pilots with flight scenarios. The pilots were given information important to the flight, such as the fuel, weight and balance, weather, aircraft serviceability, and a map, and decided whether they should go on the flight. Time pressure was manipulated by limiting the pilots to seven, four, or two minutes to access information. Any differences in decision outcome and decision-making process between experienced, intermediate, and novice pilots was only apparent when the pilots were under the greatest degree of time pressure. Similarly, Calderwood et al. (1988) found that whereas novice chess players' performance was degraded with increasing time pressure, the expert chess players' performance was unaffected by time constraints.

Studies of risk-taking behaviour using static scenarios have found that individuals become more risk averse under time pressure (Ben Zur & Breznitz, 1981; Busemeyer, 1985; Wright, 1974). For example, Wright (1974) presented participants with written descriptions of cars, and participants rated the attractiveness of the cars and the likelihood
of buying the car. When they performed the task under time pressure, participants focused on the negative attributes more often than when they were not under time pressure. This suggests that the participants became risk averse when faced with time constraints.

When the decision-making task is dynamic, participants take more risks under time pressure, compared to when not under time pressure (Kerstholt, 1994). In Kerstholt's task, participants imagined that they were attending to an athlete running a race. A simulation of the athlete running was presented to the participants, with continuously updated information about the athlete's health (temperature, heart rate, and dehydration). The participants' goal was to maintain the athlete's health at 100% by applying the appropriate treatment. For example, if the athlete became dehydrated, the participant was to give water. The participants had an incentive to keep the athlete healthy. The participant would lose money if the athlete collapsed, as well as losing money for information requests and losing money for every incorrect action (e.g., resting the athlete when he or she needed water). Time pressure was operationalised as rate at which the athlete's health declined. While under time pressure, the participants took more risks compared to when they were not under time pressure, as they spent less time waiting to see if the treatment was successful, which led to the athlete collapsing more often.

The relationship between time pressure and risk-taking may depend on the nature of the task (Kerstholt, 1994). If the task is static, the decision maker will take fewer risks when under time pressure than when not under time pressure. However, if the task is dynamic, the decision maker will take more risks under time pressure compared to when not under time pressure. According to Kerstholt (1994), this may be due to the different methods of inducing time pressure in static and dynamic tasks. In static tasks time pressure is usually induced by imposing time constraints. Meanwhile, during dynamic tasks, rather than artificially inducing time pressure, time pressure is linked to the changing situation. For example, while flying, a pilot may face increasing pressure to quickly make a decision when the weather deteriorates.

Increased workload can also lead to increased risk-taking in dynamic tasks (O'Hare et al., 2007). Increasing time pressure can lead to increased subjective workload (Bustamante, Fallon, Bliss, Bailey, & Anderson, 2005) and increased errors (Hart & Bortolussi, 1984). O'Hare et al. had pilots complete three simulated flights into adverse weather. Half of the pilots completed the flights while doing the Verbal Sternberg Task.
(high workload condition). During this task, participants indicate whether a word was present in a previously presented list. The rest of the pilots completed the flights without the concurrent task (low workload condition). In-flight decision-making was related to workload. The participants in high workload condition were more likely to continue into adverse weather compared to the pilots in the low workload condition. For example, for one flight, 91% of the pilots in the high workload condition continued past the critical decision point, compared to 27% of the pilots in the low workload condition. This study, together with Kerstholt’s (1994) study suggest that time pressure and workload can lead to increased risk-taking in dynamic tasks.

Time Pressure and Implicit Associations

Fazio and colleagues (e.g., Fazio, 1990; Fazio & Towles-Schwen, 1999) believe that the influence of attitudes and associations on behaviour depends on the time available to make a decision. If under time pressure, the individual will arrive at his or her decision through automatic activation of an attitude, without any conscious deliberation. The majority of decisions will be made in this manner. This would especially be true of pilots, who are often under time pressure. However, if the individual has sufficient motivation and time, the decision will be made following a cost-benefit analysis. The MODE model, Motivation and Opportunity as DEterminants (Fazio, 1990; Fazio & Towles-Schwen, 1999) predicts when explicit and implicit attitudes and associations best predict behaviour. If the individual has sufficient motivation and opportunity, then the explicit attitude or association should be the best predictor of behaviour. However, if the individual does not have sufficient opportunity or motivation, then the implicit attitude or associations should be the best predictor of behaviour. For example, if a pilot is under time pressure to make a decision, his or her unconscious (implicit) attitudes or associations should better predict their behaviour.

The predictions of the MODE model have been tested by Asendorpf et al. (2002), Friese, Wänke, and Plessner (2005), and O’Hare et al. (2007). Asendorpf et al. (2002) measured participants’ implicit and explicit associations of self with shyness. Implicit shyness was measured using an IAT, where the participants categorised words relating to self (e.g., ‘you’) and others (e.g., ‘they’), and words meaning shy (e.g., ‘inhibited’) and non-shy (e.g., ‘daring’). Meanwhile, during the measure of explicit shyness, participants indicated how well 40 bipolar adjective pairs described their personality. A ‘shyness
MODE model. Participants will complete two IATs measuring implicit risk perception and implicit anxiety. The pilots' implicit attitudes should more closely relate to in-flight decision-making when under time pressure, compared to when they are not under time pressure.
Chapter 6: Risk Management and In-Flight Decision-Making

Risk perception and risk tolerance are key components of ADM (AOPA, 2003). In my first four studies, I have developed measures of risk perception and risk tolerance in GA pilots. The aim of the next study is to examine the role of risk management and implicit associations in in-flight weather-related decision-making. Qualified pilots will complete a simulated ‘scud-running’ flight and four computer-based questionnaires. This flight will be similar to that used by O’Hare et al. (2007) and O’Hare and Wiegmann (2003). The simulated weather upon take-off will be above VFR minimums. Forty nautical miles into the flight, the cloud-base will lower to 1,500 feet. If the pilot continues flying beyond this point, he or she will experience a further weather change 90 nautical miles into the flight, where the cloud-base lowers to 800 feet. If, after this point, the pilot has not decided to turn around or divert to another airport, the pilot will be considered to have continued, and the simulation will be stopped. This will be considered the critical decision point. In O’Hare and colleagues’ previous studies, between 50 and 59% of the pilots continued beyond the weather deteriorating to 800 feet.

Pilots will then complete four questionnaires: the CWS procedure developed in study three, measuring expertise in aeronautical weather-related risk perception; the risk tolerance measure developed in study four; and two IATs, measuring implicit anxiety and risk perception towards adverse weather. My first aim is to assess the relationship between risk perception and risk tolerance and flying into adverse weather during a simulated flight. Specifically, I will assess if pilots who continue flying differ in expertise in aeronautical weather related risk perception and risk tolerance. Hypothesis 1 stated that the pilots who continue flying beyond the critical decision point will be less expert at perceiving aeronautical weather-related risks, and will be more tolerant of risks, compared to the pilots who turn back or divert to another airport.

My second aim is to assess the relationship between implicit anxiety and implicit risk perception towards adverse weather, and in-flight WRDM. Hypothesis 2 stated that the pilots who continue flying beyond the critical decision point will be less implicitly anxious towards adverse weather, and implicitly perceive less risk in adverse weather, compared to the pilots who divert.

My third aim in this study is to further investigate the effect of time-pressure on in-flight WRDM. Time pressure will be operationalised as the time it takes to cover the distance of the flight. Participants will either fly the route at normal-speed (low time...
pressure) or double-speed (high time pressure). If the simulation speed is set at double-speed, the pilots will take half the time to cover the distance than if the simulation is set at normal-speed. Under both conditions, the weather will change at the same geographical location. The pilots in the double-speed condition will experience the weather deterioration twice as quickly as that which will be experienced by the pilots in the normal-speed condition. The critical decision point will be 5 and 2.5 minutes after the second weather change for the pilots in the low and high time pressure conditions, respectively. Therefore, the critical decision point will occur at the same geographical location, but the pilots in the high time pressure condition will have less time to make a decision compared to the pilots in the low time pressure condition.

Following Kerstholt's (1994) study of the effects of risk-taking in a dynamic task, I predict that inducing time pressure will lead to more risk-taking. That is, I expect that the pilots in the double-speed condition will be more likely to continue beyond the critical decision point than the pilots in the normal-speed condition. Since experience protected pilots (Wiggins et al., 2002) and chess players (Calderwood et al., 1988) from the negative effects of time pressure on decision-making, I predict that the effects of time pressure on in-flight decision-making will be moderated by flight experience. Specifically, for the pilots in the double-speed condition, those who continue will be less experienced compared to those who divert. Meanwhile, for the participants in the normal speed condition, the pilots who continue and divert should not differ in flight experience.

Finally, I aim to assess the prediction of the MODE model (Fazio, 1990; Fazio & Towles-Schwen, 1999). When under time pressure, the pilots’ implicit associations will better predict decision-making compared to when not under time pressure. The pilots’ implicit anxiety and risk perception towards adverse weather will be compared for the pilots who continue and divert, and for pilots in the normal and double speed conditions. When under high time pressure, the pilots who perceive less risk and feel less anxiety should be more likely to continue than the pilots who perceive more risk and more anxiety towards adverse weather. When under low time pressure, there should be little difference in the implicit associations between the pilots who continue and divert. This would suggest that implicit attitudes better predict pilots’ spontaneous decision-making than controlled decision-making.
Method

Participants

The participants were 32 pilots from New South Wales, Australia. The two females and 30 males were aged between 18 and 65 years ($M = 33.97, SD = 15.65$). The pilots came from a range of backgrounds, and included flight instructors, student pilots, recreational pilots, agricultural pilots, and a retired airline pilot. The total hours flying experience ranged from 80 to 9,000 hours ($M = 993.92, SD = 2,128.47$). The pilots had an average of 761.94 hours pilot-in-command ($range = 20 – 8,600, SD = 1,835.76$) and a mean of 684.34 hours were pilot-in-command on cross-country flights ($range = 12 – 7,200, SD = 1,735.05$). The pilots had an average of 25.31 hours flying in the last 90 days ($range = 0 – 150, SD = 35.91$) and a mean of 15.06 hours cross country in the last 90 days ($range = 0 – 100, SD = 23.34$). A PPL, CPL, and ATPL were the highest certification held by 16, 14, and 1 pilot, respectively. One pilot held a student licence. This licence was held for an average of 5.29 years ($range = .15 – 23, SD = 6.65$). Of the 32 pilots, six held an instrument and instructor rating, four held an instrument but not an instructor rating, two held an instructor rating but not an instrument rating, and 20 held neither rating.

As a manipulation of time pressure, half of the pilots completed the simulated flight at double speed, while the other half completed it at normal speed. The mean age and experience levels of the two groups of pilots were compared and are presented in Table 6.1. Non-parametric tests were conducted and the two groups did not significantly differ in age or experience levels.
Table 6.1.

Age and Experience of Participants in the Normal and Double-Speed Conditions.

<table>
<thead>
<tr>
<th></th>
<th>Speed</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Double</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>33.38 (15.13)</td>
<td>34.56 (16.62)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Hours</td>
<td>679.28 (1823.78)</td>
<td>1308.56 (2413.59)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours Pilot-in-command (PIC)</td>
<td>411.08 (1226.58)</td>
<td>1112.81 (2279.81)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIC cross-country</td>
<td>505.63 (1785.59)</td>
<td>863.06 (1721.92)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PIC 90-days</td>
<td>16.81 (25.24)</td>
<td>33.81 (43.30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-country 90-days</td>
<td>10.19 (14.57)</td>
<td>19.94 (29.37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length held</td>
<td>4.52 (6.04)</td>
<td>5.96 (7.27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Apparatus

Flight simulator: The dual-control flight simulator was set in a white-fibre glass frame, similar in shape to the front of a light aircraft. The flight simulator was controlled by Precision Flight Controls, Inc. The control panel was configured like a Cessna 172, and replicated the electronic switch, flaps, gears, throttle, yoke, and rudder pedals. The instrument panel was presented on two LCD screens, and replicated the instrument panel of a Cessna 172S. The instrument panel included a moving-map Garmin GNS-430 Global Positioning System (GPS). The GPS was pre-programmed with the route, and included information on nearby airports, distance travelled, and distance to travel. The simulated flight was created and run using X-Plane (version 8) by Laminar Research. The outside view was displayed on a 2.2*3m projection screen. The graphics were created using a Nvidia GeForce 7950 GX2 graphic card. Figure 6.1 depicts the set up of the flight simulator.
Computer-based questionnaires: The four computer-based questionnaires were presented on an Acer Travel-Mate 270 laptop computer. The questionnaires were created and presented using the Inquisit 1.3 program by Millisecond software.

Materials

Simulated flight: The flight simulated a ‘scud-running flight’. This is when pilots lower their altitude to remain flying VFR. The flight was approximately 154nm along the west coast of the United States of America. Participants departed from Quillayute (Washington) and flew a 150 degree direct heading to Tillamook (Oregon). The simulated weather at Quillayute State Airport (KUIL) was overcast at 2,500 feet, with an eight knot wind at 268° and 25 km visibility. The temperature and dew point were 18° and 5°, respectively, and the QNH was 29.92 inches. The first weather change occurred 40nm into the flight, where the cloud base decreased to 1,500 feet. If the pilot continued flying, he or she would experience another weather change. This occurred 90nm into the flight, where the cloud base decreased to 800 feet. All other aspects of the weather remained the same.
Flight information: Before completing the simulated flight, pilots were presented with a 2.5 page booklet (see appendix H) to assist in flight planning, which included a description of the flight scenario and a simplified flight plan. Participants were informed that it was 5pm on a Sunday afternoon, that they were returning home to Tillamook with a friend, and that they were expected at work on Monday morning. Participants were told that the flight was 154nm and that they would be flying a Cessna 172. The pilots’ task was to fly the pre-planned direct route between Quillayute and Tillamook using the laminated sectional chart and equipment provided, and that they were to fly this route with the aid of the moving-map GPS system installed in the aircraft. Participants were then informed that they had 73 litres of fuel; providing for a 77 minute flight time and a 45 minute reserve.

The flight plan included a weather forecast. Participants were informed that the cloud was overcast at 2000 feet, with 15 to 20 km surface visibility. The winds aloft at 1000 feet were 5 knots at 270°, at 3000 feet were 9 knots at 290°, and at 5000 feet were 12 knots at 270°. The freezing level was at 8000 feet. The simplified flight plan included waypoints, recommended altitude, and heading. It also included fuel burn, true speed, and time en-route.

Risk perception: Each participant was presented with the same 16 flight scenarios that were used in Study Three. Participants were also provided with aeronautical maps illustrating the routes. These were the same aeronautical maps as those presented to participants in Study Two.

Risk tolerance: Participants completed the same risk tolerance measure which was developed in Study Four. Some of the data from Study Four were lost due to participants answering the questionnaire incorrectly; by marking two points on the Likert scale, marking a half-way point, or omitting scenarios. To avoid the loss of data, the questionnaire was computer based, and the pilots could only proceed to the next scenario when they had answered the scenario correctly. Other than the mode of presentation, the questionnaire was identical to that used in Study Four. Since the participants were from Australia, and one quarter of the scenarios gave information about the route, participants were provided with an aeronautical map illustrating the Christchurch to Timaru route (30 cm x 43 cm). This ensured that the participants knew that this was a coastal flight over the Canterbury Plains of New Zealand.
IAT: The participants completed two IATs: Anxiety IAT and Risky IAT. The Anxiety IAT measured the pilots’ implicit fearfulness towards VMC and IMC, by measuring the associations between words meaning afraid and unafraid, and pictures of VMC and IMC. The Risky IAT measured the implicit risk associated with VMC and IMC, by measuring the associations between words meaning risky and safe, and pictures of VMC and IMC. Following Hughson (2002), Pauley (2003), and O’Hare et al. (2007), the categories were represented by five words and pictures, each word and picture being presented twice.

The pictures were colour photographs of VMC and IMC weather (600 x 450 pixels, see appendix I). The IMC and VMC pictures were those used in Hughson (2002), Pauley (2003), and O’Hare et al. (2007). The pictures used in Hughson (2002) were pre-rated to ensure that they were representative of VMC and IMC weather. All pictures used were rated as being good examples of VMC and IMC.

The program instructions were presented in Arial Text font, size 36, in red upper case letters. The attribute exemplars and the category labels were presented in blue and green uppercase letters, respectively. Both the attribute and category labels were presented in Arial font, size 20. All words and pictures were presented against a white background. The words and pictures which the participants classified were presented in the middle of the screen, while the category labels were situated at the top left and right of the screen. The attribute words for the IAT Anxiety were taken from Hughson (2002), while the attribute words for IAT Risky were taken from Pauley (2003) and O’Hare et al. (2007). The attribute words used in the two IATs are shown in Table 6.2.

Table 6.2.

<table>
<thead>
<tr>
<th>Category Labels and Attribute Exemplars</th>
</tr>
</thead>
<tbody>
<tr>
<td>IAT</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Anxiety</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Risky</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Demographic questionnaire: The pilots completed the same questionnaire as described in Study One.

Procedure.

All participants were tested individually. Participants read the information sheet and signed the consent form. Since the computer-based questionnaires concerned weather, the participants always completed the simulated flight first. This was to ensure that the participants did not realise that the study concerned weather-related decision-making.

Simulated flight: Participants were then given the flight plan, the map of the route, and shown where Quillayute and Tillamook were on the map. The pilots were given as much time as required to examine the flight plan, and were encouraged to ask any questions.

Pilots were given the following verbal instructions:

You are the pilot-in-command and are free to fly how you wish. However, please keep in mind the physical limitations of the aircraft and fly as you do in real-life. Don’t be alarmed if you notice that the aircraft appears to be travelling faster than the indicated airspeed; this is expected. You have the use of a GPS; the route is already programmed in. Let me know if you decide to do anything other than continue to your original destination.

Participants then sat in the left (pilot-in-command) seat and were shown how to interpret the GPS. The pilots were told to assume that the pre-flight inspection had already been conducted, and were then shown the controls. When ready to begin, the pilot started the ignition and prepared to take off. If the pilot was in the double-speed condition, the simulator was set on double-speed after the pilot had completed take-off.

The simulation was paused if the participants decided to divert to another airport or turn back to Quillayute, and the pilots were told to turn the aircraft towards this direction. If the participant had not made a decision before the critical decision point, he or she was classified as a continuer. If the pilot was in the normal speed condition, the critical decision point was 5 minutes after the second weather change. Meanwhile, for the participants in the double-speed condition, the critical decision point was 2.5 minutes after the second weather change.

Participants completed the four computer-based questionnaires after the simulated flight. The order of the IATs, and the risk perception and risk tolerance measures was counterbalanced.
Risk perception: The participants were given the same instructions as given in Study Three. The procedure was identical to that followed in Study Three.

Risk tolerance: Participants were presented with the same scenarios as Study Four and were given the same instructions. They recorded their response by pressing the corresponding key on the numerical keyboard.

IATs: The IATs were also counterbalanced, so that half of the participants completed the Risky IAT first, while the others did the Anxiety IAT first. For the Risky IAT, the participants were given the following instructions on the computer.

Our research investigates cognitive processes that are used in decisions that involve memory. We are seeking to develop and test theories of the cognitive processes that occur inside and outside of awareness in the routine use of memory. Stimuli will be presented on this display screen, and your responses will be entered on the keyboard.

The research assumes that you can read English fluently, and that your vision is normal or corrected to normal. If you do not consider yourself fluent in English, or if your vision is not normal or corrected to normal, AND ESPECIALLY IF YOU ARE HAVING SOME DIFFICULTY READING THIS DESCRIPTION, PLEASE ask the experimenter now whether you should continue (you will receive payment in any case). Your identity as a subject is confidential. Further, you are free to discontinue participation at any time, without penalty.

For each of several sorting tasks you will be shown pictures or words one at a time in the middle of the computer screen. Your task is to sort each item into its correct category as fast as you can by pressing EITHER the ‘d’ key or the “k” key. IMPORTANT: Press the ‘d’ key using your left index finger, or ‘k’ key using your right index finger. The categories associated with the ‘d’ and ‘k’ keys will be shown at the top of each screen. Please pay close attention to these category labels – they change for each sorting task!

For one of the sorting tasks you will be classifying images as being either ‘VMC’ or ‘IMC’. In the other sorting task you will be classifying words as meaning either ‘RISKY’ or ‘SAFE’. For each task, please judge each item on the basis of which group it appears to belong to. If you make an error you will see a red ‘X’ displayed on or under the stimulus. When this occurs you need to make the correct response to proceed.

The participants were given the same instructions for the Anxiety IAT, but were told that they were sorting words as meaning ‘AFRAID’ or ‘UNAFRAID’.

The IATs used in the present experiment are a variation of the IAT created by Greenwald et al. (1998). Following Hughson (2002), the present experiment used 40 data collection trials per block instead of the 50 used by Greenwald et al. (1998). Both the Anxiety and Risky IAT comprised 5 blocks. The order of the blocks was counterbalanced,
so that half of the participants were assigned order one, and the other half were assigned order two. The order of the blocks for the Anxiety IAT and the Risky IAT are presented in Table 6.3 and Table 6.4, respectively. Participants responded by pressing “D” or “K” if the word or picture belonged in the left or right category, respectively.

Table 6.3.
The Order of the Blocks for the Anxiety IAT

<table>
<thead>
<tr>
<th>Block</th>
<th>Left (“D”)</th>
<th>Right (“K”)</th>
<th>Left (“D”)</th>
<th>Right (“K”)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IMC</td>
<td>VMC</td>
<td>VMC</td>
<td>IMC</td>
</tr>
<tr>
<td>2</td>
<td>Unafraid</td>
<td>Afraid</td>
<td>Unafraid</td>
<td>Afraid</td>
</tr>
<tr>
<td>3</td>
<td>IMC+Unafraid</td>
<td>VMC+Afraid</td>
<td>VMC+Unafraid</td>
<td>IMC+Afraid</td>
</tr>
<tr>
<td>4</td>
<td>VMC</td>
<td>IMC</td>
<td>IMC</td>
<td>VMC</td>
</tr>
<tr>
<td>5</td>
<td>VMC+Unafraid</td>
<td>IMC+Afraid</td>
<td>IMC+Unafraid</td>
<td>VMC+Afraid</td>
</tr>
</tbody>
</table>

Table 6.4.
The Order of the Blocks for the Risky IAT

<table>
<thead>
<tr>
<th>Block</th>
<th>Left (“D”)</th>
<th>Right (“K”)</th>
<th>Left (“D”)</th>
<th>Right (“K”)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VMC</td>
<td>IMC</td>
<td>IMC</td>
<td>VMC</td>
</tr>
<tr>
<td>2</td>
<td>Safe</td>
<td>Risky</td>
<td>Safe</td>
<td>Risky</td>
</tr>
<tr>
<td>3</td>
<td>VMC+Safe</td>
<td>IMC+Risky</td>
<td>IMC+Safe</td>
<td>VMC+Risky</td>
</tr>
<tr>
<td>4</td>
<td>IMC</td>
<td>VMC</td>
<td>VMC</td>
<td>IMC</td>
</tr>
<tr>
<td>5</td>
<td>IMC+Safe</td>
<td>VMC+Risky</td>
<td>VMC+Safe</td>
<td>IMC+Risky</td>
</tr>
</tbody>
</table>

In the first two blocks the participants learnt which category the words and pictures belonged to. The first block is referred to as target sorting. In this block the participants sorted pictures of fine and adverse weather into IMC and VMC categories. The 10 pictures (5 of IMC, and 5 of VMC), each shown twice, made up the 20 practice trials. The second block is referred to as the attribute sorting block, and had the participants sorting words into the appropriate categories. The 10 words, 5 risky and 5 safe for the Risky IAT, and 5 afraid and 5 unafraid for the Anxiety IAT, were each shown twice and made up the 20
practice trials. The third block is called the first combined task. It differed from the first two blocks, as there were two category labels on each side of the computer screen. For each trial, the participant was presented with either words or pictures, and he or she had to categorise them appropriately. This block consisted of 24 practice and 40 data collection trials. The fourth block was referred to as reversed attribute sorting, which was the same as block 2, except the category labels were on the opposite sides. The final block, the second combined task, was similar to block 3. Again there were 2 labels on each side of the screen. The participants were presented with words and pictures and their task was to sort them into the correct category.

Blocks 3 and 5 were the critical blocks. Block 3 for those in order one, and block 5 for those in order two, was referred to as the compatible condition. In the Anxiety IAT, the participants classified VMC pictures and words meaning unafraid using their left hand, and IMC pictures and words meaning afraid using their right hand. With the Risky IAT, the participants classified VMC pictures and words meaning safe using their left hand, and IMC pictures and words meaning risky using their right hand. This block was also referred to as the compatible condition, as the pairings of the words and pictures are assumed to be compatible with the normal association in memory.

Block 5 for those in order one, and block 3 for those in order two, was referred to as the incompatible condition. In the Anxiety IAT, the participants classified VMC pictures and words meaning afraid using their left hand, and IMC pictures and words meaning unafraid using their right hand. With the Risky IAT, the participants classified VMC pictures and words meaning risky using their left hand, and IMC pictures and words meaning safe using their right hand. This block was also referred to as the incompatible condition, as the pairings of the words and pictures are assumed to be incompatible with the normal association in memory. The compatible and incompatible conditions for the Anxiety and Risky IATs are shown in Table 6.5.

<table>
<thead>
<tr>
<th>IAT</th>
<th>Compatible</th>
<th>Incompatible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>VMC+Unafraid/IMC+Afraid</td>
<td>VMC+Afraid/IMC+Unafraid</td>
</tr>
<tr>
<td>Risky</td>
<td>VMC+Safe/IMC+Risky</td>
<td>VMC+Risky/IMC+Safe</td>
</tr>
</tbody>
</table>

Table 6.5. The Compatible and Incompatible Conditions for the Anxiety and Risky IATs
Participants were given instructions regarding which categories were to be presented and side of presentation, prior to each block. An incorrect response resulted in a red ‘X’ underneath the word or picture. The ‘X’ stayed on the screen until the participant made the correct response.

After finishing the four computer-based questionnaires, the participants completed the demographic and aeronautical practices questionnaire. Participants were instructed to complete this as accurately as possible. At the end of the experiment the participants were debriefed, paid SA40, and thanked for their participation. The experiment took between 2 and 2.5 hours.

Results

Data Inspection

Before parametric tests were performed, the continuous variables were assessed for normality. If the residuals were not normally distributed, the variables were transformed using a square root transformation and outliers more than 3SDs away from the mean were removed. If this did not render the data normally distributed, a non-parametric test was used, with outliers included. If applicable, both raw and transformed data are reported, represented by the subscript “r” and “t”, respectively. All inferential tests are calculated using the transformed data. All statistical tests were evaluated against an alpha level of .05. Following O’Hare and Owen (1999), due to the complex nature of pilot decision-making, and the small sample size, results with a p-value less than .09 were considered useful.

Effect sizes for significant and marginally significant results were calculated using Cohen’s (1988) $d$ and partial eta squared ($\eta_p^2$). Cohen’s $d$ is calculated as the difference between the two means, divided by the pooled standard deviation. According to Cohen (1988), effect sizes of 0.2, 0.5, and 0.8 are regarded as small, medium, and large effects, respectively. Cohen’s $d$ also provides an estimate of the percentile standing of the experimental group, and the percent of non-overlap between the two distributions. Cohen’s $d$ was calculated when comparing two groups. Partial eta squared was calculated when an ANOVA was used. Partial eta squared represents the level of variance of the dependent variable predicted by the independent variable. According to Cohen (1988), .01 (1% of variance), .06 (6% of variance), and .14 (14% of variance) represent a small, medium, and large effect size, respectively.
Action at the First Weather Change

The first weather change occurred 40nm into the flight, just past Copalis State airport. At this point the cloud ceiling descended to 1,500 feet. Nineteen pilots continued beyond this weather change, while 13 pilots discontinued the flight before or soon after the first weather change. Table 6.6 illustrates the number of participants who continued or diverted beyond the first weather change in the normal and double speed conditions.

Table 6.6.
Actions of the Pilots in the Normal and Double-Speed Conditions after the First Weather Change

<table>
<thead>
<tr>
<th>Group</th>
<th>Normal</th>
<th>Double</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diverted</td>
<td>8</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>Continued</td>
<td>8</td>
<td>11</td>
<td>19</td>
</tr>
</tbody>
</table>

The proportion of continuers and diverters in each speed condition were compared, using a Pearson Chi-Square with a continuity correction. There was no significant difference in the number of continuers and diverters, between the two speed conditions, \( \chi^2(1) = 1.13, p = .288 \).

Action at the Critical Decision Point

The second weather change occurred 90nm into the flight, which was slightly south of Martin. At this point the cloud ceiling descended to 800 feet. If the pilots were flying at normal speed, they were given 5 minutes to make a decision. Meanwhile, if the pilots were flying at double speed, they were given 2.5 minutes to make a decision. If the pilots were still flying after this point, they were assumed to have continued. Table 6.7 illustrates the number of pilots who continued, diverted, or crashed beyond this point. Of the 32 pilots, 18 had diverted either before this weather change, or before the critical decision point, and were classified as diverters. Meanwhile, 12 pilots had not made a decision to divert by the critical decision point, and were classified as continuers. Due to difficulty controlling the aircraft, two pilots crashed the plane. One pilot crashed after the first weather change, but before the weather deteriorated further. The other crash occurred after the second weather
change, but before the critical decision point. The table also illustrates the decisions made by the pilots in the normal and double-speed conditions.

Table 6.7.

*Actions of the Pilots in the Normal and Double-Speed Conditions by the Critical Decision Point*

<table>
<thead>
<tr>
<th>Group</th>
<th>Normal</th>
<th>Double</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diverted</td>
<td>10</td>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>Continued</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Crashed</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Included in the group of pilots who diverted were the 13 pilots who had diverted soon after the first weather change. It would have been interesting to make comparisons between four groups of pilots: those who diverted before the first weather change (13 pilots); those who continued after the first weather change, but diverted before the critical decision point (5 pilots); those who continued beyond the critical decision point (12 pilots); and those pilots who crashed (2 pilots). However, there would have been few pilots per group, which would have further reduced the power of the statistics. Therefore, I have collapsed the number of pilots who diverted before the critical decision point into one group.

The proportion of continuers, diverters, and crashers in each group were compared, using a Pearson Chi-Square test. Fisher’s exact test was used, as the expected count in two cells was less than five. There was no significant difference in the proportion of pilots in each speed condition that continued, diverted, or crashed ($p = .620$, Fisher’s exact test).

*Age and Experience*

The age and experience level of the diverters and continuers is presented in Table 6.8.
Table 6.8.

Age and Experience of the Pilots Who Continued or Diverted After the First Weather Change.

<table>
<thead>
<tr>
<th></th>
<th>Diverted</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>29.54 (11.56)</td>
<td>37.00 (17.57)</td>
</tr>
<tr>
<td>Total Hours</td>
<td>742.65 (1404.24)</td>
<td>1165.84 (2531.98)</td>
</tr>
<tr>
<td>Hours PIC</td>
<td>553.08 (1211.82)</td>
<td>904.85 (2184.45)</td>
</tr>
<tr>
<td>PIC CC</td>
<td>490.23 (1165.88)</td>
<td>817.16 (2057.32)</td>
</tr>
<tr>
<td>PIC 90-days</td>
<td>22.38 (25.79)</td>
<td>27.32 (42.04)</td>
</tr>
<tr>
<td>CC 90-days</td>
<td>11.08 (11.95)</td>
<td>17.79 (28.69)</td>
</tr>
<tr>
<td>Length held (years)</td>
<td>4.33 (6.28)</td>
<td>6.02 (7.02)</td>
</tr>
</tbody>
</table>

The non-parametric equivalent of an independent t-test showed no differences in age or any measure of experience, between those who continued or diverted beyond the first weather change. The number of pilots with a student licence, PPL, CPL, ATPL, and who were instructor and instrument rated, who continued or diverted after the first weather change is presented in Table 6.9.

Table 6.9.  

Flight Certification of the Pilots Who Continued or Diverted After the First Weather Change

<table>
<thead>
<tr>
<th>Certification</th>
<th>Diverted</th>
<th>Decision</th>
<th>Continued</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>0</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>PPL</td>
<td>6</td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>CPL</td>
<td>7</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>ATPL</td>
<td>0</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Instrument Rated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>4</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>No</td>
<td>9</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Flight Instructor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>4</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>No</td>
<td>9</td>
<td></td>
<td>15</td>
</tr>
</tbody>
</table>
Table 6.11.

Flight Certification of the Pilots Who Continued or Diverted by the Critical Decision Point

<table>
<thead>
<tr>
<th>Certification</th>
<th>Diverted</th>
<th>Continued</th>
<th>Crashed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>PPL</td>
<td>8</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>CPL</td>
<td>9</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>ATPL</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Instrument Rated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>6</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>12</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Flight Instructor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>13</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>

The proportion of pilots holding each class of licence and rating, who diverted or continued, was assessed using a Pearson’s Chi Square with a continuity correction. Fisher’s exact test was used when the expected count in any cell was less than five. Two pilots were excluded from the chi-square for certification, as one held a student licence, and one held an ATPL. There was no difference between the proportion of pilots holding each licence and rating who continued, diverted, or crashed at the second weather change. Therefore, the pilots who continued, diverted, or crashed after the second weather change did not differ in age or any measure of experience.

Following Wiggins and colleagues (Wiggins et al., 1996; Wiggins et al., 2002; Wiggins & O’Hare, 1995), the pilots were categorised as experienced, intermediate, and novices. Those with less than 100 hours (7 pilots), between 101 and 1,000 hours (20 pilots), and more than 1,000 hours flight experience (5 pilots) were classified as novice, intermediate, and experienced, respectively. Pearson’s Chi Square with Fisher’s exact test revealed that the proportion of pilots who continued or diverted did not differ across the three experience levels, for either decision point.

A series of Pearson’s Chi Square with Fisher’s exact test were conducted to assess whether there was any relationship between experience and in-flight decision-making for the participants who were and were not under time pressure. The proportion of pilots who continued and diverted at the first weather change, in each experience group, and when under normal and double speed conditions is shown in Table 6.12.
Table 6.12.  
*The Proportion of Pilots Who Continued and Diverted at the First Weather Change, in Each Experience Group, and When Under Normal and Double-Speed Conditions.*

<table>
<thead>
<tr>
<th>Experience Group</th>
<th>Normal Continued</th>
<th>Normal Diverted</th>
<th>Double Continued</th>
<th>Double Diverted</th>
<th>Overall Continued</th>
<th>Overall Diverted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>57</td>
<td>43</td>
</tr>
<tr>
<td>Intermediate</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>2</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>Expert</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>60</td>
<td>40</td>
</tr>
</tbody>
</table>

There was no difference in the proportion of pilots who continued and diverted across the three levels of experience, for the pilots who flew at normal or double speed. The proportion of pilots who continued or diverted by the critical decision point, in each experience group, and when under normal and double speed conditions is shown in Table 6.13. The pilots who crashed were excluded from the analyses.

Table 6.13.  
*The Proportion of Pilots who Continued and Diverted by the Critical Decision Point, in Each Experience Group, and When Under Normal and Double-Speed Conditions.*

<table>
<thead>
<tr>
<th>Experience Group</th>
<th>Normal Continued</th>
<th>Normal Diverted</th>
<th>Double Continued</th>
<th>Double Diverted</th>
<th>Overall (%) (Continued)</th>
<th>Overall (%) (Diverted)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Intermediate</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Expert</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>25</td>
<td>75</td>
</tr>
</tbody>
</table>

There was no significant difference in the proportion of pilots who continued and diverted across the three levels of experience, for the pilots who flew at normal or double speed. However, the there was an overall tendency for the more experienced pilots to be more likely to divert. The lack of relationship between experience and in-flight decision-
making, for the pilots in the normal and double speed conditions may have been due to the small number of pilots in some of the cells.

A series of t-tests were conducted to assess whether there was any relationship between total hours experience and decision made, for the participants who were and were not under time pressure. At the first weather change, there was no difference in total hours experience between the continuers and diverters at normal speed, \( t(13) = -0.36, p = 0.725 \), or at double speed, \( t(4.564, ) = -1.09, p = 0.330 \). At the second weather change, there was no difference in the experience level between the continuers and diverters who flew at normal speed, \( t(13) = -0.71, p = 0.489 \). However, there was a marginally significant difference in the experience level of the continuers and diverters who were at double speed, \( t(7.46) = -2.17, p = 0.064 \). The pilots who continued tended to have flown fewer hours (\( M_r = 171.4, SD_r = 77.46, M_t = 12.81, SD_t = 3.02 \)) compared to the pilots who diverted (\( M_r = 1187.50, SD_r = 1693.25, M_t = 28.72, SD_t = 20.36 \)). The mean total hours experience for the pilots who flew normal and double speed, and for the pilots who continued and diverted is shown in Figure 6.2.

![Figure 6.2](image_url)

**Figure 6.2.** Mean total hours experience for the pilots who flew at normal and double speed and continued and diverted by the critical decision point.
Hazardous Event Scale

Pilots indicated how often they had experienced 12 potentially hazardous aeronautical situations in the past 24-months, on a scale from 0 to 4 or more. The pilots' involvement in hazardous events was calculated in the same manner as Studies One, Two, and Four. The pilots had been involved in a mean of 5.2 hazardous events in the past 24-months (range = 0 – 27, SD = 5.24). Two pilots did not complete this section of the questionnaire, one because he had not flown in the last 24-months, and the other because he ran out of time. These two pilots were excluded from the following analysis.

Involvement in hazardous events was compared between the pilots who continued and diverted at the first weather change, and the pilots who continued, diverted, and crashed at the second weather change. The pilots who continued or diverted at the first weather change did not differ in involvement in hazardous events. At the critical decision point, there was a significant difference in experience in past aeronautical events $F(2,26) = 5.67, p = .009, \eta_p^2 = .304$. The pilots who crashed had been involved in significantly more hazardous events ($M = 11, SD = 0$), than those who continued ($M = 3.80, SD = 2.57$), (mean difference = 7.20, $SE = .81, p < .001$) and those who diverted ($M = 4.06, SD = 3.09$), (mean difference = 6.94, $SE = .75, p <.001$). However, as there are only two people who crashed, a level of caution must be exercised when interpreting this finding. The relationship between action at the critical decision-point and involvement in hazardous events is depicted in Figure 6.3.
Figure 6.3. Mean number of hazardous events experienced by the pilots who continued, diverted and crashed at the first weather change.

Three of the HES items specifically addressed weather-related decision-making. The pilots were asked how many times they had flown VFR-into-IMC, became disorientated after entering IMC conditions, and had turned back or diverted after encountering adverse weather. The pilots’ involvement in weather-related aeronautical incidents was tallied in the same manner as their involvement in general aeronautical hazardous events. On average, the pilots had been involved in 2.4 weather-related aeronautical incidents (range = 0 – 8, SD = 2.40). There was no relationship between involvement in weather-related aeronautical incidents and in-flight decision-making after the first and second weather change.

Risk Perception

CWS: Each participant gave two risk ratings for each scenario. Discrimination, inconsistency, and the CWS score were calculated in the same manner as in Studies One, Two, and Three. One participant did not make a response for one scenario, so was excluded from the following analyses. The CWS scores ranged from .702 to 23.48 ($M = 6.59, SD = 5.55$). The CWS scores were adjusted using a square root transformation. The CWS scores did not significantly differ between the pilots who continued or diverted after
the first weather change, or between the pilots who had continued, diverted, or crashed by the second weather change.

The discrimination scores ranged from 21.46 to 2192.87 ($M = 1122.49, SD = 575.85$). Inconsistency ranged from 7.81 to 709.50 ($M = 257.60, SD = 194.02$). The discrimination and inconsistency scores did not significantly differ between the pilots who continued or diverted after the first weather change, or between the pilots who continued, diverted, or crashed by the critical decision point.

The pilots’ CWS, discrimination, and inconsistency scores did not significantly correlate with involvement in hazardous aeronautical events, or with weather-related hazardous aeronautical events.

CWS, discrimination, and inconsistency scores were not related to the age or experience of the pilot. The one exception was the number of hours flown pilot-in-command. After the outliers were removed, non-parametric correlations were performed. There was a positive relationship between discrimination and total number of hours ($r_s(32) = .454, p = .009$), indicating that the more hours experience, the more discrimination the pilots showed in the hypothetical scenarios. There was also a negative relationship between inconsistency and the number of hours flown pilot-in-command over the last 90 days ($r_s(32) = -.427, p = .015$), indicating that the more hours flown in the last 90 days, the more consistent the participants were.

The CWS scores were compared for pilots with a PPL and a CPL. The pilots with a CPL had a significantly higher CWS score ($M_t = 8.45, SD_t = 6.25, M_c = 2.75, SD_c = .99$) compared to those with a PPL ($M_t = 4.38, SD_t = 3.24, M_c = 1.96, SD_c = .76$). $F(1, 27) = 5.78, p = .023, \eta^2_p = .176, d = .90$. This is shown below in Figure 6.4. The mean CWS score for the pilots with a CPL was in the 82nd percentile of the pilots with a PPL, with 41.6% non-overlap between the two distributions. There was no significant difference in the discrimination and inconsistency scores between the pilots with a PPL and CPL.
The CWS scores of the pilots with and without an instructor and instrument rating were compared. There was a marginally significant effect of having an instructor rating, $F(1, 29) = 3.79, p = .061, \eta^2_p = .116, d = .74$. The flight instructors ($M_i = 9.79, SD_i = 7.46, M_t = 2.94, SD_t = 1.15$) tended to have higher CWS scores compared to the pilots without a flight instructor rating ($M_i = 5.47, SD_i = 4.39, M_t = 2.18, SD_t = .88$). This is shown below in Figure 6.5.

![Figure 6.4. Mean raw CWS score for pilots with a PPL and CPL](image)

![Figure 6.5. Mean CWS score for pilots with and without a flight instructor rating.](image)
There was no effect of having an instrument rating on the CWS scores. There were no significant differences in the discrimination and inconsistency scores between the pilots with and without an instrument rating and instructor rating.

*Average risk rating:* The average risk rating given by the participants who continued or diverted was compared. An average risk rating was calculated for each participant, for each replication. The higher the rating, the more risk the participant perceived in the scenario. The mean ratings for replication one and two were 62.18 (SD = 14.69) and 59.71 (SD = 16.32), respectively. A repeated measures ANOVA was conducted, with replication as the within-subjects variable and decision made as the between-subjects variable. There was no main effect of replication, and no decision*replication interaction. There was a trend for action made at the first weather change, $F(1,29) = 3.02, p = .093, \eta^2_p = .09$. The pilots who diverted soon after the first weather change tended to give higher risk ratings (replication 1: $M = 67.97, SD = 11.83$; replication 2: $M = 66.36, SD = 13.95$), compared to the pilots who continued after the first weather change (replication 1: $M = 54.77, SD = 19.12$; replication 2: $M = 52.15, SD = 18.05$). A repeated measures ANOVA, with a Sidak correction, showed a main effect of action made at the critical decision-point, $F(2,28) = 3.79, p = .035, \eta^2_p = .213$. The pilots who diverted before the critical decision point, gave higher average risk ratings (replication 1: $M = 67.74, SD = 11.86$; replication 2: $M = 63.56, SD = 15.26$), therefore perceiving more risk, compared to the pilots who crashed (replication 1: $M = 39.91, SD = 27.53$; replication 2: $M = 44.59, SD = 26.74$) (Mean difference = 23.40, $p = 10.21$). This is shown in Figure 6.6.
Figure 6.6. The average risk rating given by pilots who continued, diverted, and crashed, for replication one and two.

There was no replication*decision interaction. There was also no relationship between the average risk ratings and any indices of experience, other than certification. There was a marginally significant trend for the pilots with a CPL ($M = 65.27, SD = 10.24$) to give higher ratings of risk compared to pilots with a PPL ($M = 54.04, SD = 21.35$), $t(20.41) = 1.82, p = .083, d = .67$.

There was no relationship between the average risk rating and involvement in potentially hazardous aeronautical incidents. There was a marginally significant relationship between the average rating given for the first replication and the pilots’ involvement in potentially hazardous weather-related aeronautical incidents, $r_s(29) = -.331, p = .079$. The negative relationship suggests that the participants who were involved in more hazardous weather-related incidents tended to rate the hypothetical scenarios as less risky.
Risk Tolerance

To assess the participants' level of risk tolerance, participants were presented with the same set of scenarios that were used in Study Four. A multiple regression equation was calculated for each participant, indicating the influence of opportunity and threat on the rated likelihood of taking-off. The multiple regression equations were calculated in the same manner as Study Four.

Overall risk tolerance: A multiple regression equation was calculated for each participant, across all 36 scenarios, representing the influence of all facets of opportunity and threat on the pilots' decision to take-off. Table 6.14 illustrates the within-subject regression coefficients for opportunity and threat, and $R^2$ for each participant. The within-subject regression coefficients indicate the relationship between the likelihood of going on the flight and the independent variables (level of opportunity and threat). $R^2$ indicates the variation explained by the independent variables. According to Fritzsche et al. (2000), $R^2$ values should be greater than .50. If not, the participant's policy has not been captured, and he or she will be excluded from the analyses.
Table 6.14.
*Within-Subject Regression Coefficients for Overall Threat and Opportunity.*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.69**</td>
<td>-.14</td>
<td>.49</td>
</tr>
<tr>
<td>2</td>
<td>-.85**</td>
<td>-.07</td>
<td>.72</td>
</tr>
<tr>
<td>3</td>
<td>-.85**</td>
<td>.01</td>
<td>.73</td>
</tr>
<tr>
<td>4</td>
<td>-.72**</td>
<td>.04</td>
<td>.52</td>
</tr>
<tr>
<td>5</td>
<td>-.68**</td>
<td>-.25*</td>
<td>.52</td>
</tr>
<tr>
<td>6</td>
<td>-.89**</td>
<td>-.07</td>
<td>.80</td>
</tr>
<tr>
<td>7</td>
<td>-.91**</td>
<td>.07</td>
<td>.84</td>
</tr>
<tr>
<td>8</td>
<td>-.78**</td>
<td>-.03</td>
<td>.61</td>
</tr>
<tr>
<td>9</td>
<td>-.70**</td>
<td>-.05</td>
<td>.49</td>
</tr>
<tr>
<td>10</td>
<td>-.70**</td>
<td>-.05</td>
<td>.49</td>
</tr>
<tr>
<td>11</td>
<td>-.80**</td>
<td>-.02</td>
<td>.65</td>
</tr>
<tr>
<td>12</td>
<td>-.88**</td>
<td>.05</td>
<td>.78</td>
</tr>
<tr>
<td>13</td>
<td>-.88**</td>
<td>-.06</td>
<td>.78</td>
</tr>
<tr>
<td>14</td>
<td>-.75**</td>
<td>-.12</td>
<td>.58</td>
</tr>
<tr>
<td>15</td>
<td>-.77**</td>
<td>-.09</td>
<td>.61</td>
</tr>
<tr>
<td>16</td>
<td>-.70**</td>
<td>.00</td>
<td>.49</td>
</tr>
<tr>
<td>17</td>
<td>-.89**</td>
<td>.03</td>
<td>.80</td>
</tr>
<tr>
<td>18</td>
<td>-.81**</td>
<td>-.04</td>
<td>.66</td>
</tr>
<tr>
<td>19</td>
<td>-.86**</td>
<td>.07</td>
<td>.75</td>
</tr>
<tr>
<td>20</td>
<td>-.72**</td>
<td>-.06</td>
<td>.52</td>
</tr>
<tr>
<td>21</td>
<td>-.80**</td>
<td>.06</td>
<td>.65</td>
</tr>
<tr>
<td>22</td>
<td>-.70**</td>
<td>-.20</td>
<td>.54</td>
</tr>
<tr>
<td>23</td>
<td>-.91**</td>
<td>-.03</td>
<td>.82</td>
</tr>
<tr>
<td>24</td>
<td>-.76**</td>
<td>-.05</td>
<td>.58</td>
</tr>
<tr>
<td>25</td>
<td>-.49**</td>
<td>.11</td>
<td>.25</td>
</tr>
<tr>
<td>26</td>
<td>-.77**</td>
<td>.10</td>
<td>.60</td>
</tr>
<tr>
<td>27</td>
<td>-.90**</td>
<td>.04</td>
<td>.81</td>
</tr>
<tr>
<td>28</td>
<td>-.81**</td>
<td>-.11</td>
<td>.67</td>
</tr>
<tr>
<td>29</td>
<td>-.83**</td>
<td>-.12</td>
<td>.70</td>
</tr>
<tr>
<td>30</td>
<td>-.89**</td>
<td>-.08</td>
<td>.80</td>
</tr>
<tr>
<td>31</td>
<td>-.77**</td>
<td>-.06</td>
<td>.60</td>
</tr>
<tr>
<td>32</td>
<td>-.83**</td>
<td>.00</td>
<td>.68</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01.

In all cases, the pilots' responses were significantly influenced by the level of threat in the situation. The within-subject regression coefficients ranged from -.91 to -.49, indicating that the greater the threat in the situation, the less likely it was that the pilots would go on the flight. Meanwhile, the majority of pilots' responses were not significantly influenced by the level of opportunity in the situation. The within-subject regression coefficients ranged from -.25 to .11, indicating that for the majority of pilots, there was
little relationship between the level of opportunity and the decision to go on the flight. However, for one pilot there was a significant relationship between the level of opportunity and the likelihood of going on the flight, within-subject regression coefficient = -.25, $t = -2.07$, $p = .047$. The negative relationship shows that the higher the level of opportunity in the situation, the less likely it is that this pilot would go on the flight. The $R^2$ values ranged from .25 to .82, with five pilots with a $R^2$ less than .50. These participants were excluded from subsequent analyses.

There was a significant difference for the within-subject coefficient for threat between the participants who continued or diverted after the first weather change. For the participants who continued, there was a weaker relationship between the level of threat and the decision to take-off ($M = -.79$, $SD = .07$) than the participants who diverted ($M = -.85$, $SD = .05$), $F(1, 25) = 6.41$, $p = .018$, $d = .99$, $\eta^2_p = .20$. The average within-subject regression coefficient (threat) for the diverters, was in the 82nd percentile of the within subject regression coefficients (threat) for the continuers, with a 51.6% non-overlap between the two distributions. This is shown in Figure 6.7. The pilots who continued or diverted after the first weather change did not significantly differ in the influence of opportunity on the decision to take-off.
There was a marginally significant difference at the critical decision point in the influence of threat on the decision to take-off, \( F(2,24) = 3.09, p = .064 \). The difference was between those who continued or diverted, \( F(1,23) = 4.16, p = .053, \eta^2_p = .153, d = .79 \). The participants who continued tended to be less influenced by threat (\( M = -.78, SD = .08 \)) compared to the participants who diverted (\( M = -.84, SD = .05 \)). There was no difference in the influence of opportunity on the decision to take off, between the pilots who continued, diverted, or crashed by the critical decision point. The mean within-subject regression coefficients for the pilots who diverted, continued, and crashed by the critical decision point are shown in Figure 6.8.
Using Spearman’s rank order correlations, I found no relationship between the pilots’ involvement in hazardous events and the influence of opportunity and threat on the pilots’ decision to take-off.

In the scenarios, different facets of opportunity and threat were varied. Opportunity was operationalised as the potential to gain social approval, excitement, income, and career advancement. Threat was operationalised as the threat of loss from within the pilot, aircraft, environment, and time pressure. Since I was interested in the influence of each facet of opportunity and threat on the decision to take-off, I calculated separate multiple-regression equations for each way of operationalising opportunity and threat.

Social approval/Pilot: In these scenarios, high, medium, and low opportunity for social approval were represented by transporting a critically injured passenger, taking a friend on a flight, and taking a Friday afternoon off work to go on a flight, respectively. Meanwhile, high pilot-related threat was represented by the pilot having had 5 hours sleep in the previous 24-hours, drinking heavily the night before, on medication for a cold, and some personal problems. Medium pilot-related threat was represented by the pilot having had 8
hours sleep in the previous 24-hours, having not drunk any alcohol in the couple of days, on medication for a cold, and some personal problems. Low pilot-related threat was represented by the pilots having had 8 hours sleep in the previous 24-hours, having not drunk any alcohol in the last week, and feeling healthy with no personal problems. The within-subject regression coefficients for when the opportunity was for social approval and the threat was within the pilots are displayed in Table 6.15.

Table 6.15.

*Within-Subject Regression Coefficients for Pilot-Related Threat and Opportunity for Social Approval.*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>2</td>
<td>-.92**</td>
<td>.00</td>
<td>.85</td>
</tr>
<tr>
<td>3</td>
<td>-.91**</td>
<td>.00</td>
<td>.84</td>
</tr>
<tr>
<td>4</td>
<td>-.86**</td>
<td>.07</td>
<td>.74</td>
</tr>
<tr>
<td>5</td>
<td>-.99**</td>
<td>.00</td>
<td>.98</td>
</tr>
<tr>
<td>6</td>
<td>-.89**</td>
<td>.00</td>
<td>.79</td>
</tr>
<tr>
<td>7</td>
<td>-.97**</td>
<td>-.14</td>
<td>.96</td>
</tr>
<tr>
<td>8</td>
<td>-.91**</td>
<td>.12</td>
<td>.85</td>
</tr>
<tr>
<td>9</td>
<td>-.94**</td>
<td>.00</td>
<td>.87</td>
</tr>
<tr>
<td>10</td>
<td>-.86**</td>
<td>.31</td>
<td>.84</td>
</tr>
<tr>
<td>11</td>
<td>-.86**</td>
<td>.18</td>
<td>.77</td>
</tr>
<tr>
<td>12</td>
<td>-.90**</td>
<td>.13</td>
<td>.83</td>
</tr>
<tr>
<td>13</td>
<td>-.99**</td>
<td>-.07</td>
<td>.98</td>
</tr>
<tr>
<td>14</td>
<td>-.80*</td>
<td>-.23</td>
<td>.69</td>
</tr>
<tr>
<td>15</td>
<td>-.92**</td>
<td>.00</td>
<td>.85</td>
</tr>
<tr>
<td>16</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>17</td>
<td>-.89**</td>
<td>.00</td>
<td>.79</td>
</tr>
<tr>
<td>18</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>19</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>20</td>
<td>-.88**</td>
<td>-.08</td>
<td>.78</td>
</tr>
<tr>
<td>21</td>
<td>-.88**</td>
<td>.13</td>
<td>.79</td>
</tr>
<tr>
<td>22</td>
<td>-.86**</td>
<td>-.23</td>
<td>.79</td>
</tr>
<tr>
<td>23</td>
<td>-.90**</td>
<td>.26</td>
<td>.88</td>
</tr>
<tr>
<td>24</td>
<td>-.91**</td>
<td>.00</td>
<td>.84</td>
</tr>
<tr>
<td>25</td>
<td>-.79*</td>
<td>.22</td>
<td>.68</td>
</tr>
<tr>
<td>26</td>
<td>-.99**</td>
<td>.00</td>
<td>.99</td>
</tr>
<tr>
<td>27</td>
<td>-.99**</td>
<td>.00</td>
<td>.99</td>
</tr>
<tr>
<td>28</td>
<td>-.97**</td>
<td>.00</td>
<td>.93</td>
</tr>
<tr>
<td>29</td>
<td>-.96**</td>
<td>.00</td>
<td>.92</td>
</tr>
<tr>
<td>30</td>
<td>-.91**</td>
<td>.07</td>
<td>.83</td>
</tr>
<tr>
<td>31</td>
<td>-.89**</td>
<td>-.06</td>
<td>.74</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01.
Participant 25 indicated that he would definitely go on all 9 social/pilot scenarios. Since there was no variation in this pilot’s responses, a regression equation could not be calculated. Therefore, this pilot was excluded from the following analyses.

In all cases the pilots’ responses were significantly influenced by the level of threat in the situation. The within-subject regression coefficients ranged from -.99 to -.79, indicating that the more pilot-related threat in the situation, the less likely it was that the pilot would go on the flight. Meanwhile, in all cases the pilots’ responses were not significantly influenced by the level of opportunity in the situation. The within-subject regression coefficients ranged from -.23 to .31, indicating that there was little relationship between the level of social approval in the scenarios and the decision to go on the flight. The $R^2$ values for all of the pilots were greater than 0.5, indicating that the model captured the decision-making policies of all of the pilots. Subsequently, only participant 25 was excluded from the following analyses.

There was no significant difference in the influence of threat and opportunity on the decision to take-off between the pilots who continued or diverted after the first weather change. At the critical decision point, the two pilots who crashed (numbers 14 and 26), had the two highest (least negative) within-subject regression coefficients for threat. A non-parametric test showed that these pilots had significantly higher within-subject regression coefficients compared to the other two groups, (Crashers: $M = -.795$, $SD = .001$; Diverters: $M = -.917$, $SD = .05$, Continuers: $M = .901$, $SD = .036$) $Z(28) = 5.972$, $p = .05$, $d = 3.45$. Meanwhile, there were no significant differences between the continuers and diverters. The mean within-subject regression coefficient (threat) for the pilots who crashed was in the 97.7th percentile of the within-subject coefficients (threat) for the pilots who diverted, with 81.1% non-overlap between the two distributions. The mean within-subject regression coefficients for pilot-related threat for the pilots who continued, diverted, and crashed after the second weather change is depicted in Figure 6.9.
There was no significant difference in the influence of opportunity on the decision to take-off between the participants who continued, diverted, or crashed by the critical decision point. There was no relationship between the number of hazardous events which the participants had been involved in and the within-subject coefficients for opportunity for social gain and pilot-related threat.

*Excitement/Aircraft*: In these scenarios, high opportunity for excitement was represented by a cross-country flight the pilot had never been on, to a destination that the pilot had never been to. Medium opportunity for excitement was represented by a cross-country flight that the pilot had never been on, to a destination that the pilot had driven to. Low opportunity for excitement was represented by a cross-country flight that the pilot had been on many times before. High aircraft-related threat was represented by an aircraft that last had its annual inspection 11-months ago and a magneto drop of 200rpm. Medium aircraft-related threat was represented by an aircraft with its last inspection 9-months ago and a magneto drop of 100rpm. Low aircraft-related threat was represented by an aircraft which had its inspection 2-months ago and a magneto drop of 50rpm. The within-subject

---

*Figure 6.9.* Mean within-subject regression coefficients for pilot-related threat for the pilots who continued, diverted, and crashed by the second weather change.
regression coefficients for scenarios with opportunity for excitement and aircraft-related threat are displayed in Table 6.16.

Table 6.16.

*Within-Subject Regression Coefficients for Aircraft-Related Threat and Opportunity to Gain Excitement.*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.63</td>
<td>-.39</td>
<td>.55</td>
</tr>
<tr>
<td>2</td>
<td>-.95**</td>
<td>-.08</td>
<td>.90</td>
</tr>
<tr>
<td>3</td>
<td>-.89**</td>
<td>.06</td>
<td>.79</td>
</tr>
<tr>
<td>4</td>
<td>-.72**</td>
<td>.04</td>
<td>.74</td>
</tr>
<tr>
<td>5</td>
<td>-.97**</td>
<td>.00</td>
<td>.93</td>
</tr>
<tr>
<td>6</td>
<td>-.89**</td>
<td>-.15</td>
<td>.82</td>
</tr>
<tr>
<td>7</td>
<td>-.88**</td>
<td>.24</td>
<td>.83</td>
</tr>
<tr>
<td>8</td>
<td>-.93**</td>
<td>.07</td>
<td>.86</td>
</tr>
<tr>
<td>9</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>10</td>
<td>-.91**</td>
<td>.06</td>
<td>.84</td>
</tr>
<tr>
<td>11</td>
<td>-.61</td>
<td>-.34</td>
<td>.49</td>
</tr>
<tr>
<td>12</td>
<td>-.88**</td>
<td>.00</td>
<td>.78</td>
</tr>
<tr>
<td>13</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>14</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>15</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>16</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>17</td>
<td>-.88**</td>
<td>-.06</td>
<td>.79</td>
</tr>
<tr>
<td>18</td>
<td>-.91**</td>
<td>-.06</td>
<td>.84</td>
</tr>
<tr>
<td>19</td>
<td>-.98**</td>
<td>.15</td>
<td>.97</td>
</tr>
<tr>
<td>20</td>
<td>-.68</td>
<td>-.11</td>
<td>.48</td>
</tr>
<tr>
<td>21</td>
<td>-.89**</td>
<td>.06</td>
<td>.79</td>
</tr>
<tr>
<td>22</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>23</td>
<td>-.85**</td>
<td>-.13</td>
<td>.74</td>
</tr>
<tr>
<td>24</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>25</td>
<td>-.87**</td>
<td>-.22</td>
<td>.80</td>
</tr>
<tr>
<td>26</td>
<td>-.88**</td>
<td>.00</td>
<td>.78</td>
</tr>
<tr>
<td>27</td>
<td>-.95**</td>
<td>.00</td>
<td>.89</td>
</tr>
<tr>
<td>28</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>29</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>30</td>
<td>-.95**</td>
<td>.00</td>
<td>.90</td>
</tr>
<tr>
<td>31</td>
<td>-.87**</td>
<td>-.27</td>
<td>.84</td>
</tr>
<tr>
<td>32</td>
<td>-.87**</td>
<td>.00</td>
<td>.87</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01.

For all but three cases, the pilots’ responses were significantly influenced by the level of threat in the situation. These within-subject regression coefficients ranged from -.98 to -.61, indicating that the more aircraft-related threat in the situation, the less likely the pilot would go on the flight. For the three pilots whose responses were not significantly
influenced by the level of threat in the situation, the within-subject regression coefficients were close to significance, participant 1: $t = -2.304, p = .061$; participant 11: $t = -2.102, p = .08$, participant 20: $t = -2.314, p = .06$. Meanwhile, in all cases, the pilots' responses were not significantly influenced by the level of opportunity in the situation. The within-subject regression coefficients ranged from -.39 to .24, indicating that the pilots' responses were not significantly influenced by the opportunity to gain excitement. Two pilots had $R^2$ values of less than 0.5, and these pilots were excluded from the following analyses.

There was no difference in the influence of opportunity or threat on the decision to take-off between the pilots who continued or diverted after the first weather change, or between the pilots who had continued, diverted, or crashed by the second weather change. There was also no relationship between the number of hazardous events previously experienced and the influence of threat or opportunity on the decision to take-off.

*Income/Weather:* The within-subject regression coefficients for the scenarios with weather-related threat and the opportunity to earn income are presented in Table 6.17. Opportunity to earn income was operationalised as the number of passengers on board the aircraft, with high, medium, and low opportunity represented by 5, 3, and 2, respectively. High weather-related threat was represented by isolated thunderstorms and rain, 1000 metres visibility, overcast cloud at 800 feet, and 35 knot winds. Medium weather-related threat was represented by 30km visibility, no precipitation, scattered cloud at 5000 feet, and 10 knot winds. Low weather-related threat was represented by 80km visibility with no precipitation, clear skies, and a 5 knot wind.
Table 6.17.

Within-Subject Regression Coefficients for Weather-Related Threat and the Opportunity to Earn Income.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.93**</td>
<td>.08</td>
<td>.86</td>
</tr>
<tr>
<td>2</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>3</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>4</td>
<td>-.68*</td>
<td>.53*</td>
<td>.73</td>
</tr>
<tr>
<td>5</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>6</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>7</td>
<td>-.86**</td>
<td>.22</td>
<td>.79</td>
</tr>
<tr>
<td>8</td>
<td>-.47</td>
<td>.07</td>
<td>.23</td>
</tr>
<tr>
<td>9</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>10</td>
<td>-.78*</td>
<td>.32</td>
<td>.70</td>
</tr>
<tr>
<td>11</td>
<td>-.86**</td>
<td>.07</td>
<td>.74</td>
</tr>
<tr>
<td>12</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>13</td>
<td>-.91**</td>
<td>.06</td>
<td>.84</td>
</tr>
<tr>
<td>14</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>15</td>
<td>-.86**</td>
<td>.06</td>
<td>.74</td>
</tr>
<tr>
<td>16</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>17</td>
<td>-.89**</td>
<td>.06</td>
<td>.79</td>
</tr>
<tr>
<td>18</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>19</td>
<td>-.75*</td>
<td>.33</td>
<td>.67</td>
</tr>
<tr>
<td>20</td>
<td>-.89**</td>
<td>.06</td>
<td>.79</td>
</tr>
<tr>
<td>21</td>
<td>-.85**</td>
<td>.12</td>
<td>.74</td>
</tr>
<tr>
<td>22</td>
<td>-.90**</td>
<td>.06</td>
<td>.82</td>
</tr>
<tr>
<td>23</td>
<td>-.92**</td>
<td>-.19</td>
<td>.89</td>
</tr>
<tr>
<td>24</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>25</td>
<td>-.27</td>
<td>.47</td>
<td>.28</td>
</tr>
<tr>
<td>26</td>
<td>-.80**</td>
<td>.33</td>
<td>.74</td>
</tr>
<tr>
<td>27</td>
<td>-.82**</td>
<td>.30</td>
<td>.76</td>
</tr>
<tr>
<td>28</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>29</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>30</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>31</td>
<td>-.68*</td>
<td>.38</td>
<td>.60</td>
</tr>
<tr>
<td>32</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
</tbody>
</table>

*p<.05, **p<.01.

In all but two cases, the pilots' responses were significantly influenced by the level of threat in the situation. These within-subject regression coefficients ranged from -.93 to -.27, indicating that the pilots were less likely to go on the flight when the weather was severe. Meanwhile, within-subject regression coefficients for opportunity ranged from -.19 to .53, indicating that for the majority of the participants, the opportunity to earn income did not significantly influence their decision to take-off. For one participant, the within-
subject regression coefficient for opportunity was .53, \((t = 2.49, p = .047)\) and the within-subject regression coefficient for threat was -.68 \((t = -3.20, p = .018)\), indicating that both weather-related threat and the opportunity to earn income influenced the likelihood that the participant would go on the flight. The higher the opportunity to earn income, and the lower the weather-related threat, the more likely it was that the pilot would take-off. Two pilots had \(R^2\) values of less than 0.5, and these pilots were excluded from the following analyses.

There was no difference in the influence of opportunity or threat on the decision to take-off between the pilots who continued or diverted after the first weather change, or between the pilots who had continued, diverted, or crashed by the critical decision point. There was also no relationship between the number of hazardous events previously experienced and the influence of threat or opportunity on the decision to take-off.

*Career advancement/Time pressure:* Opportunity for career advancement was operationalised as the type of aircraft flown, with more advanced aircraft representing more opportunity for career gain. High, medium, and low opportunity for career advancement were represented by a Piper Seneca, Piper Arrow, and Piper Warrior, respectively. High, medium, and low threat of time pressure were represented having 5 minutes, 30 minutes, and one hour to do the pre-flight checks, respectively. The within-subject regression coefficients for scenarios with threat of time pressure and opportunity for career advancement are presented in Table 6.18.
Table 6.18.

*Within-Subject Regression Coefficients for Threat of Time Pressure and the Opportunity for Career Advancement.*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Threat</th>
<th>Opportunity</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-.43</td>
<td>-.43</td>
<td>.38</td>
</tr>
<tr>
<td>2</td>
<td>-.82**</td>
<td>-.31</td>
<td>.76</td>
</tr>
<tr>
<td>3</td>
<td>-.87**</td>
<td>.00</td>
<td>.75</td>
</tr>
<tr>
<td>4</td>
<td>-.55</td>
<td>-.55</td>
<td>.56</td>
</tr>
<tr>
<td>5</td>
<td>-.13</td>
<td>-.97**</td>
<td>.95</td>
</tr>
<tr>
<td>6</td>
<td>-.93**</td>
<td>-.15</td>
<td>.88</td>
</tr>
<tr>
<td>7</td>
<td>-.99**</td>
<td>.00</td>
<td>.98</td>
</tr>
<tr>
<td>8</td>
<td>-.15</td>
<td>-.44</td>
<td>.21</td>
</tr>
<tr>
<td>9</td>
<td>-.87**</td>
<td>-.26</td>
<td>.82</td>
</tr>
<tr>
<td>10</td>
<td>-.07</td>
<td>-.77*</td>
<td>.59</td>
</tr>
<tr>
<td>11</td>
<td>-.94**</td>
<td>.00</td>
<td>.87</td>
</tr>
<tr>
<td>12</td>
<td>-.77*</td>
<td>.28</td>
<td>.67</td>
</tr>
<tr>
<td>13</td>
<td>-.95**</td>
<td>-.14</td>
<td>.93</td>
</tr>
<tr>
<td>14</td>
<td>-.82**</td>
<td>-.27</td>
<td>.75</td>
</tr>
<tr>
<td>15</td>
<td>-.68*</td>
<td>-.38</td>
<td>.60</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>-.97**</td>
<td>.15</td>
<td>.95</td>
</tr>
<tr>
<td>18</td>
<td>-.73*</td>
<td>-.13</td>
<td>.56</td>
</tr>
<tr>
<td>19</td>
<td>-.95**</td>
<td>-.14</td>
<td>.92</td>
</tr>
<tr>
<td>20</td>
<td>-.41</td>
<td>-.41</td>
<td>.33</td>
</tr>
<tr>
<td>21</td>
<td>-.68</td>
<td>-.07</td>
<td>.47</td>
</tr>
<tr>
<td>22</td>
<td>-.14</td>
<td>-.72*</td>
<td>.54</td>
</tr>
<tr>
<td>23</td>
<td>-.98**</td>
<td>-.07</td>
<td>.97</td>
</tr>
<tr>
<td>24</td>
<td>-.65</td>
<td>-.28</td>
<td>.49</td>
</tr>
<tr>
<td>25</td>
<td>-.65</td>
<td>.24</td>
<td>.48</td>
</tr>
<tr>
<td>26</td>
<td>-.62</td>
<td>-.13</td>
<td>.39</td>
</tr>
<tr>
<td>27</td>
<td>-.90**</td>
<td>-.16</td>
<td>.83</td>
</tr>
<tr>
<td>28</td>
<td>-.72**</td>
<td>-.56*</td>
<td>.83</td>
</tr>
<tr>
<td>29</td>
<td>-.67**</td>
<td>-.67**</td>
<td>.86</td>
</tr>
<tr>
<td>30</td>
<td>-.89**</td>
<td>-.37*</td>
<td>.94</td>
</tr>
<tr>
<td>31</td>
<td>-.64*</td>
<td>-.46</td>
<td>.62</td>
</tr>
<tr>
<td>32</td>
<td>-.89**</td>
<td>-.06</td>
<td>.79</td>
</tr>
</tbody>
</table>

* p<.05, ** p<.01.

Participant 16 indicated that he would definitely go on all nine career/time pressure scenarios. Since there was no variation in this pilot's responses, a regression equation could not be calculated. Therefore, this pilot was excluded from the following analyses.

The within-subject regression coefficients for threat ranged from -.07 to -.98. While all of the within-subject regression coefficients were negative, for 11 out of the 31 participants, the level of threat in the situation did not significantly affect the likelihood of
the pilot going on the flight. For the other 20 participants, there was a negative relationship between the likelihood of going on the flight and the level of threat in the situation, where the greater the time pressure in the scenario, the less likely it was that the pilots would go on the flight. The within-subject regression coefficients for opportunity ranged from -.97 to .28. For 6 out of the 31 participants, there was a significant negative relationship between the likelihood of going on the flight and the level of opportunity in the situation. For these pilots, the greater the opportunity for career gain, the less likely it was that the pilots would go on the flight. For the other 15 pilots, there was no relationship between the level of opportunity in the situation and the likelihood of going on the flight. Seven pilots had a $R^2$ value of less than .50. These pilots were excluded from the subsequent analyses.

There was no difference in the influence of opportunity or threat on the decision to take-off between the pilots who continued or diverted after the first weather change. One of the participants who crashed the simulator (participant 26) was excluded from the analyses. Therefore, the influence of opportunity and threat on the likelihood of taking-off was compared only for the participants who continued or diverted by the critical decision point. A t-test, when equal variances were not assumed, showed that there was a marginally significant difference between the pilots who continued and diverted at the second weather change. The pilots who continued tended to have a higher mean within-subject regression coefficient for threat than the pilots who diverted. That is, the pilots who continued had a weaker relationship between the level of threat in the situation ($M = -.54$, $SD = .37$) and the likelihood of going on the flight than the pilots who diverted ($M = -.85$, $SD = .13$), $t(7.88) = 2.24$, $p = .056$, $d = 1.08$. The average within-subject regression coefficient (threat) for the diverters was in the 84th percentile of the within-subject regression coefficients (threat) for the continuers, with 55.4% non-overlap between the two distributions. The mean within-subjects coefficients for the threat of time pressure for the pilots who continued, diverted, or crashed by the second weather change is shown in Figure 6.10. There was no difference in the influence of opportunity on the decision to take-off and in-flight decision-making at the critical decision-point.
Figure 6.10. Mean within-subject regression coefficient for the threat of time pressure for the pilots who continue or divert by the second weather change.

Number of flights agreed to go on: One of the reasons why I used a 6-point Likert scale was to force the participants to make a go/no-go decision. It was of interest to assess whether there was any relationship between going on the hypothetical scenarios in the risk tolerance exercise and in-flight decision-making in the simulated flight. A rating of between 1 (definitely not) and 3 was coded as ‘no-go’. A rating of between 4 and 6 (definitely yes) was rated as ‘go’. The number of flights that the participants indicated that they would go on was calculated. The maximum was 36 flights. The pilots indicated that they would go a mean of 19.53 (54.25%) flights (range = 10 (27.8%) - 34 (94.4%), SD = 4.77 (13.24%)). There was no difference in the percentage of flights the pilots agreed to go on and the decision made after the first or second weather change.

The mean likelihood that each pilot would go on the flight was also calculated across all of the scenarios. This was represented as a number out of 6. The mean rating ranged from 2.28 to 5.42 (M = 3.61, SD = .647). The pilots who diverted after the first weather change gave a marginally significant lower average likelihood of going on the flights (M = 3.40, SD = .48) compared to the pilots who continued (M = 3.75, SD = .72), Z(32) = -1.71, p = .087, d = .57. The average Likert rating for the continuers was in the 69th percentile of
the Likert ratings for the diverters, with a 33.0% non-overlap between the two distributions. That is, the pilots who continued after the first weather change were marginally more likely to rate the likelihood of taking off as higher than the pilots who diverted. The average likelihood of going on the flight for the pilots who continued or diverted after the first weather change is shown in Figure 6.11.

![Figure 6.11](image)

**Figure 6.11.** The mean likelihood of going on the flights for the pilots who continued and diverted after the first weather change.

There were no differences in any of the variables when looking at the decision made by the critical decision point. The pilots who continued, diverted, or crashed did not differ in the percentage of flights the participants agreed to go on or the mean likelihood of going on the flight. There was no relationship between previous involvement in hazardous aeronautical events and the number of flights the participants agreed to go on or the mean likelihood of going on the flight.
Implicit Association Test

Participants completed two IATs measuring implicit risk perception and implicit anxiety felt towards adverse and fine weather. The calculation of the IAT effect followed recommendations by Greenwald, Nosek, and Banaji (2003), whereby implicit attitudes are measured using D-ASIS. D-ASIS was calculated as:

\[
\frac{\text{Mean Reaction Time Incompatible} - \text{Mean Reaction Time Compatible}}{\text{Standard Deviation of all Latencies}}
\]

D-ASIS includes the data from both the practice and test blocks of the combined tasks, while omitting trials with latencies greater than 10 seconds. This measure is recommended when the participants are required to correct their errors, and when latency is recorded until the participant does so. Therefore, the IAT procedure includes a built-in error penalty.

IAT-Anxiety: For IAT Anxiety the compatible block was IMC+AFRAID/VMC+UNAFRAID, and the incompatible block was IMC+UNAFRAID/VMC+AFRAID. Negative D-ASIS scores indicate that the participants were faster at responding when words meaning AFRAID were paired with pictures of IMC weather conditions, and words meaning UNAFRAID were paired with pictures of VMC weather conditions. Positive D-ASIS scores indicate that the participants were faster at responding when words meaning AFRAID were paired with pictures of VMC weather conditions, and words meaning UNAFRAID were paired with pictures of IMC weather conditions. Figure 6.12 shows the distribution of D-ASIS Anxiety scores.
Only 2 participants had positive D-ASIS scores, indicating that the vast majority of participants associated bad weather with feeling afraid. The D-ASIS scores ranged from -1.15 to .144 (M = -.59, SD = .33). To assess the relationship between implicit anxiety and the decision made on the first weather change, a 2*2 (speed*decision) ANOVA was conducted, with D-ASIS as the dependent variable. There were no main effects or interactions.

When the participants were divided into three groups (continued, diverted, or crashed by the critical decision point), and a 2*3 (speed*decision) ANOVA was conducted there was a significant difference between the three groups of pilots, $F(2,27) = 4.35, p = .023, \eta_p^2 = .244$, where the participants who crashed had a weaker IAT-Anxiety effect ($M = .04, SD = .11$) compared to those who continued ($M = -.69, SD = .27$), (Mean Difference = .65, $SE = .23, p = .024, d = 3.11$). There was also a marginally significant difference between the pilots who crashed and diverted, where the participants who crashed had a weaker IAT-Anxiety effect compared to the pilots who diverted (Mean Difference = .54, $SE = .23, p = .062, d = 1.6875$). The difference in implicit anxiety between the pilots who continued, crashed, and diverted by the second weather change is shown in Figure 6.13.
Figure 6.13. Mean D-ASIS anxiety scores for the pilots who continued, diverted, and crashed by the second weather change.

The pilots who crashed were excluded and a 2*2 (speed*decision made by the second weather change) between-subjects ANOVA was conducted. There were no main effects or interaction.

IAT-Risky: For the IAT risky, the compatible block was IMC+RISKY/VMC+SAFE, and the incompatible block was IMC+SAFE/VMC+RISKY. Negative D-ASIS scores for the Risky IAT indicated that the participants were faster at responding during the RISKY+IMC and SAFE+VMC condition compared to the RISKY+VMC and SAFE+IMC condition. Figure 6.14 shows the distribution of D-ASIS Risky scores.
All of the participants had a negative D-ASIS score, indicating that the participants associated adverse weather with risk, and good weather with safety. The D-ASIS scores ranged from -1.18 to -.140 ($M = -.68$, $SD = .26$).

A 2*2 (speed condition*decision at first weather change) between subjects ANOVA with D-ASIS Risky scores as the dependent variable was conducted. There was no significant difference between the implicit attitudes for the pilots who continued and diverted after the first weather change. There was no interaction with speed and decision; however, there was a main effect of speed. The pilots in the double speed condition had a lower (more negative) IAT effect ($M = -.79$, $SD = .24$), than the pilots in the normal speed condition ($M = -.57$, $SD = .25$), $F(1,32) = .69$, $p = .014$, $\eta^2_p = .198$, $d = .91$, suggesting that the pilots who were flying at double speed had a stronger implicit attitude towards adverse weather being risky than the pilots in the normal speed condition. The pilots who were flying at double speed had implicit attitudes that were in the 82$^{nd}$ percentile of the participants who were flying at normal speed, with 51.6% non-overlap in the distributions.

A 2*3 (speed*decision by the second weather change) between-subjects ANOVA was conducted. The pilots who had diverted, continued, or crashed by the decision point did not differ in implicit risk perception. There was a significant effect of speed, $F(1,27) =$
4.67, \( p = .04, \eta_p^2 = .147 \). The pilots in the double-speed condition implicitly perceived more risk in adverse weather than the pilots in the normal speed condition. The mean implicit risk perception of the pilots who continued, diverted, or crashed, for both speed conditions are presented in Figure 6.15.

![Graph showing mean D-ASIS risky scores](image)

**Figure 6.15.** Mean D-ASIS risky scores for the pilots who continued, diverted, and crashed by the critical decision point.

The pilots who crashed were excluded and a 2*2 (speed*decision at second weather change) between-subjects ANOVA was conducted. There was a main effect of speed, \( F(1,26) = 4.57, \ p = .042, \ \eta_p^2 = .149, \ d = .84 \). The pilots in the double speed condition implicitly perceived more risk in adverse weather (\( M = -.78, SD = .25 \)) compared to the pilots in the normal speed condition (\( M = -.57, SD = .25 \)). The pilots who were flying at double speed had implicit attitudes that were in the 79\textsuperscript{th} percentile of the participants who were flying at normal speed, with 47.4\% non-overlap in the distributions. There was no main effect of decision made or interaction between speed and decision.

It was of interest to see if the number of hazardous events which the participants were previously involved in was related to the IAT scores. After the outlier was removed, the HES scores were normally distributed. There was a significant relationship between the number of hazardous events which the participants were involved in and the D-ASIS score
for Anxiety IAT, $r(29) = .50, p = .006$. The positive relationship suggested that the more hazardous events which the participants were previously involved in, the higher the D-ASIS score. Participants with a higher IAT effect had a weaker association between IMC + Afraid and VMC + Unafraid. Meanwhile, there was no relationship between Risky-IAT scores and the number of hazardous events which the participants were previously involved in. There was also a relationship between the pilot’s involvement in weather-related incidents and implicit anxiety, $r(29) = .393, p = .035$. Again, this suggests that the more hazardous events the pilots had been involved in, the less implicitly anxious the pilots were.

Participants were asked if bad weather had ever forced them to land at an airport other than their original destination. A one-way ANOVA was conducted to assess whether the pilots who had experienced a forced landing differed in the IAT scores. The participants who had experienced a forced landing had a significantly weaker association between adverse weather and feeling anxious ($M = -.47, SD = .24$) than the participants who had not experienced a forced landing ($M = -.74, SD = .29$), $F(1,30) = 5.95, p = .021$, $\eta^2_p = .166, d = 0.99$. This suggests that the pilots who had experienced a forced landing were less implicitly afraid of bad weather, compared to the pilots who had not experienced a forced landing. The mean D-ASIS score for the pilots who had experienced a forced landing was in the 82nd percentile of the D-ASIS scores for the pilots who had not experienced a forced landing, with 41.6% non-overlap in the two distributions. This is shown in Figure 6.16.
A series of Spearman’s rank order correlations were performed to assess relationship between implicit attitudes and experience. Implicit risk perception was not associated with age or any measure of experience, other than the number of hours flown as pilot-in-command over the last 90-days. The fewer hours the pilot had flown PIC over the last 90-days, the stronger the association between adverse weather and risk. Implicit anxiety was related to age, hours flown PIC over the last 90-days, and hours flown cross-country in the last 90-days. The older the pilot, and the fewer hours flown PIC and cross-country in the last 90-days, the stronger the association between bad weather and feeling afraid. This is shown in Table 6.19.
Table 6.19.
Spearman’s Rank Order Correlations between Measures of Experience and Age, and Implicit Risk Perception and Anxiety.

<table>
<thead>
<tr>
<th>Measures</th>
<th>IAT Risky</th>
<th>IAT Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.093</td>
<td>-.459**</td>
</tr>
<tr>
<td>Total Hours</td>
<td>-.138</td>
<td>-.115</td>
</tr>
<tr>
<td>PIC</td>
<td>-.047</td>
<td>-.094</td>
</tr>
<tr>
<td>PIC Cross-country</td>
<td>.012</td>
<td>-.047</td>
</tr>
<tr>
<td>PIC last 90-days</td>
<td>.370*</td>
<td>.537**</td>
</tr>
<tr>
<td>Cross-country last 90-days</td>
<td>.314</td>
<td>.491**</td>
</tr>
<tr>
<td>Length of certification</td>
<td>-.302</td>
<td>-.271</td>
</tr>
</tbody>
</table>

*. Correlation is significant at the .05 level (2-tailed)

**. Correlation is significant at the .001 level (2-tailed)

Implicit risk perception did not differ according to level of certification or whether the pilot held an instrument rating. Likewise, implicit anxiety did not differ according to level of certification. However, the pilots who held an instructor rating had a marginally weaker implicit association between bad weather and feeling afraid ($M = -.40$, $SD = .33$), compared to the non-instructors ($M = -.65$, $SD = .31$), $F(1,30) = 3.85$, $p = .059$, $\eta^2_p = .114$. This is shown in Figure 6.17.
Control of Aircraft

It is possible that the pilots who flew the simulation at double-speed found it more difficult to control the simulated aircraft compared to the pilots who flew the simulation at normal-speed. To assess this and the relationship between difficulty controlling the aircraft and in-flight decision-making, the difficulty each pilot had flying was measured. This was measured by calculating the root mean square error (RMSE) for heading, altitude, and airspeed. RMSE is a measurement of how accurately the pilots were flying. It was measured by calculating the individual errors (which was how much the pilot deviated from the instructed heading, altitude, and speed) for a given distance (in this case, 5 nm). Each error was then squared, added together, divided by the total number of errors, and the square root was calculated. This 5nm period occurred before the first weather change. Two pilots diverted during this 5nm period, and one participant’s data was lost. These three participants were excluded from the following analyses.

If the residuals were normally distributed, or became normally distributed following transformation, and the variances were not significantly different, a 2*2 (decision*speed, for the first weather change), or a 3*2 (decision*speed, for the second weather change) ANOVA was calculated, assessing the main effects of decision and speed on difficulty

Figure 6.17. Mean D-ASIS Anxiety for the pilots with and without a flight instructor rating.
flying the simulator, and the interaction between decision and speed. If the residuals were not normally distributed, non-parametric equivalent of one-way ANOVAs were conducted, assessing the main effects of decision and speed on difficulty flying the simulator. With non-parametric data, the interaction between decision and speed could not be assessed.

*Altitude RMSE:* The errors were calculated as the deviation from 2,000 feet, which was the altitude they were recommended to fly at. On average, the pilots deviated by 345.70 feet (range = 38.37 - 1,113.89, SD = 304.16). There were no significant differences between the pilots who flew at normal and double speed. The pilots who continued or diverted after the first weather change did not differ in altitude error. There was no interaction between speed and decision made. When the decisions made by the critical decision point were examined, the pilots who crashed did not find it more difficult to maintain 2,000 feet (at least during this 5nm period). However, the pilots who continued found it marginally more difficult to maintain 2,000 feet compared to the pilots who diverted (continuers: $M_t = 465.34$, $SD_t = 346.87$, $M_d = 20.12$, $SD_d = 8.12$; diverters: $M_t = 249.63$, $SD_t = 239.47$, $M_d = 14.30$, $SD_d = 6.95$), $F(1,26) = 4.00$, $p = .057$, $\eta_p^2 = .148$, $d = .77$. The mean error in altitude for the pilots who continued was in the 76th percentile of the errors in altitude for the pilots who diverted, with 43% non-overlap between the distributions. The relationship between decision made and altitude error is shown in Figure 6.19. There was no interaction between speed and decision made the critical decision point.
Figure 6.18. Mean altitude error for the pilots who continued, diverted, and crashed by the critical decision point.

**Heading RMSE:** Errors were calculated as the deviation from the assigned heading of 150 degrees. On average, the pilots deviated from the instructed heading by 12.60 degrees (range = .001 – 98.65, SD = 20.16). The residuals were normally distributed, but the variances were unequal. There was a significant difference between the pilots in the normal and double speed conditions, $t(13.184) = 2.217, p = .045, d = .60$. The pilots in the double speed condition found it more difficult to maintain the heading than the participants in the normal speed condition (normal: $M = 4.38, SD = 2.36$, double: $M = 15.856, SD = 26.95$). The mean heading error for the pilots who were in the double-speed condition was in the 73rd percentile of the errors for the pilots who were in the normal speed condition, with 38.2% non-overlap between the two distributions. The mean heading error for the pilots in the normal and double speed conditions is shown in Figure 6.19.
Figure 6.19. Mean heading error for the pilots in the normal and double-speed conditions. There was no difference in the mean heading error between the pilots who continued or diverted after the first weather change or the critical decision point.

Airspeed RMSE: Airspeed RMSE was calculated as deviation from 110 knots. On average, the pilots deviated 8.43 knots (range = 1.92 – 21.28, SD = 5.13). There was no difference between the pilots in the normal and double speed conditions, or between the pilots who continued and diverted after the first weather change or the critical decision point.

Personal Minimums
Pilots were asked to indicate the minimum visibility and cloud ceiling they would accept before taking off on a local or cross-country flight. Over half of the pilots would not take-off if visibility was less than 5 km on a local flight and 10 km on a cross-country flight. Meanwhile, over half of the pilots would not take-off if the cloud ceiling was less than 2,000 feet on a local flight and 3,000 feet on a cross-country flight. Interestingly, all pilots took-off on the simulated flight, even though the cloud ceiling was forecast as 2,000 feet. There was no difference in the mean personal minimums between the pilots who continued and diverted (or crashed) at either decision point.
Usual Aeronautical Practices

The pilots rated the percentage of time they participated in nine aeronautical behaviours. Of the behaviours, seven were phrased so that agreement with the statement indicated safe behaviour. The other two items were reversed scored, so that 100% indicated that the pilots participated in the safe behaviour all of the time. This was averaged across all items for each pilot to indicate how often the pilot was involved in safe practices. On average, the pilots indicated they were involved in safe practices 74.70% of the time ($range = 31.67 - 100.00$, $SD = 16.33$). The mean percentage of time that the pilots participated in safe behaviour did not differ between the pilots who continued or diverted (or crashed) by the first weather change or the critical decision point.

Logistic Regression

A logistic regression was calculated to assess the influence of the measured variables on in-flight decision-making. According to Tabachnick and Fidell (1996), much like linear regression, logistic regression measures the influence of predictor variables on an outcome. Unlike linear regression, the outcome does not need to be a continuous variable, the predictors can be a mix of continuous and categorical variables, and there are no underlying assumptions about the distribution of any of the variables. Therefore, logistical regression is an ideal way to predict which parameters influence the likelihood of the pilots continuing and diverting.

For both the first decision point (the first weather change) and the critical decision point (second weather change), the following variables were entered into the equation:

- PIC Cross-country
- Within-subject coefficient for opportunity
- Within-subject coefficient for threat
- IAT risky
- IAT anxiety
- CWS

The Homer and Lemeshow test assesses whether the model predicts pilot decision-making. If the significance of the test is less than .05, then the model failed to adequately fit the data. With the above variables in the model, the model adequately predicted pilot-
decision making at the first weather change, \(\chi^2(8) = 4.38, p = .821\). Table 6.20 presents the results of the logistic regression for decision-making at the first weather change.

Table 6.20.
Logistic Regression for First Weather Change

<table>
<thead>
<tr>
<th>Variable</th>
<th>Wald</th>
<th>df</th>
<th>p</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opportunity</td>
<td>.23</td>
<td>1</td>
<td>.631</td>
<td>.05</td>
</tr>
<tr>
<td>Threat</td>
<td>3.89</td>
<td>1</td>
<td>.049</td>
<td>.00</td>
</tr>
<tr>
<td>IAT Risky</td>
<td>.66</td>
<td>1</td>
<td>.416</td>
<td>.16</td>
</tr>
<tr>
<td>IAT Anxiety</td>
<td>4.73</td>
<td>1</td>
<td>.028</td>
<td>113.19</td>
</tr>
<tr>
<td>CWS</td>
<td>3.20</td>
<td>1</td>
<td>.074</td>
<td>.81</td>
</tr>
<tr>
<td>PIC cross-country</td>
<td>1.02</td>
<td>1</td>
<td>.313</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Two variables significantly predicted pilot decision-making: D-ASIS Anxiety, Wald(1) = 4.73, p = .028; and threat, Wald(1) = 3.89, p = .049. The pilots' CWS scores marginally predicted decision-making at the first weather change, Wald(1) = 3.20, p = .074. The expected B score measures the direction of the relationship. When this value is less than one, an increase in the variable decreases the odds of the event’s occurrence. When this value is greater than one, increasing values increases the odds of the event’s occurrence. The decision to continue was coded as 0, while the decision to divert was coded as 1. The expected B score for IAT anxiety was 113.19. Increasing values of D-ASIS Anxiety relates to lesser association between feeling anxious and bad weather. Therefore, the participants who diverted were less implicitly afraid of bad weather, compared to those who continued. The expected B value for threat was .001. Since higher threat scores indicate that the participants were less influenced by threat, the participants who diverted were more influenced by threat. The expected value for CWS was .81. The higher the CWS score, the better the pilot was at perceiving weather-related aeronautical risks. Therefore, pilots who continued were better at perceiving weather-related aeronautical risks. Table 6.21 presents the results of the logistic regression for decision-making at critical decision point.
The two participants who crashed were excluded from the analysis, as the logistic regression was unable to be calculated with only two participants in one group. With the variables in the model, the model adequately predicted pilot-decision making at the critical decision point. With the variables in the model, the model adequately predicted pilot-decision making at the critical decision point. 

**Discussion**

Approximately 25% of pilots admit to flying VFR-into-IMC (Hunter, 1995; O’Hare & Chalmers, 1999), which can lead to spatial disorientation (Nagel, 1988). The aim of the present study was to examine the role of risk management and implicit associations in flying into adverse weather. In the present experiment, pilots completed a simulated flight where the weather deteriorated below VFR-minimums and four computer based-questionnaires, measuring expertise in aeronautical weather-related risk perception, risk tolerance, and implicit anxiety and risk perception towards adverse weather.

During the simulated flight, 13 pilots diverted before or soon after the first weather change. These pilots revised their plan to fly to Tillamook after the weather deteriorated to 1,500 feet cloud ceiling. Meanwhile, five pilots continued beyond the first weather change, but revised their plan before the critical decision-point. Twelve pilots were still flying after...
the critical decision point. These pilots had not made a decision to divert or turn back to Quillayute, even though the cloud ceiling had lowered to 800-feet. After the weather deteriorated to 1,500 feet cloud ceilings, most of the pilots changed their altitude to 1,400 feet. The pilots who continued beyond the critical decision point either continued flying through the cloud at 1,400 feet, or descended below the cloud and flew approximately 500 feet above the ground. After the second weather change, the pilots who diverted after the first weather change, but before the critical decision point usually maintained their altitude at 1,400 feet, and did a 180° turn with the assumption that they would soon get out of the clouds. After exiting the clouds, the pilots would usually divert to the nearest airport.

Finally, two pilots ‘crashed’ the simulator. For one pilot the ‘crash’ occurred between the first and second weather changes. For the other pilot, the ‘crash’ occurred after the second weather change, but before the critical decision point. Both of these pilots were having difficulty controlling the simulator before crashing.

Risk perception: One reason why pilots may fly VFR-into-IMC may be because they do not realise the risk involved in doing so, that is they have poor risk perception. The CWS procedure measured the level of expertise in judging aeronautical weather-related risks. The higher the CWS score, the more consistently the participants discriminated between different weather-related scenarios. Since past research has shown that risk perception is related to weather-related decision-making (Goh & Wiegmann, 2001a, 2001b, 2002a; Hunter, 2002, 2005, 2006; O’Hare & Wiegmann, 2003), I hypothesised that expertise in weather-related aeronautical risk perception would be related to in-flight decision-making. Specifically, I expected that the pilots who continued after the critical decision point will have had lower CWS scores compared to the pilots who diverted before the critical decision point.

When I conducted an ANOVA, I found no difference between the CWS scores of the pilots who continued or diverted at either decision point. This suggests that expertise in weather-related aeronautical risk perception is not related to in-flight decision-making. However, when a logistic regression analysis was conducted, I found a marginally significant relationship between in-flight decision-making and CWS scores. The pilots who continued beyond the first weather change had higher CWS scores than the pilots who diverted, suggesting that the pilots who continued were more expert at perceiving aeronautical weather-related risks compared to those who diverted. While this is contrary
to the prediction, it may be because amongst those who continued after the initial weather change were pilots who diverted before the critical decision point. These pilots may have initially recognised that the weather was deteriorating, and that it was becoming riskier, but they may have felt that they could cope with 1,500 feet cloud ceiling. When a logistic regression was conducted on the relationship between risk management, implicit associations, and experience on in-flight decision-making at the critical decision point, the continuers and diverters did not differ in the CWS scores. Overall, the results suggest a lack of relationship between expertise in aeronautical weather-related risk perception and in-flight decision-making.

There are two explanations for the lack of relationship between the CWS scores and in-flight decision-making: the task may not adequately measure aeronautical risk perception, or it may measure aeronautical risk perception, but risk perception may not be related to in-flight WRDM. Both of these will be discussed in turn.

It is possible that this task is not related to aeronautical risk perception. If the CWS procedure was a valid measure of risk perception, it follows that the CWS scores should have been related to experience. That is, the pilots with the most experience in WRDM should have had the highest CWS scores. Some indices of experience were related to CWS scores. The certification of the pilot was related to the CWS score, as the pilots with a CPL had a higher mean CWS score compared to the pilots with a PPL. Also, the pilots with a flight instructor rating had a higher CWS scores compared to the pilots without a flight instructor rating. This suggests that, for these two indices of experience, the pilots with the most flight experience were the most expert at perceiving aeronautical weather-related risks.

If the CWS task is a valid measure of risk perception, it follows that there should be a relationship between the pilots’ involvement in past aeronautical events, measured using Hunter’s (1995) HES, and the CWS scores. Hunter (2002, 2005, 2006) found a relationship between risk perception and involvement in hazardous events. However, in the present experiment, there was no relationship between expertise in aeronautical weather-related risk perception and involvement in past hazardous aeronautical events. This lack of relationship may have occurred because the HES measured involvement in hazardous aeronautical events in general, whilst the CWS measured expertise in aeronautical weather-related risk perception in particular. Since it is difficult to find a relationship between an attitude and behaviour when they are measured at different levels of specificity (Ajzen &
Fishbein, 1997), I also calculated involvement in weather-related aeronautical events. There was no relationship between involvement in weather-related aeronautical events and the CWS scores. Due to the lack of relationship between the CWS scores and involvement in hazardous aeronautical events, it may be that this measure does not adequately measure risk perception.

The second possible reason for the lack of relationship between in-flight decision-making in the simulated flight and the CWS scores is that risk perception is not related to WRDM. Previous studies have found mixed results regarding the relationship between risk perception and weather-related decision-making in a simulated flight. Goh and Wiegmann (2001a) and O'Hare and Wiegmann found a relationship between risk perception and flying VFR-into-IMC during a simulated flight, although Goh and Wiegmann’s result did not reach statistical significance. Meanwhile, O'Hare (1990) and Molesworth et al. (2006) found no relationship between risk perception and pre-flight and in-flight decision-making, respectively.

In the present study, I found no relationship between expertise in aeronautical weather-related risk perception and in-flight decision-making. However, I did find a relationship between risk perception and in-flight decision-making, when risk perception was measured as the average risk ratings given for the hypothetical scenarios. This measure of risk perception is analogous to that devised by Hunter (2002, 2005, 2006). In the present study, there was a trend for the pilots who continued after the first weather change to give lower ratings of risk in the hypothetical scenarios compared to the pilots who diverted. This difference was significant when the ratings were compared for the pilots who diverted, continued, and crashed by the critical decision point.

The pilots who diverted before the critical decision point gave higher risk ratings, compared to the pilots who crashed. This suggests a relationship between risk perception and in-flight decision-making, and further highlights that one reason why pilots may fly VFR-into-IMC is because they do not recognise the risk involved in flying into adverse weather. Further support for the relationship between risk perception and risk-taking was shown with the marginally significant relationship between risk judgments given in the second replication and involvement in hazardous events. This supports Hunter (2002, 2005, 2006) who found a similar relationship between risk perception and responses on the HES.
However, a problem with measuring risk perception in this manner is that it is difficult to know what the appropriate level of risk is for a given scenario. That is, it is difficult to assign a correct rating out of 100 for each scenario. This is partially due to the compensatory method of pilot decision-making (e.g., Driskill et al., 1995). Also, a pilot may give lower ratings of risk not because he or she perceives less risk in the situation than someone else, but because he or she has more flight experience, and may be able to better handle an objectively risky situation compared to the average pilot. There was either no relationship between average risk ratings and experience, or the more experienced pilots perceived more risk in the hypothetical scenarios. While this suggests that experienced and inexperienced pilots did not differ in risk perception, it is difficult to predict what risk judgments are appropriate for a given pilot.

While there may be inherent problems with assessing risk perception in this manner, the average risk rating for a set of hypothetical scenarios did relate to in-flight decision-making in a simulated flight. Therefore, this may be a better measure of risk perception than the CWS, which measured expertise in aeronautical weather-related risk perception.

Risk tolerance: Along with poor risk perception, pilots may fly into adverse weather because they are willing to accept high levels of weather-related risks (Hunter, 2002, O’Hare, 1990; Platenius & Wilde 1989). Risk tolerance was measured by presenting pilots with flight scenarios varying in the degree of opportunity for gain and threat of loss. Participants indicated the likelihood that they would go on each flight. Multiple regression equations were calculated, and the within-subject regression coefficients indicated the influence of opportunity and threat on the pilots’ decision to take-off.

Across all of the scenarios, the majority of pilots were influenced by the level of threat in the scenarios but not the level of opportunity. Following Lopes’s (1987) terminology, this indicates that the pilots were risk averse, as they were only willing to go on the flight if the threat of loss was low, regardless of what there was to gain. One pilot was influenced by both the level of opportunity and threat in the scenarios. This pilot was more likely to take-off on the hypothetical flight if the level of threat was low than if the level of threat was high. This pilot was also more likely to take-off on the hypothetical flight if the level of opportunity was low compared to when the level of opportunity was high. This indicates that this pilot was risk averse, as he would only take-off if the threat of
loss was low. The negative relationship between the level of opportunity and the likelihood of going on the flight suggests that this pilot viewed at least some of the opportunities as threats. Examination of his within-subject regression coefficients for the different facets of threat and opportunity reveals that this negative relationship may have arisen from the influence of career advancement on the likelihood to take-off. His within-subject regression coefficient for the opportunity for career advancement was -.97, while his within-subject regression coefficients for the other facets of opportunity were .00, suggesting that this pilot perceived opportunity for career advancement as a threat. This will be discussed later in this chapter.

I predicted that the pilots who continued flying beyond the critical decision point would be more risk tolerant than the pilots who diverted. For the pilots who continued beyond the first weather change, there was a weaker relationship between overall threat of loss and the likelihood of going on the flight, compared to the pilots who diverted. This suggests that the pilots who continued beyond the first weather change were less risk averse compared to the pilots who diverted. This finding was replicated in the logistic regression. However, as previously discussed, amongst the continuers were pilots who continued after the first weather change, but had diverted by the critical decision point. Subsequently, the important comparison was between the pilots who continued and diverted beyond the critical decision point. There was a marginally significant difference between these two groups of pilots. The pilots who continued beyond the critical decision point tended to be less influenced by the level of threat in situations, compared to the pilots who diverted. This suggests that the pilots who continued beyond the second weather change were less risk averse compared to the pilots who diverted. However, this relationship was not found in the logistic regression. Nonetheless, there seemed to be some relationship between being risk averse and in-flight decision-making.

There was no relationship between the influence of opportunity on the likelihood of going on the flight and in-flight decision-making. However, as discussed in chapter four, this lack of relationship may have resulted from the low variation in the within-subject regression coefficients for opportunity. There was also no relationship between involvement in hazardous aeronautical events and the influence of threat and opportunity on the decision to take-off.

Across all facets of opportunity and threat, most of the participants were influenced by the level of threat in the situation but not opportunity. This indicates that these pilots
were risk averse, as the pilots were more willing to go on the flight if the threat of loss was low, regardless of what there was to gain. When the opportunity was to earn income and the threat was weather-related, one pilot showed risk tolerant tendencies. The larger the potential profit, the more likely it was that this pilot would take-off. The pilot was more willing to accept worsening weather if the reward was great enough.

When the opportunity was for career gain, and the threat was time pressure, six pilots showed a relationship between the level of opportunity and the likelihood of going on the flight. For these six pilots, the more opportunity for career gain, the less likely it was that the pilots would take-off. This suggests that these pilots treated the opportunity for career advancement as a threat, and that they may have perceived more risk in flying in more advanced aircraft. This would account for the negative relationship between opportunity and the likelihood of going on the flights. As discussed in chapter four, while I took care to ensure that the levels of threat and opportunity were orthogonal, by specifying that a flight instructor was present on each flight, it is possible that these pilots ignored this information.

The only facets of threat and opportunity which were related to in-flight decision-making were the threat within the pilot and the threat of time pressure. The pilots who diverted before the critical decision point were more influenced by pilot related threat and the threat of time pressure compared to the pilots who continued.

Overall, there was some relationship between risk tolerance and in-flight decision-making. The pilots who continued were less risk averse than the pilots who diverted. Since there was no relationship between the within-subject coefficients for opportunity, I cannot conclude that the pilots who continued were more risk tolerant compared to those who diverted.

One limitation was that I measured pre-flight risk tolerance, but in-flight decision-making. That is, for the risk tolerance measure, the pilots responded according to whether they would take-off, while in the simulated flight I measured whether the pilots continued flying into adverse weather. Future studies should keep the type of decision-making consistent. Either a new tolerance measure should be created, which measures risk tolerance using in-flight decision-making problems, or the present measure of risk tolerance should be validated using measures of pre-flight decision-making.
The second aim of this study was to examine the relationship between implicit anxiety and implicit risk perception and in-flight decision-making. The pilots’ implicit associations were measured using two IATs: anxiety IAT and risky IAT.

All but two pilots had a negative anxiety IAT effect. That is, the pilots’ reaction time was slower in the incompatible condition (IMC+UNAFRAID/VMC+AFRAID) compared to the compatible condition (IMC+AFRAID/VMC+UNAFRAID). According to Greenwald et al. (1998), this indicates that these pilots had an implicit association between IMC weather and feeling afraid, and VMC weather and feeling unafraid. The remaining two pilots had an IAT effect of 0.04 and 0.12, indicating a very weak association between IMC weather and not feeling afraid, and VMC weather and feeling afraid.

I predicted that the pilots who diverted would have a stronger association between adverse weather and feeling afraid, compared to the pilots who continued. If shown, this would suggest a relationship between implicit anxiety and in-flight decision-making. However, the results did not show a significant difference in implicit anxiety between the pilots who continued and diverted. When a logistic regression was performed, implicit anxiety significantly predicted whether the pilots continued or diverted after the first weather change. Meanwhile, a logistic regression showed that implicit anxiety marginally predicted whether the pilots continued or diverted by the critical decision point. At both decision points, the result was opposite of that predicted. The pilots who diverted had a weaker association between bad weather and feeling afraid compared to the pilots who continued, suggesting that the pilots who diverted were less implicitly afraid of adverse weather than the pilots who continued. Meanwhile, the two pilots who crashed had a weaker association between feeling afraid and adverse weather, compared to the pilots who diverted or continued. This suggests that the pilots who crashed were less implicitly afraid of bad weather compared to the other pilots.

The finding that the pilots who continued were more implicitly anxious towards adverse weather may be explained by the difference in experience during the simulated flight. That is, the pilots who continued flying may have had a more negative experience during the simulated flight, as they would have encountered more severe weather. However, according to Egloff and Schmukle (2002) and Friese et al. (2006) conscious distortion should not have affected implicit attitudes. Furthermore, if the negative experiences of the flight simulation led to the pilots who continued feeling more implicitly afraid of adverse weather, the pilots who crashed should have the strongest association.
between adverse weather and feeling afraid, as they presumably had the most negative experience in the flight simulator. However, the pilots who crashed had a significantly weaker association between adverse weather and feeling afraid, compared to the other pilots.

The finding that the pilots who flew into adverse weather were not less implicitly anxious towards adverse weather suggests that implicit anxiety cannot account for why pilots fly VFR-into-IMC. However, this contradicts past findings by O’Hare and Wiegmann (2003). These researchers found that the pilots who continued into adverse weather during a simulated flight did not experience as marked an increase in heart rate when the weather deteriorated, as compared to the pilots who diverted before the critical decision point. O’Hare and Wiegmann’s (2003) research suggested a relationship between implicit anxiety and in-flight decision-making.

I was also interested in the relationship between implicit attitudes and involvement in hazardous aeronautical events. There was a significant relationship between previous involvement in hazardous aeronautical events and implicit anxiety. The fewer hazardous aeronautical events the pilots had been involved in, the stronger the association between feeling anxious and adverse weather. This suggests that the fewer hazardous events the pilots had been involved in, the more implicit anxiety the pilots felt towards adverse weather. However, the HES measured involvement in hazardous aeronautical events in general, while the IAT measured implicit anxiety towards adverse weather. Therefore, it would be more important and interesting if implicit anxiety related to the pilots’ involvement in weather-related aeronautical events. The same pattern was found when assessing the pilots’ involvement in weather-related hazardous aeronautical events. The pilots who had been involved in more aeronautical weather-related incidents had weaker implicit associations between adverse weather and feeling anxious, suggesting that one of the reasons why pilots are involved in such events may be because these pilots are less implicitly afraid of flying in hazardous weather.

All pilots had a negative risky IAT effect. That is, the pilots’ reaction time was slower in the incompatible condition (IMC+SAFE/VMC+RISKY) than the compatible condition (IMC+RISKY/VMC+SAFE). According to Greenwald et al. (1998), this indicates that the pilots implicitly associated adverse weather with high risk and fine weather with safety, suggesting that pilots implicitly perceived that flying in adverse weather involved more risk than flying in fine weather.
I predicted that the pilots who diverted would have had a stronger association between adverse weather and high risk, compared to the pilots who continued. If this was shown, it would suggest that implicit risk perception was related to in-flight decision-making. However, the results did not show a relationship between implicit risk perception and in-flight decision-making in the simulated flight. The logistic regression revealed a marginally significant difference between the pilots who continued and diverted by the critical decision point. The pilots who continued tended to have a weaker association between adverse weather and risk, compared to the pilots who diverted. This suggests a relationship between in-flight decision-making and implicit risk perception. That is, one of the reasons why pilots continued flying may have been because they implicitly perceived less risk in flying in adverse weather compared to the pilots who diverted. Pilots may be more likely to fly VFR-into-IMC if they perceive less risk in flying in adverse weather. This supports Pauley (2003), who found that pilots’ pre-flight decision-making was related to implicit risk perception.

According to Adams and Ericsson (2000), experienced pilots make decisions on an intuitive level. Therefore, I should have found a relationship between implicit attitudes and experience. There was a relationship between some indices of experience and implicit risk perception and anxiety. The fewer hours the pilot had flown pilot-in-command over the previous 90-days, the stronger the associations between adverse weather and risk. The fewer hours the pilot had flown pilot-in-command and cross-country over the previous 90-days, the stronger the associations between adverse weather and feeling anxious. However, there was no relationship between implicit attitudes and total hours, hours pilot-in-command, and hours cross-country over the previous 90-days. Therefore, the results did not show that the more experienced pilots differed in implicit attitudes. Rather, the pilots with less recent flight experience implicitly perceived less risk, and felt less anxious, towards adverse weather.

The third goal of the present study was to assess the effect of time pressure on in-flight decision-making. Time pressure was operationalised as simulation speed. Half of the pilots completed the task with the simulator set at double speed. For these pilots, the weather deteriorated at twice the rate compared to when the simulator was set at normal speed. Following Kersthoft’s (1994) finding that time pressure led to increased risk-taking, I predicted that the pilots in the high time pressure condition would be more likely to continue flying into adverse weather compared to the pilots in the low time pressure
condition. However, I found that the proportion of pilots who continued and diverted did not differ between the two speed conditions, suggesting that inducing time pressure did not lead to increased risk-taking.

There are two possible confounding variables which may account for the lack of a relationship between time pressure and risk-taking: difficulty flying and sunk costs. Each of these will be discussed in turn.

The more difficulty the pilot had controlling the flight simulator, the more attention the pilot would have to have given towards the primary flight tasks such as controlling the simulator, and the less attention the pilot would be able to give to cognitive tasks, such as situational awareness (e.g., noticing that the clouds were lowering in the distance), risk perception (e.g., understanding the lowering clouds represented a greater risk), and subsequent decision-making (e.g., deciding that he or she should have turned back or diverted to another airport). Through observation of the pilots’ flight control under the two speed conditions, it appeared that the pilots who were under high time pressure found it more difficult to fly than the pilots in the low time pressure condition. When the simulator was set at double speed, any movement of the flight controls will have resulted in a rate of change double the magnitude compared when the simulator was set at normal speed. Therefore, I expected that the pilots who flew at double speed would have found it more difficult to control the simulator compared to the pilots who flew at normal speed.

The difficulty the pilots had controlling the simulator was measured by calculating the deviation of altitude, heading, and speed from the recommended altitude, heading, and speed. The greater the deviation from the recommended flight parameters, the more difficulty the pilot would have had controlling the simulator. The pilots under high time pressure varied more from the recommended heading (150 degrees) compared to the pilots under low time pressure. This indicates that, at least for the 5nm period assessed, the pilots in the double speed condition had more difficulty maintaining the heading compared to the pilots in the normal speed condition. Therefore, the two speed conditions did not only differ in the degree of time pressure imposed on the pilots, but also in the ability for the pilots to control the simulator.

If the difficulty controlling the aircraft affected the amount of attention available there should be a relationship between difficulty handling the aircraft and in-flight decision-making. That is, the pilots who had difficulty controlling the aircraft should have been more likely to continue compared to the pilots who did not have difficulty controlling
the aircraft. The decision to divert to another airport or turn back to the departure point requires an active decision. The pilots who found it difficult to handle the simulator may not have had the opportunity to make such a decision, as they were too busy concentrating on flying the simulator. The only flight parameter for which this was found was the amount of deviation in altitude from 2,000 feet. The pilots who continued beyond the critical decision point deviated more from 2,000 feet compared to those who diverted. This supports O’Hare et al. (2007), who found that participants in the high workload condition were more likely to continue compared to the participants in the low workload condition. The participants in the high workload condition would have less opportunity for decision-making compared to the pilots in the low workload condition.

While difficulty flying may be a confounding variable, it does not explain why the pilots in the high time pressure condition were no more likely to continue compared to the pilots in the low time pressure condition. Both difficulty flying and increased time pressure should increase the likelihood of continuing into adverse weather, so if anything, it should have enhanced the relationship between time pressure and in-flight decision-making.

Along with difficulty flying, the two speed conditions may differ in the level of sunk-costs. As discussed in chapter one, the sunk-cost heuristic is when individuals attend to prior investments of time, money, or effort, which can lead to continuation of the original plan, even if the individual would be better off giving up or revising the plan (Arkes & Blumer, 1985). In the context of WRDM, investment of time or effort invested in the flight may lead to pilots continuing flying when they would be better off turning back or diverting to another airport. Previous research has found that the further the pilot had flown, the more likely it was that the pilot would have continued into adverse weather (Batt & O’Hare, 2005; O’Hare & Owen, 2002).

In the present study, by the critical decision point, the pilots in the normal-speed condition would have had invested more time and effort flying compared to the pilots in the double-speed condition. While the second weather change occurred approximately 90 nm into the flight regardless of speed condition, for the pilots in the normal speed condition, the weather change occurred approximately 50 minutes after take-off. Meanwhile, for the pilots in the double speed condition, this weather change occurred approximately 25 minutes after take-off. Therefore, the pilots in the normal speed condition had invested more time in the flight by the weather change, compared to the pilots in the double speed condition. According to the sunk cost hypothesis, the pilots in
the normal speed condition should be more likely to continue compared to those in the double speed condition. While there was no difference in the proportion of pilots who continued and diverted between the two speed conditions, the sunk cost effect may have diminished the predicted effect of time pressure on decision-making.

While inducing time pressure did not lead to poorer decision-making, the effect of time pressure on pilot decision-making depended on the pilots' level of flight experience. For the pilots in the normal speed condition, there was no difference in total hours flight experience between the pilots who continued and diverted beyond the critical decision point. However, for the pilots in the double-speed condition, the pilots who had diverted by the critical decision point were marginally more experienced compared to the pilots who continued. This supports past findings that time pressure affected inexperienced pilots' (Wiggins et al., 2002) and chess players' (Calderwood et al., 1988) performance and decision-making strategy.

The decision to divert or return to the departure point involves the recognition that the weather has deteriorated and that it would be dangerous to continue. Meanwhile, continuing does not necessarily involve making a decision, rather the pilot is just continuing with the original plan. Not only does diverting involve more effort but it also involves accepting a certain loss of the time and effort already spent flying. Meanwhile, continuing involves accepting an unknown risk, as the precise chance of a pilot crashing when flying into adverse weather is unknown. Since making the decision to divert involves more effort than continuing flying, and involves accepting a certain loss, making the decision to divert is more difficult compared to continuing the flight.

The RPD model (Klein, 1993, 1997) posits that experienced decision-makers will use a pattern matching strategy when making decisions. That is, the experienced pilots in the present study may have quickly recognised that the weather was deteriorating, and that it was no longer safe to continue flying. Therefore, the experienced pilots should not have been affected by time pressure, as they will quickly recognise that they should divert. Meanwhile, since the inexperienced pilots will not use such a decision strategy, their decision-making will be affected by time pressure.

While the inexperienced pilots in the normal-speed condition were not more likely to continue than the experienced pilots, the inexperienced pilots in the double speed condition were more likely to continue than the more experienced pilots. This provides support for
the RPD model, as it suggests that experienced pilots make decisions more rapidly than inexperienced pilots. It also suggests that time pressure does not equally affect all pilots.

The final aim of the present study was to assess the predictions of the MODE model (Fazio, 1990; Fazio & Towles-Schwen, 1999). This model predicts that inducing time pressure should decrease the pilots' ability to make decisions through conscious deliberation, and instead the pilots will intuitively make decisions. While under time pressure, the pilots' decisions should correspond more to their implicit attitudes compared to the pilots in the normal speed condition. Therefore, I expected that while under time pressure, the pilots who implicitly felt more anxious towards adverse weather, and implicitly perceived more risk in adverse weather, should have been more likely to divert compared to the other pilots. Meanwhile, for the pilots in the normal speed condition, there should be little difference in the implicit associations between the pilots who continued and diverted. Contrary to the prediction, I did not find an interaction between simulation speed and decision outcome. This suggests that being under time pressure did not lead pilots to make decisions consistent with their implicit associations. While this does not provide support for the MODE model, it does support O'Hare et al. (2007), who did not find an interaction between workload condition and in-flight decision-making. This lack of an interaction found in the present experiment may have resulted from the confounding variables discussed earlier, such as sunk costs.

The present study found a main effect of simulation speed on implicit risk perception. That is, the pilots in the double speed condition implicitly perceived more risk in adverse weather compared to the pilots in the normal speed condition. There are two possible explanations: the pilots' implicit risk perception may have been affected by their experience in the flight simulator, or the two groups of pilots may have differed in their implicit risk perception to begin with.

The pilots all completed the flight simulation before completing the four computer-based questionnaires. Therefore, it is possible that the simulated flight affected the pilots' implicit associations. Since the pilots who completed the flight at double speed had more difficulty controlling the simulator, the more negative experience of flying in simulated adverse weather may have led the pilots to feel that flying into adverse weather is riskier compared to the pilots who flew the simulation at normal speed. However, this contradicts the research findings of Egloff and Schmukle (2002), and Friese et al. (2006), that conscious distortion did not affect implicit attitudes. Furthermore, if experiences in the
flight simulator changed the participants’ implicit risk perception, then the pilots’ implicit anxiousness should have been similarly affected. Therefore, it is unlikely that the pilot’s experience in the flight simulator affected implicit risk perception. An alternative explanation is that the pilots differed in implicit risk perception to begin with. If this is the case, then the random assignment of the pilots into the two speed conditions was unsuccessful.

ADM is a very complex process (Wickens, 1999). Subsequently, it is difficult to discover which factors influence pilot behaviour, such as flying VFR-into-IMC. This is compounded by the small subject pool, which reduced the power of statistical tests. In the present study, I explored the influence of risk management and implicit associations on in-flight weather-related decision-making. Despite the low power, I found that the pilots who continued into adverse weather were less risk averse and tended to implicitly perceive less risk in flying in adverse weather. The pilots who diverted before the critical decision point perceived more risk in the hypothetical scenarios than those who crashed. This suggests a relationship between flying into adverse weather, risk tolerance, and implicit and explicit risk perception.
measured expertise in aeronautical risk perception, the participants’ CWS scores should be related to experience in ADM. Hunter’s (1995) HES, and in-flight decision-making in a simulated flight. The relationship between experience and expertise in ADM was assessed in two ways: by comparing the CWS scores of naïve (non-pilots), novice (student pilots), and experienced participants (qualified pilots); and by assessing the relationship between the pilots’ CWS scores and flight experience (e.g., total number of hours flight experience).

In my first study, I measured expertise in perceiving aeronautical risks in participants who were naïve (undergraduate psychology students) and experienced (qualified pilots) in ADM. Participants were presented with pen-and-paper flight scenarios, differing in the level of pilot, aircraft, and environmental-related risk, as well as risk from external pressures. The participants’ CWS scores, measuring expertise in aeronautical risk perception, were calculated from these risk ratings.

The CWS scores were partially related to experience in ADM. The pilots had a higher mean CWS score compared to the non-pilots, suggesting that the pilots were more expert at perceiving aeronautical risk in the hypothetical scenarios than the non-pilots. However, there was no relationship between the pilots’ CWS scores and flight experience. This may have been due to the experience level of the pilots. Most of the pilots were highly experienced, with an average of almost 10,000 total hours flight experience, and the majority of the pilots held an ATPL. This may explain the apparent lack of relationship between aeronautical risk perception and flight experience. The highly experienced pilots would have flown the majority of hours with an airline, while the type of experience predictive of experience in a GA setting should be hours flown in a GA aircraft (Wiggins & O’Hare, 2003a). Therefore, while there seemed to be no relationship between flight experience and CWS scores, this may have been due to the type of pilots in the sample.

Hunter (2002, 2005, 2006) found a relationship between pilots’ risk judgements and involvement in hazardous aeronautical events, using the HES (Hunter, 1995). The fewer hazardous events experienced by the pilots during the previous 24-months, the more risk perceived in the flight scenarios. According to Hunter, this indicates that risk perception plays a role in the pilots’ involvement in such incidents. For example, one reason why pilots fly VFR-into-IMC may be because they do not perceive that doing so is risky. The HES was administered to the pilots in Study One. I found no relationship between
involvement in hazardous events and the CWS scores, suggesting that expertise in aeronautical risk perception was not related to past involvement in hazardous events.

One of the criticisms of this study was that the scenarios were not presented in a format consistent with real-life pilot decision-making. For example, the en-route visibility was described as poor or good. Rather than being explicitly told whether the weather is suitable for flying, pilots usually gain understanding of the en-route weather from the internet or phoning a forecaster (NTSB, 2005). Furthermore, information regarding the visibility is usually presented in feet (U.S.) or metres as the horizontal distance at which the pilot can readily perceive objects (Fenwick, 1982). The visibility in this study was described in a qualitative manner to aid the non-pilots’ understanding of weather information.

To improve the realism of the scenarios, and to shift the focus onto WRDM, I designed a new task measuring expertise in weather-related aeronautical risk perception. The information was presented in a manner consistent with how pilots perceive weather-related information in real-life. Participants who were naïve (geography students), novice (student pilots) and experienced (qualified pilots) in ADM were presented with flight scenarios, differing in the level of weather-related risk. Since the naïve participants had weather-related knowledge, the information was presented in a more technical and detailed manner. Participants rated the risk involved in each scenario. The participants’ CWS scores were calculated from these ratings.

The CWS scores were not related to experience in weather-related decision-making. The three groups of participants did not differ in their CWS scores, although the pilots were more discriminating and the geography students were more consistent. Furthermore, there was little relationship between the pilots’ flight experience and their CWS scores. The one exception was a marginally significant relationship between the CWS scores and hours flown cross-country in the previous 90-days. However, this relationship was opposite to that predicted, with the more hours flown cross-country in the previous 90-days, the lower the CWS scores. These results suggest that expertise in weather related aeronautical risk perception was not related to experience in WRDM.

As well as completing the CWS procedure, the qualified and student pilots completed the HES. There was no relationship between the pilots’ involvement in hazardous aeronautical events and their CWS scores. However, this lack of relationship may have been due to the way in which the HES and CWS scores were measured. The CWS
procedure measured weather-related risk perception, while the HES measured involvement in a variety of hazardous events. There was some relationship between the pilots' CWS scores and involvement in weather-related hazardous events. The pilots who had experienced a forced landing due to bad weather had higher CWS scores compared to the pilots who had not experienced such a forced landing. Making a forced landing involves recognition that the weather is deteriorating and that there is risk involved in continuing flying. Therefore, this relationship between experience in making a forced landing and the CWS scores was expected, and suggests a relationship between involvement in weather-related aeronautical events and expertise in weather-related aeronautical risk perception.

The apparent lack of relationship between expertise in weather-related aeronautical risk perception and experience in WRDM may have occurred because people with a good memory may have been more consistent. Since the CWS score includes a measure of consistency, people who remember the rating given to a prior scenario will be more consistent than those who do not remember. The naïve participants were all postgraduate university students, and may have more experience remembering seemingly irrelevant information. Therefore, the geography students may have had an advantage in this task. This may explain why the geography students were more consistent compared to the qualified pilots and student pilots. Following this, the CWS procedure was modified to reduce the reliance on memory. The same scenarios, in the same order, were used, but a blocking task was included to stop the participants from rehearsing the information contained in each scenario, and therefore, to decrease the reliance on memory.

This new CWS procedure was initially tested on a small number of geography students and qualified pilots. While, due to the small sample size, no strong conclusions can be drawn, the new procedure seemed to worsen the performance of the geography students, while improving the performance of the pilots. As part of my fifth study, 32 pilots from Australia completed the CWS procedure, with the blocking task.

I assessed weather-related aeronautical risk perception in two samples of participants. In Study Two, I assessed expertise in weather-related risk perception in 11 qualified pilots, 11 student pilots, and 11 non-pilots (geography students) from the lower half of the South Island, New Zealand. In my fifth study, I assessed weather-related risk perception in 32 pilots from Sydney, Australia. The two studies also differed in the exact nature of the task. The New Zealand participants only completed the CWS measure and the demographic questionnaire. Meanwhile, the Australian pilots completed the task as part of a 2.5 hour
experiment, involving a simulated flight and a battery of questionnaires. Furthermore, the
Australian pilots had a blocking task between each of the scenarios.

The CWS scores were compared between the trained pilots in the two groups. The
pilot in Study Five without a PPL was excluded from this analysis. After data
transformation, the variances were not significantly different and the residuals were
normally distributed. The pilots from the Australian sample had a significantly higher
CWS score ($M_r = 6.21$, $SD_r = 5.217$; $M_t = 2.31$, $SD_t = .95$) than the pilots from the New
Zealand sample ($M_r = 2.98$, $SD_r = 1.41$; $M_t = 1.68$, $SD_t = .40$), $F(1,40) = 4.47$, $p = .041$, $\eta^2_p = .103$, $d = .87$. The mean Australian pilot’s CWS score was in the 79th percentile of the
New Zealand pilots’ scores, with a 47.4% non-overlap in the distributions. Although, after
data transformation, the variances were not significantly different, $F(1,39) = 3.95$, $p = .054$, it is clear that the Australian sample of pilots did vary more than the New Zealand
pilots. The mean CWS scores for the participants in the study two and study five are shown
in Figure 7.1.

![Figure 7.1. Mean CWS scores of the pilots in study two and study five.](image)

This result cannot be explained by experience levels. There was no difference
between the two groups of pilots in total number of hours flown, $Z(42) = -1.159$, $p = .257$;
number of hours flown PIC, Z(42) = -1.531, p = .131; hours flown cross-country, Z(42) = -.903, p = .381; hours flown cross-country in the last 90 days, Z(42) = -.698, p = .498; or the length certification was held, Z(41) = -.773, p = .459. However, the New Zealand pilots had flown more hours as pilot-in-command over the last 90 days (M = 55.76, SD = 55.64) compared to the Australian pilots (M = 25.50, SD = 37.07), Z(41) = -2.32, p = .019, d = .64. The mean number of hours flight experience as pilot-in-command over the previous 90-days for the New Zealand pilots was in the 73rd percentile of that of the Australian pilots, with 38.2% non-overlap between the two distributions. The two samples did not differ in level of certification, or number of pilots with an instrument rating. There was a trend towards the New Zealand pilots being more likely to hold an instructor rating, χ²(1) = 2.786, p = .070.

If the CWS procedure with the blocking task was a more valid measure of expertise in aeronautical weather-related risk perception, pilots should have higher scores in the new CWS compared to the old CWS. Blocking rehearsal should have improved performance because pilots will have been forced to make an intuitive judgment, rather than trying to remember what rating they had given to previous scenarios. The participants in the fifth study had significantly higher CWS scores, and moderately larger variance, than the participants in the second study. This difference in expertise could not be explained by differences in experience. For all but two indices of experience, there were no significant differences in experience level between the pilots in the two studies. However, the pilots in the second study had flown more hours pilot-in-command over the previous 90-days and tended to be more likely to hold an instructor rating.

The difference in CWS scores between the two samples may have been found because blocking rehearsal lead to better performance. There was more variation in Study Five’s pilots’ scores compared to the pilots in Study Two. As previously mentioned in the discussion for Study One, there should be a large variation in the pilots’ abilities to make weather-related judgments. Empirical research and accident reports suggest that some pilots are not very good at perceiving aeronautical risks (Hunter, 2002; Goh & Wiegmann, 2001b, 2002a), so I expected variation in the CWS scores. Together with the results of Study Three, these results suggest that the CWS with the blocking task may be a better measure of expertise in aeronautical weather-related risk perception than the standard procedure.
If the blocking task improved the validity of the CWS procedure, there should be a stronger relationship between the CWS scores and experience in weather-related decision-making in the fifth study, compared to Study Two. This was supported by the results. In Study Two, there was no relationship between CWS and experience, other than the number of hours flown cross-country in the previous 90-days. This relationship was the opposite of that predicted, as the more hours flown cross-country in the past 90-days, the lower the CWS score. Meanwhile, in study five, there was some relationship between experience and CWS scores. The certification of the pilot was related to CWS scores. The pilots with a CPL had a higher mean CWS score compared to the pilots with a PPL. Also, the pilots with a flight instructors rating had a higher mean CWS score compared to the pilots without a flight instructors rating.

The participants in Study Five also completed the HES. There was no relationship between expertise in weather-related risk perception and involvement in past aeronautical events in general, and involvement in weather-related hazardous events in particular.

If the CWS procedure measured expertise in weather-related risk perception, and if pilots sometime fly into adverse weather because of poor risk perception, then there should have been a relationship between in-flight decision-making in the simulated flight and the CWS scores. The pilots who continued beyond the first weather change (when the cloud ceiling decreased to 1,500 feet), had marginally higher CWS scores compared to the pilots who diverted after the second weather change. There was no relationship between the CWS scores and decision-making at the critical decision-point (when the cloud ceiling decreased to 800 feet). Therefore, there was either no relationship between expertise in weather-related risk perception and in-flight decision-making, or the relationship was opposite to that predicted. While the CWS procedure with the blocking task seemed to improve pilots’ expertise in aeronautical risk perception, and there was a relationship between the CWS scores and experience in ADM, the pilots who continued into adverse weather were not worse at this task compared to the pilots who diverted.

This lack of a relationship may indicate that risk perception is not related to in-flight weather-related decision-making. However, risk perception was also assessed through the average risk ratings that the participants had given for the scenarios. This is similar to Hunter’s (2002, 2005, 2006) measure of risk perception. If the participants’ average risk ratings were a good measure of risk perception, then there should be a relationship between the risk ratings and HES, and they should relate to in-flight decision-making. Unlike the
CWS procedure, there should not be a relationship between the average risk ratings and experience in ADM, as the average risk ratings did not measure expertise in risk perception.

In Study One there was a marginally significant correlation between the pilots’ average risk ratings and their HES scores. However, the relationship was in the opposite direction to that predicted. The more hazardous incidents experienced by the pilots, the higher their ratings of risk. In Study Two there was no relationship between the pilots’ average risk ratings and their HES scores. The results of both of these studies suggest that involvement in hazardous aeronautical events was not related to risk perception. However, in Study Five there was a marginally significant relationship between the pilots’ risk ratings and their HES scores. The more hazardous incidents experienced by the pilots, the lower their ratings of risk. Furthermore, there was a relationship between risk perception and in-flight decision-making. The pilots who diverted gave higher ratings of risk in the scenarios, compared to the pilots who crashed, suggesting a relationship between weather-related in-flight decision-making and risk perception.

While expertise in weather-related risk perception was not related in-flight WRDM, there was a relationship between the pilots’ average risk ratings and in-flight WRDM. This suggests that one of the reasons why pilots may fly VFR-into-IMC is because they do not perceive as much risk in flying in adverse weather compared to other pilots. This supports past research showing a relationship between risk perception and WRDM (Goh & Wiegmann, 2001b, 2002a; Hunter, 2002, 2005, 2006; O’Hare & Wiegmann, 2002).

Risk tolerance: People may take risks because they are willing to accept a high level of risk to accomplish a goal (Hunter, 2002). For example, one reason why pilots continue into adverse weather may be because they are willing to accept the risk of an accident in pursuit of their goal of arriving at their destination. There has been mixed evidence regarding the relationship between pilots’ risk tolerance and risk-taking. Some research has shown that pilots who take risks (e.g., flying VFR-into-IMC) are more risk tolerant compared to pilots who do not take such risks (Hunter, 2001; O’Hare, 1990; O’Hare & Wiegmann, 2003; Platenius & Wilde, 1989; Wiggins et al., 1996), while other studies have found no relationship (Hunter, 2002, 2005; Knecht et al., 2004).

Risk tolerance has been measured in many different ways: involvement in risky activities (Sicard et al., 2003; Turner & McClure, 2004), gap acceptance (e.g., Horswill &
Coster, 2001; Hunter, 2002, 2005), risky choice selection (Lejuez et al., 2002; Thomson et al., 2004), self-report (O'Hare, 1990; Platenius & Wilde, 1989; Wiggins et al., 1996), and personal minimums (Kirkbride et al., 1996; O'Hare et al., 2007). I used Lopes's (1987) theory to create a measure of risk tolerance. Lopes believes that a risk tolerant person is motivated by what he or she gain from risk-taking, while a risk averse person is motivated by what they can lose from risk-taking.

Qualified pilots were presented with scenarios varying in the opportunity for gain and the threat of loss and rated the likelihood that they would go on each flight from one (definitely would not go) to six (definitely would go). Multiple regression equations were calculated for each pilot, indicating the influence of opportunity and threat on the decision to take-off. A risk tolerant pilot would have been influenced by the opportunity for gain (e.g., the amount of money available to earn), while a risk averse pilot would have been influenced by the threat of loss (e.g., the threat of adverse weather).

Using this measure, the risk tolerance of qualified pilots was assessed in two studies. In Study Four, 27 pilots from New Zealand completed the risk tolerance measure and Hunter's (1995) HES. As part of Study Five, 32 pilots from Australia completed a simulated flight into adverse weather, the risk tolerance measure, and the HES. Across both studies, the pilots were largely risk averse as they were influenced by the level of threat in the situation and not the level of opportunity. That is, the majority of pilots would only take-off if the threat of loss was low, regardless of what there was to gain. When the threat was weather-related and the opportunity was for income, two pilots in Study Four, and one pilot in Study Five were risk tolerant. These pilots were more likely to take-off if there was the opportunity to earn a large amount of money, even if the weather-related threat was high.

If risk tolerance is related to risk-taking, this measure should be related to the HES and in-flight decision-making in a simulated flight. That is, the risk tolerant pilots should have been involved in more hazardous aeronautical events, and should be more likely to continue into adverse weather, compared to the risk averse pilots. In Study Four, the two pilots who were risk averse when the opportunity was to earn income and the threat was weather-related, were involved in significantly more hazardous aeronautical events compared to the risk averse pilots. This analysis was not conducted with the results from Study Five, as there was only one risk tolerant pilot. In Study Five, there was a relationship between risk tolerance and in-flight weather-related decision-making. The pilots who
diverted before the critical decision point were marginally more risk averse compared to the pilots who continued. Together, these results suggest a relationship between risk tolerance and risk-taking.

Personal minimums are another measure of pilots' risk tolerance. This was measured across all five studies. The pilots were asked to estimate the least visibility and lowest cloud ceiling which they would accept before taking-off on a local and cross-country flight. The qualified pilots' personal minimums across the five studies are presented in Table 7.1.

Table 7.1.
Personal Minimums for Visibility and Cloud Ceiling across the Five Studies.

<table>
<thead>
<tr>
<th></th>
<th>Study 1</th>
<th>Study 2</th>
<th>Study 3</th>
<th>Study 4</th>
<th>Study 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibility (L)</td>
<td>5km</td>
<td>5km</td>
<td>5km</td>
<td>5km</td>
<td>5km</td>
</tr>
<tr>
<td>Visibility (CC)</td>
<td>10km</td>
<td>7km</td>
<td>10km</td>
<td>10km</td>
<td>10km</td>
</tr>
<tr>
<td>Cloud-ceiling (L)</td>
<td>1500ft</td>
<td>1000ft</td>
<td>1000ft</td>
<td>1000ft</td>
<td>2000ft</td>
</tr>
<tr>
<td>Cloud-ceiling (CC)</td>
<td>1000ft</td>
<td>1500ft</td>
<td>1500ft</td>
<td>2000ft</td>
<td>3000ft</td>
</tr>
</tbody>
</table>

Note: L = local flight; CC = cross-country flight.

The participants in Study Five had stricter personal minimums for cloud ceiling compared to the participants in the other studies. That is, in Study Five, the median level of cloud ceiling the pilots required before taking off on a local and cross-country flight was 2,000ft and 3,000ft, respectively. In comparison, for the pilots in the other studies, the median level of cloud ceiling required before the pilots would take-off was between 1,000 to 1,500ft (local flight) and 1,000 to 2,000 feet (cross-country flight). The pilots in Study Five were not stricter in their personal minimums for visibility. The pilots in Study Five were from Sydney, Australia, whereas the pilots from the other three studies were from Canterbury, Otago, or Southland in New Zealand. The pilots from New Zealand may have been more lenient in their personal minimums for cloud ceiling because the weather is poorer in New Zealand compared to Australia. If the pilots from New Zealand were too strict in their personal minimums for cloud ceiling, perhaps they would never fly at all.

If personal minimums are a good measure of risk tolerance, and if risk tolerance is related to in-flight WRDM, there should have been a relationship between the pilots’
personal minimums and in-flight decision-making in Study Five. However, the pilots who continued beyond the critical decision point did not differ in their personal minimums compared to the pilots who diverted before the critical decision point. This supports O'Hare et al. (2007), who found no relationship between the pilots’ personal minimums for visibility and cloud ceiling and in-flight WRDM. While personal minimums may be a good method of training pilots to make non-compensatory decisions (Hunter et al., 2003), self-reported personal minimums may not reflect the pilots’ risk tolerance. There are two lines of evidence to support this.

For the pilots in Study One, the median acceptable cloud ceiling was lower for a cross-country compared to a local flight. Pilots should accept more stringent personal minimums on a local flight compared to a cross-country flight. Furthermore, although all 32 pilots of the pilots in Study Five took off on the simulated cross-country flight, with cloud ceiling forecast as 2,000 feet, the median personal minimums ceiling for a cross-country flight was 3,000 feet. That is, over half of the pilots would only take-off if the cloud ceiling was at least 3,000 feet, which was 1,000 feet lower than that forecast in the simulated flight. This may indicate that the pilots did not take the simulated flight seriously. Alternatively, it could indicate that the personal minimums question is not adequately measuring the pilots’ actual personal minimums.

The results of Studies Four and Five suggest that risk tolerance is related to involvement in hazardous events and in-flight WRDM. One of the reasons why pilots fly into adverse weather may be because they are willing to accept more risks to reach their goals compared to other pilots. This supports past research showing a relationship between risk tolerance and WRDM (Hunter, 2001; O’Hare, 1990; O’Hare & Wiegmann, 2003; Platenius & Wilde, 1989; Wiggins et al., 1996).

Implicit associations: Pilots’ feelings towards hazards, such as adverse weather, may be related to pilot decision-making. For example, a pilot may feel that flying in adverse weather is risky, and this may reduce the likelihood of the pilot taking-off in adverse weather. These feelings may exist on an implicit, or unconscious, level. In the fifth study, the pilots’ implicit risk perception and implicit anxiety were measured using the IAT (Greenwald et al., 1998). All of the pilots implicitly perceived more risk in IMC than VMC. Meanwhile, the majority of pilots felt more implicit anxiety towards IMC compared to VMC.
If implicit risk perception and implicit anxiety are related to WRDM, the pilots who continued into adverse weather should have perceived less risk in, and felt less anxiety towards, flying into adverse weather compared to the pilots who diverted. There was some relationship between implicit associations and in-flight WRDM. The pilots who crashed the simulator were less implicitly anxious towards adverse weather compared to the other pilots. The pilots who continued beyond the critical decision point were more implicitly anxious towards adverse weather compared to the pilots who diverted. Finally, the pilots who diverted had marginally stronger implicit attitudes towards adverse weather and risk compared to those who continued, suggesting a relationship between implicit risk perception and in-flight decision-making. One reason why pilots fly into adverse weather may be because they implicitly perceive less risk in doing so compared to other pilots.

The MODE model (Fazio, 1990; Fazio & Towles-Schwen, 1999) predicts that spontaneous behaviour is best predicted by implicit associations, and controlled behaviour is best predicted by explicit associations. This prediction was tested in Study Five. As a manipulation of time pressure, pilots flew the simulated flight at double-speed or normal-speed. Since the pilots in the double-speed condition had less time to make a decision when the weather deteriorated, their in-flight decision should have been more closely related to their implicit attitudes compared to those in the normal speed condition. However, an interaction between simulation speed and decision made was not found, suggesting that being under time pressure did not lead the pilots to make a spontaneous decision.

While time constraints can lead to poorer decision-making and increased risk taking in dynamic situations (Kerstholt, 1994), I found no difference in in-flight decision-making between the pilots in the double and normal speed conditions. This lack of relationship may have been due to increased sunk-cost in the normal-speed condition compared to the double-speed condition. The pilots in the normal-speed condition had invested more time in the flight when the weather deteriorated. Following the sunk-cost heuristic, the pilots in the normal-speed condition should have been more likely to continue compared to the pilots in the double-speed condition.

While I did not find a main effect of time pressure on in-flight decision-making, flight experience protected the pilots from the negative effect of time pressure on pilot decision-making. For the pilots in the normal-speed condition, there was no difference in the total hours flight experience between those who continued and diverted beyond the
critical decision point. However, for the pilots in the double-speed condition, the pilots who had diverted by the critical decision point were marginally more experienced compared to the pilots who had continued.

Limitations of Research

The most obvious limitation to the current research is the small number of participants. The sample size ranged from 27 to 33 participants, with only 10 and 11 qualified pilots in the first and second studies, respectively. Consequently, there was very low power in the statistics, and an increased probability of type two errors. Due to the small number of pilots in New Zealand, it was very difficult to obtain large numbers of pilots. However, as previously mentioned, Study Four would be ideally suited to the internet. Furthermore, pilots’ implicit risk perception could be assessed using internet-based IATs. Pilots could complete IATs measuring implicit risk perception and implicit anxiety, as well a questionnaire measuring pilots’ experience flying VFR-into-IMC. Such internet based measures would allow the collection of data from many pilots, without being restricted geographically.

A potential limitation with the CWS procedure was that the scenarios were created using a fractional factorial design. One of the aims of Study One was to assess whether the experts (qualified pilots) were more likely to use cues which were relevant to the safety of the hypothetical flight when making risk judgements, than the naïve participants (non-pilots). However, the two groups of participants were equally likely to use relevant cues when they made the risk judgements. Since the scenarios were created using a fractional factorial design, the main effects of factor were confounded by two-way and three-way interactions. The difficulty of measuring the main effects of the cues is further illustrated by the types of cues attended to by the non-pilots. Some non-pilots attended to factors which required technical knowledge to understand how they could affect the safety of the flight, such as whether the pilot filed a flight plan and the magneto drop. It is unlikely that the non-pilots actually used these factors to make risk judgements. However, the use of a fractional factorial design was necessary to create complex scenarios with a large number of factors. The main aim of Studies One, Two, and Three was to present complex scenarios to groups of participants who differed in the degree of experience in ADM, justifying the use of fractional factorial design.
In the risk tolerance measure, some of the facets of opportunity may have been perceived by the pilots as threats. For some pilots in Study Five, there was a negative relationship between the opportunity for career advancement and the likelihood of going on the hypothetical flights. These pilots would only go on the flight if the opportunity for career advancement was low, suggesting that the pilots viewed flying more advanced aircraft as a threat. According to Karren and Barringer (2002), it is important that the cues used in policy capturing research are independent, or orthogonal. This independence enables the researcher to assess the effect of each independent variable (e.g., threat and opportunity) on the dependent variable (e.g., likelihood of going on the flight). While I attempted to keep the level of threat constant across the level of opportunity, the results suggest that for some pilots, opportunity for career advancement was not independent of the level of threat in the situation.

Although the simulated flight was immersive, it may have lacked realism. While there are severe consequences of crashing a real aircraft, the only consequences of crashing the simulator were embarrassment and a damaged ego. Since the pilot would not sustain injuries or the simulated aircraft would not be damaged following a crash, it is possible that some of the pilots did not take the task seriously. This may have resulted in the pilots being more likely to take risks in the simulated flight compared to real flight. However, while the simulated flight may have lacked realism, it is not practical or ethical to send pilots on a real flight into simulated weather. The best that researchers can do is measure the relationship between risk management and WRDM using a variety of methods: pen-and-paper scenarios, cognitive task analysis, simulator studies, and accident and incident databases.

Measuring ADM in a laboratory setting can also lead to the possible problem of compliance. During Study Five, the pilots’ expectations of the experiment may have influenced their decision-making. Some of the pilots may have thought the goal of the simulated flight was to complete the flight, which may have made them more inclined to continue flying than they would be in real-life. However, I attempted to avoid this issue by making it clear that the pilots were pilot-in-command, and therefore, could fly as they please. The pilots were also told to inform the experimenter if they decided to do anything other than continue this flight. Therefore, they should have been aware that diverting to another airport or turning back to Quillayute was an option.
Theoretical Implications

The present research provides support for three theories relevant to pilot decision-making and risk management: Lopes’s (1987) model of risk tolerance, Klein’s (1993, 1997) RPD model, and Deery’s (1999) model of risk management. Each of these will be discussed in turn.

For any decision involving risk, there will always be opportunity for gain and threat of loss. Lopes (1987) believes that an individual who attends to the opportunity for gain will be more likely to take risks than an individual who attends to the threat of loss. The results of Studies Four and Five provide support for this theory. In Study Four, the pilots who were influenced by the opportunity to earn money had been involved in more hazardous incidents in the previous 24-months, suggesting a relationship between risk tolerance and risk-taking in a real-life context. In Study Five, the pilots who continued beyond the critical decision point were marginally less influenced by threat compared to the pilots who diverted before the critical decision point, suggesting a relationship between risk tolerance and risk-taking in a simulated flight. Together, these results suggest that pilots who attend to the opportunity for gain, or those who attend less to the threat of loss, are more likely to take risks while flying, supporting Lopes’ theory.

The RPD model (Klein, 1993, 1997) predicts that expert decision-makers (e.g., experienced pilots and fire commanders) will use prior experience to rapidly make decisions. Experts will only consider the costs and benefits associated with each option if there is time available. An implication of this model is that expert performance should be less affected by time pressure compared to novice performance. This has been found with expert pilots (Wiggins et al., 2002) and chess players (Calderwood et al., 1988). This was also found in Study Five. The pilots who continued under time pressure were marginally less experienced compared to the pilots who diverted while under time pressure. Meanwhile, for the pilots who were not under time pressure, there was no relationship between flight experience and in-flight decision-making. This supports the RPD model, as well as models described by Stokes et al. (1997) and O’Hare (1992). Both Stokes et al. (1997) and O’Hare (1992) believe that the experienced pilots should make decisions rapidly, through immediate access to their LTM. Therefore, experienced pilots should not have to weight up the costs and benefits associated with each option before making a decision.
Deery's (1999) model of risk management describes how risk management relates to decision-making. When faced with a potential hazard, a number of factors can influence the operator's (e.g., driver or pilot) decision and subsequent behaviour. The operator must first notice the hazard (e.g., storm clouds), must decide whether the hazard represents a safety risk, and decide whether he or she is skilled enough to handle the hazard. Finally, the operator must decide whether he or she is willing to accept the level of risk to obtain a goal. Only if the operator notices the hazard, decides that the hazard represents a safety risk and that he or she cannot cope with such a risk, and the risk is greater than that he or she is willing to accept, will the operator make a decision to avoid the hazard (e.g., turn back, divert, or navigate around the clouds).

In Study Five, I assessed the relationship between risk perception and risk tolerance on pilot decision-making. There was some relationship between risk perception and risk tolerance on WRDM during a simulated flight. While Deery's (1999) theory concerned driver decision-making, the present research has shown that WRDM in pilots may partially depend on the pilots' risk perception and risk tolerance. In Study Five, I also explored the relationship between implicit associations towards adverse weather and in-flight decision-making. There was a marginally significant relationship between implicit risk perception and in-flight decision-making, where the pilots who continued implicitly perceived marginally less risk in adverse weather compared to those who diverted before the critical decision point. This supports Pauley (2003), who found a relationship between implicit risk perception and pre-flight decision-making using pen-and-paper scenarios. Together, these results suggest that implicit risk perception plays an important role in pilot decision-making. Deery's model should be extended to include implicit and explicit risk perception. This is illustrated in Figure 7.2.
Figure 7.2. Deery's (1999) model of risk management (black lines) extended to include implicit and explicit risk perception (blue lines).

Practical Implications for the Aviation Industry

One of the important goals of ADM research is to improve pilot decision-making, through recommendations for training and selection. Due to the preliminary nature of the current research, more research is needed before any strong recommendations can be made. However, tentative implications to the aviation industry will be discussed regarding the role of risk tolerance and implicit risk perception.

According to Lopes (1987), pilots who attend to the opportunity for gains should be more likely to take risks compared to pilots who attend to the threat of loss. This was supported by the findings of Studies Four and Five. When making a pre-flight or in-flight decision, pilots should be encouraged to focus on the potential threats (e.g., adverse weather) and ignore the possible gains (e.g., the amount of potential earnings). For example, if faced with deteriorating weather, a pilot may be more likely to take-off on a medical rescue flight to rescue a critically injured patient than a patient with non-threatening injuries (NTSB, 2006). Therefore, medical rescue pilots are advised to make a go/no-go decision without regard for the patient’s injuries (NTSB, 2006; www.lifeflight.org.nz). Transport Canada Safety and Security (1998) recommend that medical rescue pilots should not be told the condition of the patient. This will allow the
pilot to make a go/no-go decision by considering the threats involved in the flight, without regard to the potential gains.

The current research did not find a relationship between personal minimums and in-flight decision-making. Despite the lack of relationship, the development of pilots' personal minimums is essential. This is especially true for private pilots, who do not work for companies or airlines. These pilots are solely responsible for any decisions made, and are often not given guidelines. Personal minimums make risk factors explicit, and encourage pilots to make decisions using a non-compensatory decision strategy (Hunter et al., 2003). The personal minimums program, developed by Jensen et al. (1998) should be taught in flying schools across New Zealand.

In Study Five, I found some relationship between implicit risk perception and in-flight decision-making. If implicit risk perception influences pilot decision-making, pilots should be made aware that they may possess implicit associations which may influence their decision-making. If there is a role of implicit risk perception in WRDM, then how do we use this information to train pilots in WRDM skills? Implicit associations are thought to arise from early experiences, while explicit associations arise from recent experiences (Rudman, 2004). For example, the experience of flying in low cloud and poor visibility may enhance pilots' understanding that these weather conditions are risky. This understanding may change pilots implicit and explicit risk perception. Such weather-related experience may be gained through simulator training.

Future Research

The present studies highlight the role of risk management in WRDM. Together, with past research, the present research suggests that risk perception and risk tolerance are key components of ADM and WRDM. I will discuss ideas for future research, which will increase our understanding of the role of implicit associations in decision-making, and how risk management skills develop.

In Studies Four and Five, I found a relationship between risk tolerance and risk-taking. The risk tolerant pilots in Study Four (those who attended to the level of opportunity in the hypothetical scenarios) had been involved in more hazardous aeronautical events over the previous 24-months, compared to the risk averse pilots (those who attended to the level of threat in the hypothetical scenarios). In Study Five, the pilots who were more risk averse were less likely to continue into adverse weather in a simulated
flight, compared to the pilots who were less risk averse. This illustrates the relationship between risk tolerance and risk-taking. However, this was found using hypothetical written scenarios.

An important next step would be to assess whether in real life situations some pilots are motivated by the threat of loss in the situation, while other pilots are motivated by the opportunity for gain. This could be assessed using cognitive task analyses, to assess which cues pilots use when making a decision (e.g., go/no-go decisions, in-flight decisions). Furthermore, it would be useful and interesting to assess whether the pilots who attend to the level of opportunity in the situation are more likely to take risks (e.g., taking-off when the weather is forecast to deteriorate). If some pilots attend to the degree of opportunity for gain in real-life situations, and if these pilots are more likely to take risks than the pilots who attend to the threat of loss, what effect does training have on the pilots’ risk-taking? That is, does training pilots to attend to the threat of loss and to ignore the opportunity for gain improve decision-making?

In Study Five, I measured implicit risk perception and anxiety towards adverse weather. Past research has also assessed pilots’ implicit excitement towards flying into adverse weather (O’Hare et al. 2007; Pauley, 2003). There may be other implicit associations related to ADM, such as valence. Hughson (2002) had non-pilots classify pictures of VMC and IMC, and words with positive and negative connotations into categories. This measured the participants’ implicit valence towards adverse and fine weather. A similar IAT could be used to assess pilots’ implicit valence, and to determine whether implicit valence was related to WRDM in a simulated flight.

Risk management is an important part of driver decision-making (e.g., Deery, 1999). For example, inexperienced drivers perceive less risk in driving situations compared to more experienced drivers, which may partially explain why inexperienced drivers are involved in more accidents (Brown & Groeger, 1998). Implicit risk perception may also play a role in drivers’ decision-making and risk-taking. It would be interesting to assess whether drivers who take risks (e.g., speeding and drink-driving) implicitly perceive less risk in doing so compared to drivers who do not take such risks.

Finally, while we are beginning to understand the role of risk management in ADM, we lack understanding of the development of risk management skills. How do pilots develop implicit and explicit risk perceptions, and how do they become risk tolerant? In the current studies, I have developed methods of assessing risk perception, risk tolerance,
and implicit attitudes in pilots, and have shown that there is some relationship between risk management and in-flight weather-related decision-making. However, there is a lack understanding of how pilots develop risk management skills. To improve pilot decision-making, we need to study how these skills develop.

**Conclusion**

The overall aim of my research was to develop measures of risk perception and risk tolerance, and to use them to assess the relationship between risk management and in-flight WRDM in a simulated flight. I have successfully developed measures of risk perception and risk tolerance and I found some relationship between risk management and WRDM. In the fifth flight simulator study, the pilots who diverted gave higher ratings of risk during the CWS task compared to the pilots who crashed. The pilots who diverted also tended to be more risk averse and implicitly perceived more risk in adverse weather, compared to the pilots who continued. I also examined the role of time pressure on decision-making and implicit attitudes. While I did not find a relationship between time pressure and decision-making, I found an interesting interaction with experience. Time pressure only affected the pilots’ decision-making when they were inexperienced. Together, these results further emphasise the role of risk management in pilot decision-making. More emphasis should be placed on risk management training in aviation.
References


Appendix A: Flight Scenarios (Study One)

One Thursday, a female pilot prepared to take her two company employees for a familiarisation flight from Dunedin to Christchurch. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 5 minutes before takeoff. She did not file a flight plan. The pilot had few hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of more than 200 rpm. The en-route weather was forecast as good visibility and high cloud ceiling. The forecast for tomorrow was good.

One Thursday, a male pilot prepared to take his two company employees for a familiarisation flight from Dunedin to Milford Sound. The aircraft was a twin-engine Piper Seneca and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 5 minutes before takeoff. He did not file a flight plan. The pilot had many hours experience flying a twin-engine Piper Seneca, had 8 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of more than 200 rpm on the right hand engine. The en-route weather was forecast as good visibility and low cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a female pilot prepared to take her two company employees for a familiarisation flight from Dunedin to Christchurch. The aircraft was a twin-engine Piper Seneca and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 5 minutes before takeoff. She filed a VFR flight plan. The pilot had few hours experience flying a twin-engine Piper Seneca, had 8 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of less than 50 rpm on each engine. The en-route weather was forecast as poor visibility and high cloud ceiling. The forecast for tomorrow was poor.
One Thursday, a female pilot prepared to take her two company employees for a familiarisation flight from Dunedin to Milford Sound. The aircraft was a single-engine Cessna 172N and was made in 1984. The pilot arrived at the airport to do her pre-flight preparation 2 hours before take off. She did not file a flight plan. The pilot had many hours experience flying a single-engine Cessna 172N, had 8 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of less than 50 rpm. The en-route weather was forecast as poor visibility and high cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a female pilot prepared to take two overseas tourists on a flight from Dunedin to Milford Sound. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 5 minutes before take off. She filed a VFR flight plan. The pilot had many hours experience flying a single-engine Cessna 172N, had 8 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of more than 200 rpm. The en-route weather was forecast as good visibility and high cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a male pilot prepared to take his two company employees for a familiarisation flight from Dunedin to Christchurch. The aircraft was a twin-engine Piper Seneca and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 2 hours before take off. He did not file a flight plan. The pilot had many hours experience flying a twin-engine Piper Seneca, had 4 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of more than 200 rpm on the right hand engine. The en-route weather was forecast as poor visibility and high cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a female pilot prepared to take two overseas tourists on a flight from Dunedin to Christchurch. The aircraft was a single-engine Cessna 172N and was made in 1984. The pilot arrived at the airport to do her pre-flight preparation 2 hours before take off. She filed a VFR flight plan. The pilot had few hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of less than 50 rpm. The en-route weather was forecast as poor visibility and high cloud ceiling. The forecast for tomorrow was good.
One Thursday, a male pilot prepared to take his two company employees for a familiarisation flight from Dunedin to Christchurch. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 2 hours before take off. He filed a VFR flight plan. The pilot had many hours experience flying a single-engine Cessna 172N, had 8 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of less than 50 rpm. The en-route weather was forecast as good visibility and high cloud ceiling. The forecast for tomorrow was good.

One Thursday, a female pilot prepared to take two overseas tourists on a flight from Dunedin to Milford Sound. The aircraft was a twin-engine Piper Seneca and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 5 minutes before take off. She did not file a flight plan. The pilot had many hours experience flying a twin-engine Piper Seneca, had 4 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of less than 50 rpm on each engine. The en-route weather was forecast as poor visibility and high cloud ceiling. The forecast for tomorrow was good.

One Thursday, a male pilot prepared to take two overseas tourists on a flight from Dunedin to Milford Sound. The aircraft was a twin-engine Piper Seneca and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 2 hours before take off. He filed a VFR flight plan. The pilot had few hours experience flying a twin-engine Piper Seneca, had 8 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of more than 200 rpm on the right hand engine. The en-route weather was forecast as poor visibility and high cloud ceiling. The forecast for tomorrow was good.

One Thursday, a female pilot prepared to take her two company employees for a familiarisation flight from Dunedin to Milford Sound. The aircraft was a twin-engine Piper Seneca and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 2 hours before take off. She filed a VFR flight plan. The pilot had few hours experience flying a twin-engine Piper Seneca, had 4 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of less than 50 rpm on each engine. The en-route weather was forecast as good visibility and low cloud ceiling. The forecast for tomorrow was poor.
One Thursday, a male pilot prepared to take two overseas tourists on a flight from Dunedin to Christchurch. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 5 minutes before take off. He did not file a flight plan. The pilot had few hours experience flying a single-engine Cessna 172N, had 8 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of less than 50 rpm. The en-route weather was forecast as poor visibility and low cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a female pilot prepared to take her two company employees for a familiarisation flight from Dunedin to Christchurch. The aircraft was a single-engine Cessna 172N and was made in 1984. The pilot arrived at the airport to do her pre-flight preparation 5 minutes before take off. She did not file a flight plan. The pilot had many hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of less than 50 rpm. The en-route weather was forecast as good visibility and low cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a male pilot prepared to take his two company employees for a familiarisation flight from Dunedin to Milford Sound. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 5 minutes before take off. He filed a VFR flight plan. The pilot had many hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of less than 50 rpm. The en-route weather was forecast as poor visibility and low cloud ceiling. The forecast for tomorrow was good.

One Thursday, a male pilot prepared to take his two company employees for a familiarisation flight from Dunedin to Milford Sound. The aircraft was a twin-engine Piper Seneca and was made in 1984. The pilot arrived at the airport to do his pre-flight preparation 5 minutes before take off. He did not file a flight plan. The pilot had few hours experience flying a twin-engine Piper Seneca, had 8 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of less than 50 rpm on each engine. The en-route weather was forecast as good visibility and high cloud ceiling. The forecast for tomorrow was good.
One Thursday, a female pilot prepared to take her two company employees for a familiarisation flight from Dunedin to Christchurch. The aircraft was a twin-engine Piper Seneca and was made in 1984. The pilot arrived at the airport to do her pre-flight preparation 5 minutes before take off. She filed a VFR flight plan. The pilot had many hours experience flying a twin-engine Piper Seneca, had 8 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of more than 200 rpm on the right hand engine. The en-route weather was forecast as poor visibility and low cloud ceiling. The forecast for tomorrow was good.

One Thursday, a male pilot prepared to take two overseas tourists on a flight from Dunedin to Milford Sound. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 2 hours before takeoff. He did not file a flight plan. The pilot had few hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of less than 50 rpm. The en-route weather was forecast as good visibility and high cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a female pilot prepared to take two overseas tourists on a flight from Dunedin to Milford Sound. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 2 hours before takeoff. He did not file a flight plan. The pilot had few hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of less than 50 rpm. The en-route weather was forecast as good visibility and high cloud ceiling. The forecast for tomorrow was good.
One Thursday, a female pilot prepared to take two overseas tourists on a flight from Dunedin to Christchurch. The aircraft was a twin-engine Piper Seneca and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 2 hours before take off. She did not file a flight plan. The pilot had many hours experience flying a twin-engine Piper Seneca, had 8 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of less than 50 rpm on each engine. The en-route weather was forecast as good visibility and low cloud ceiling. The forecast for tomorrow was good.

One Thursday, a male pilot prepared to take two overseas tourists on a flight from Dunedin to Milford Sound. The aircraft was a single-engine Cessna 172N and was made in 1984. The pilot arrived at the airport to do his pre-flight preparation 2 hours before take off. He did not file a flight plan. The pilot had many hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of more than 200 rpm. The en-route weather was forecast as good visibility and low cloud ceiling. The forecast for tomorrow was good.

One Thursday, a female pilot prepared to take two overseas tourists on a flight from Dunedin to Christchurch. The aircraft was a twin-engine Piper Seneca and was made in 1994. The pilot arrived at the airport to do her pre-flight preparation 2 hours before take off. She did not file a flight plan. The pilot had few hours experience flying a twin-engine Piper Seneca, had 8 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of more than 200 rpm on the right hand engine. The en-route weather was forecast as good visibility and high cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a male pilot prepared to take two overseas tourists on a flight from Dunedin to Christchurch. The aircraft was a twin-engine Piper Seneca and was made in 1996. The pilot arrived at the airport to do his pre-flight preparation 5 minutes before take off. He filed a VFR flight plan. The pilot had few hours experience flying a twin-engine Piper Seneca, had 4 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of more than 200 rpm on the right hand engine. The en-route weather was forecast as good visibility and low cloud ceiling. The forecast for tomorrow was good.
One Thursday, a female pilot prepared to take two overseas tourists on a flight from Dunedin to Milford Sound. The aircraft was a twin-engine Piper Seneca and was made in 1984. The pilot arrived at the airport to do her pre-flight preparation 5 minutes before take off. She did not file a flight plan. The pilot had few hours experience flying a twin-engine Piper Seneca, had 4 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of more than 200 rpm on each engine. The en-route weather was forecast as poor visibility and low cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a female pilot prepared to take two overseas tourists on a flight from Dunedin to Christchurch. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 2 hours before take off. She filed a VFR flight plan. The pilot had many hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of more than 200 rpm. The en-route weather was forecast as poor visibility and low cloud ceiling. The forecast for tomorrow was poor.

One Thursday, a male pilot prepared to take his two company employees for a familiarisation flight from Dunedin to Christchurch. The aircraft was a twin-engine Piper Seneca and was made in 1984. The pilot arrived at the airport to do his pre-flight preparation 2 hours before take off. He did not file a flight plan. The pilot had few hours experience flying a twin-engine Piper Seneca, had 4 hours sleep in the last 24 hours, and was suffering from a head cold. The run-up check showed a magneto drop of less than 50 rpm on each engine. The en-route weather was forecast as poor visibility and low cloud ceiling. The forecast for tomorrow was good.

One Thursday, a female pilot prepared to take her two company employees for a familiarisation flight from Dunedin to Milford Sound. The aircraft was a single-engine Cessna 172N and was made in 1996. The pilot arrived at the airport to do her pre-flight preparation 2 hours before take off. She did not file a flight plan. The pilot had few hours experience flying a single-engine Cessna 172N, had 8 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of more than 200 rpm. The en-route weather was forecast as poor visibility and low cloud ceiling. The forecast for tomorrow was good.
One Thursday, a male pilot prepared to take two overseas tourists on a flight from Dunedin to Milford Sound. The aircraft was a twin-engine Piper Seneca and was made in 1984. The pilot arrived at the airport to do his pre-flight preparation 2 hours before take off. He filed a VFR flight plan. The pilot had many hours experience flying a twin-engine Piper Seneca, had 8 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of less than 50 rpm on each engine. The en-route weather was forecast as poor visibility and low cloud ceiling. The forecast for tomorrow was poor.
One Thursday, a male pilot prepared to take his two company employees for a familiarisation flight from Dunedin to Milford. The aircraft was a single-engine Cessna 172N and was made in 1984. The pilot arrived at the airport to do his pre-flight preparation 5 minutes before take off. He filed a VFR flight plan. The pilot had few hours experience flying a single-engine Cessna 172N, had 4 hours sleep in the last 24 hours, and was in good health. The run-up check showed a magneto drop of more than 200 rpm. The en-route weather was forecast as poor visibility and high cloud ceiling. The forecast for tomorrow was poor.
Appendix B: Extra Information Sheet (Study One)

Cessna 172N

This is a simple, single engine aircraft with fixed landing gear.

Piper Seneca

This is a complex, twin engine aircraft with retractable landing gear.

General Information

- These are VFR flights (clear of cloud and in sight of ground).
- The pilot depicted in the scenarios has a Commercial Pilot Licence.

The following paragraph is a description of a typical General Aviation operation:
When the pilot arrives at the airfield, he or she will check that the aircraft is available. The pilot will then check the weather and other information pertaining to the route. Next, the pilot completes the flight planning forms. The pilot will then approach the aircraft and ensure that there is sufficient fuel to complete the flight. The pilot and passengers will then get into the plane and the passengers will be given a safety briefing. When the aircraft is started, the engines need to be warmed up. The pilot will taxi the aircraft out. He or she will do a final engine check. Finally, the aircraft will take off.
Appendix C: Demographic Aeronautical Practices Questionnaire

Cognitive Ergonomics and Human Decision-Making Laboratory

Aeronautical Decision-Making Questionnaire

Demographics, Practices & Opinions

No.
DEMOGRAPHIC INFORMATION

Participant Number: __________  Age: __________  Gender: M F

The following questions relate to your flying experience. Please answer the questions as accurately as possible.

Total number of hours: ________ hours

Number of hours as pilot-in-command: ________ hours

Number of hours as pilot-in-command on cross country flights: ________ hours

Number of hours as pilot-in-command over the last 90 days: ________ hours

Number of hours cross-country flights in the last 90 days: ________ hours

What is the highest level of certification you currently hold? (Circle) None PPL CPL ATPL

How long have you held this certification? ________ years

Do you hold, or have you ever held an instrument rating? Yes ☐ No ☐

Do you hold, or have you ever held a flight instructors rating? Yes ☐ No ☐

Have you ever been forced to land at an airport other than your destination due to a weather situation? Yes ☐ No ☐
**USUAL AERONAUTICAL PRACTICES**

If you are making a VFR CROSS-COUNTRY FLIGHT in a general aviation aircraft, what percentage of the time would you do the following?

<table>
<thead>
<tr>
<th>Activity</th>
<th>PERCENTAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I get a briefing on the weather before take off.</td>
<td>0 10 25 50 75 90 100</td>
</tr>
<tr>
<td>I compute my weight and balance before take off.</td>
<td>0 10 25 50 75 90 100</td>
</tr>
<tr>
<td>I compute my expected fuel consumption before take off.</td>
<td>0 10 25 50 75 90 100</td>
</tr>
<tr>
<td>I use a checklist for before take-off and before landing checks?</td>
<td>0 10 25 50 75 90 100</td>
</tr>
<tr>
<td>I file a flight plan.</td>
<td>0 10 25 50 75 90 100</td>
</tr>
<tr>
<td>I request weather updates during a flight.</td>
<td>0 10 25 50 75 90 100</td>
</tr>
<tr>
<td>I fly VFR above overcast cloud layers.</td>
<td>0 10 25 50 75 90 100</td>
</tr>
<tr>
<td>I fly at less than 1000 ft AGL to maintain cloud clearance.</td>
<td>0 10 25 50 75 90 100</td>
</tr>
<tr>
<td>I verify my fuel consumption rate in flight.</td>
<td>0 10 25 50 75 90 100</td>
</tr>
</tbody>
</table>
If you wanted to make a VFR flight for some personal or business reason (not involving life or death), what are the minimum conditions under which you would begin that flight?

If the **visibility** was lower than this value you would not take off.

<table>
<thead>
<tr>
<th>VISIBILITY (KM)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>A LOCAL (30 minute) daytime flight.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A CROSS-COUNTRY daytime flight.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If the **ceiling** was less than this value you would not take off.

<table>
<thead>
<tr>
<th>CEILING (FEET)</th>
<th>1000</th>
<th>1500</th>
<th>2000</th>
<th>3000</th>
<th>4000</th>
<th>5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>A LOCAL (30 minute) daytime flight.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A CROSS-COUNTRY daytime flight.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Past Aeronautical Practices (PAP)

In the past 24 months, how often have you experienced the following situations / events? Please answer by circling one of the numbers to indicate whether you have (0) never been involved in the event or involved 1, 2, 3, or 4 or more times.

01 How many times have you run so low on fuel (NOT because of equipment failures) that you were seriously concerned about making it to an airport before you run out? 0 1 2 3 4+

02 How many times have you made a precautionary or forced landing at an airport other than your original destination? 0 1 2 3 4+

03 How many times have you made a precautionary or forced landing away from an airfield? 0 1 2 3 4+

04 How many times have you inadvertantly stalled an aircraft? 0 1 2 3 4

05 How many times have you become so disorientated that you had to land or call ATC for assistance in determining your location? 0 1 2 3 4+

06 How many times have you had a mechanical failure which jeopardized the safety of the flight? (for example, nav failure while on a cross-country; landing gear stuck in up position, engine running rough or quitting) 0 1 2 3 4+

07 How many times have you had an engine quit because of fuel starvation, either because you ran out of fuel or because of improper pump or fuel tank selection? 0 1 2 3 4+
08 How many times have you flown into areas of instrumental meteorological conditions, without an instrument rating 0 1 2 3 4+ or an instrument qualified craft?

09 How many times have you become so disoriented after entering instrument meteorological conditions that you had difficulty in maintaining control of the aircraft? 0 1 2 3 4+

10 How many times have you turned back or diverted to another airport because of bad weather while on a VFR flight? 0 1 2 3 4+

11 How many times have you made what you later considered to be a very bad decision (something you will never do again) that could easily have resulted in an accident had circumstances been slightly different? (For example, deciding to press on to your destination in the face of deteriorating weather and landing just before a severe storm front passes through.)

12 How many aircraft accidents have you been in (as a flight crew member)? 0 1 2 3 4+
Appendix D: Demographic Information Questionnaire (Non-Pilot)

Participant Number: ..................

Gender: Male/Female

Age: .........................

Please rate your level of interest in aviation:

No interest ........................ Great Interest
1 2 3 4 5 6 7

Have you ever flown a PC-based simulator (e.g., Microsoft Flight simulator)?
Yes/No

Have you ever piloted a plane?
Yes/No

Do you have a family member or close friend who is a pilot?
Yes/No

If so, who (e.g., father, friend)? ........................................
Appendix E: Flight Scenarios (Study Two)

Weather for Taieri Airport, 22nd of April at 2pm

Area forecast for Plains
Valid for 22nd of April, 10am to 11pm

FORECAST WINDS ALOFT
3000' 10 knot at 315°

VISIBILITY
20 km reducing to 5000m in showers

CLOUD
Scattered at 900' and broken at 8000'

METAR NZTU (current weather at Timaru)
Valid for 22nd of April, 2pm.

WIND
20 knots at 200°

VISIBILITY
25 km reducing to 5000m in showers

CLOUD
Broken at 850' and broken at 8000'

TEMPERATURE
16°

DEW POINT
9°

QNH (Barometer)
1002
Weather for Taieri Airport, 22nd of April at 2pm

Area forecast for Clyde
Valid for 22nd of April, 10am to 11pm.
FORECAST WINDS ALOFT
3000' 10 knot at 315°
VISIBILITY
30 km
CLOUD
Broken at 900' and broken 8000'

METAR NZQN (current weather at Queenstown)
Valid for 22nd of April, 2pm
WIND
20 knots at 005°
VISIBILITY
25 km reducing to 5000m in showers
CLOUD
Scattered at 5000' and broken at 8000'
TEMPERATURE
16°
DEW POINT
9°
QNH (barometer)
1002
Weather for Taieri Airport, 22nd of April at 2pm

**Area forecast for Clyde**
Valid for 22nd of April, 10am to 11pm
FORECAST WINDS ALOFT
3000' 10 knot at 315°
VISIBILITY
30 km
CLOUD
Broken at 4000' and broken at 8000'

**METAR NZQN** (current weather at Queenstown)
Valid for 22nd of April, 2pm
WIND
2 knots at 230°
VISIBILITY
25 km reducing to 5000m in showers
CLOUD
Broken at 5000' and broken at 8000'
TEMPERATURE
16°
DEW POINT
9°
QNH (Barometer)
1002

Weather for Taieri Airport, 22nd of April at 2pm

**Area forecast for Plains**
Valid for 22nd of April, 10am to 11pm
FORECAST WINDS ALOFT
3000' 10 knot at 315°
VISIBILITY
30 km
CLOUD
Broken at 4000' and broken at 8000'

**METAR NZTU** (current weather at Timaru)
Valid for 22nd April, 2pm
WIND
20 knots at 200°
VISIBILITY
25 km reducing to 5000m in showers
CLOUD
Broken at 850' and broken at 8000'
TEMPERATURE
16°
DEW POINT
9°
QNH (Barometer)
1002
Weather for Taieri Airport, 22nd of April at 2pm

**Area forecast for Plains**
Valid for 22nd of April, 10am to 11pm

FORECAST WINDS ALOFT
3000' 10 knot at 315°

VISIBILITY
20 km reducing to 5000m in showers

CLOUD
Broken at 900' and broken at 8000'

**METAR NZTU** (current weather at Timaru)
Valid for 22nd of April, 2pm

WIND
2 knots at 200°

VISIBILITY
40 km

CLOUD
Scattered at 5000' and broken at 8000'

TEMPERATURE
16°

DEW POINT
9°

QNH (Barometer)
1002

---

Weather for Taieri Airport, 22nd of April at 2pm

**Area forecast for Clyde**
Valid for 22nd April, 10am to 11pm

FORECAST WINDS ALOFT
3000' 10 knot at 315°

VISIBILITY
20 km reducing to 5000m in showers

CLOUD
Broken at 900' and broken at 8000'

**METAR NZQN** (current weather at Queenstown)
Valid for 22nd April, 2pm

WIND
20 knots at 230°

VISIBILITY
40 km

CLOUD
Scattered at 850' and broken at 8000'

TEMPERATURE
16°

DEW POINT
9°

QNH (Barometer)
1002
### Area forecast for Clyde
Valid for 22nd of April, 10am to 11pm

**FORECAST WINDS ALOFT**
3000' 10 knot at 315°

**VISIBILITY**
20 km reducing to 5000m in showers

**CLOUD**
Broken at 4000' and broken at 8000' feet

**METAR NZQN** (current weather at Queenstown)
Valid for 22nd of April, 2pm

- **WIND**
  - 2 knots at 005°
- **VISIBILITY**
  - 40 km
- **CLOUD**
  - Broken at 850' and broken at 8000'
- **TEMPERATURE**
  - 16°
- **DEW POINT**
  - 9°
- **QNH (Barometer)**
  - 1002

---

### Area forecast for Plains
Valid for 22nd of April, 10am to 11pm

**FORECAST WINDS ALOFT**
3000' 10 knot at 315°

**VISIBILITY**
20 km reducing to 5000m in showers

**CLOUD**
Scattered at 4000' and broken at 8000'

**METAR NZTU** (current weather at Timaru)
Valid for 22nd of April, 2pm

- **WIND**
  - 2 knots at 315°
- **VISIBILITY**
  - 25 km reducing to 5000m in showers
- **CLOUD**
  - Scattered at 850' and broken at 8000'
- **TEMPERATURE**
  - 16°
- **DEW POINT**
  - 9°
- **QNH (Barometer)**
  - 1002
Weather for Taieri Airport, 22nd of April at 2pm

Area forecast for Plains
Valid for 22nd of April, 2pm
FORECAST WINDS ALOFT
3000' 10 knot at 315°
VISIBILITY
30 km
CLOUD
Scattered at 900' and broken at 8000'

METAR NZTU (current weather at Timaru)
Valid for 22nd April, 2pm
WIND
20 knots at 335°
VISIBILITY
20 km
CLOUD
Broken at 5000' and broken at 8000'
TEMPERATURE
16°
DEW POINT
9°
QNH (Barometer)
1002

Weather for Taieri Airport, 22nd of April at 2pm

Area forecast for Plains
Valid for 22nd of April, 10am to 11pm
FORECAST WINDS ALOFT
3000' 10 knot at 315°
VISIBILITY
30 km
CLOUD
Broken at 900' and broken at 8000'

METAR NZTU (current weather at Timaru)
Valid for 22nd April, 2pm
WIND
2 knots at 336°
VISIBILITY
25 km reducing to 5000m in showers
CLOUD
Scattered at 850' and broken at 8000'
TEMPERATURE
16°
DEW POINT
9°
QNH (Barometer)
1002
Area forecast for Clyde
Valid for 22nd of April, 10am to 11pm
FORECAST WINDS ALOFT
3000' 10 knot at 315°
VISIBILITY
20 km reducing to 5000m in showers
CLOUD
Scattered at 900' and broken at 8000'

METAR NZQN (current weather at Queenstown)
Valid for 22nd of April, 2pm
WIND
2 knots at 230°
VISIBILITY
25 km reducing to 5000m in showers
CLOUD
Broken at 5000' and broken at 8000'
TEMPERATURE
16°
DEW POINT
9°
QNH (Barometer)
1002
Area forecast for Clyde
Valid for 22nd of April, 10am to 11pm
FORECAST WINDS ALOFT
3000' 10 knot at 315°
VISIBILITY
30 km
CLOUD
Scattered at 4000' and broken at 8000'

METAR NZQN (current weather at Queenstown)
Valid for 22nd of April, 2pm
WIND
20 knots at 230°
VISIBILITY
40 km
CLOUD
Scattered at 850' and broken at 8000'
TEMPERATURE
16°
DEW POINT
9°
QNH (Barometer)
1002

Area forecast for Plains
Valid for 22nd of April, 10am to 11pm
FORECAST WINDS ALOFT
3000' 10 knot at 315°
VISIBILITY
30 km
CLOUD
Scattered at 4000' and broken at 8000'

METAR NZTU (current weather at Timaru)
Valid for 22nd of April, 2pm
WIND
2 knots at 200°
VISIBILITY
40 km
CLOUD
Scattered at 5000' and broken at 8000'
TEMPERATURE
16°
DEW POINT
9°
QNH (Barometer)
1002
Appendix F: Flight Scenarios (Study Four: Phase One)

On the next four pages you will be presented with a number of different situations and will be asked to rank the situations from 1 to 6. Use each number from 1-6 only once.

For example:

Rate these TV programs according to how much you enjoy them (1 = enjoy most 6 = enjoy least)

<table>
<thead>
<tr>
<th>TV program</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desperate Housewives</td>
<td>1</td>
</tr>
<tr>
<td>Shortland Street</td>
<td>5</td>
</tr>
<tr>
<td>Intrepid Journeys</td>
<td>4</td>
</tr>
<tr>
<td>The Young And The Restless</td>
<td>6</td>
</tr>
<tr>
<td>America’s Next Top Model</td>
<td>2</td>
</tr>
<tr>
<td>Game of Two Halves</td>
<td>3</td>
</tr>
</tbody>
</table>
Opportunity

On the next 4 pages you will be presented with sets of situations, each situation varying in the amount of opportunity. You should rank each situation from most to least opportunity. **Opportunity refers to the upside, or the possibilities of gain, of a particular situation.** Any situation may vary in the likelihood of gain, and gains may include such things as excitement, interest, and career advancement, admiration of others, financial gain, and self-satisfaction.
Please rank the opportunity in each of these situations (1 = most opportunity, 6 = least opportunity). Assume that you are the owner of a tourist flight operating company. Each passenger pays $350 for the scenic flight, but each time you fly, it costs $450 in operating costs.

<table>
<thead>
<tr>
<th>Number of Passengers</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Do you have any comments?
Please rank the opportunity in each of these situations (1 = most opportunity, 6 = least opportunity). Assume that you usually fly a Cessna 172 and in all cases a qualified instructor is present on the flight.

**Aircraft that you will be flying**  

<table>
<thead>
<tr>
<th>Aircraft that you will be flying</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cessna 172. This single-engine aircraft has a fixed pitch propeller and a fixed undercarriage.</td>
<td></td>
</tr>
<tr>
<td>Piper Arrow. This single-engine aircraft has a constant-speed propeller and a retractable landing gear.</td>
<td></td>
</tr>
<tr>
<td>Piper Seneca. This twin-engine aircraft has constant-speed propellers and a retractable landing gear.</td>
<td></td>
</tr>
<tr>
<td>Cessna 152. This single-engine aircraft has a fixed pitch propeller and a fixed undercarriage.</td>
<td></td>
</tr>
<tr>
<td>Cessna 182. This single-engine aircraft has a constant-speed propeller and a fixed undercarriage.</td>
<td></td>
</tr>
<tr>
<td>Piper Warrior. This single-engine aircraft has a fixed pitch propeller and a fixed undercarriage.</td>
<td></td>
</tr>
</tbody>
</table>

Do you have any comments?
Please rank the opportunity in these situations (1 = most opportunity, 6 = least opportunity).

<table>
<thead>
<tr>
<th>Situation</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>You are working towards your CPL and decide to go on a flight to accrue more hours.</td>
<td></td>
</tr>
<tr>
<td>You operate a tourist flight operating business. One morning you get a phone call from Peter Jackson’s agent. He would like you to transport Peter to look at film locations.</td>
<td></td>
</tr>
<tr>
<td>It is a Friday afternoon and you decide to take some time off work and go on a flight.</td>
<td></td>
</tr>
<tr>
<td>You have promised to take a friend on a flight. Your friend has never been in a light aircraft before.</td>
<td></td>
</tr>
<tr>
<td>You are a medical rescue pilot and get a phone call to transport a critically injured adult male.</td>
<td></td>
</tr>
<tr>
<td>You are a medical rescue pilot and get a phone call to transport an adult male with non-life threatening injuries.</td>
<td></td>
</tr>
</tbody>
</table>

Do you have any comments?
Please rank the opportunity in these situations (1 = most opportunity, 6 = least opportunity). Assume that in all cases you are flying a Cessna 172 and have one passenger, who is also an experienced pilot.

<table>
<thead>
<tr>
<th>Situation</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuits at your home airfield.</td>
<td></td>
</tr>
<tr>
<td>A local flight around your home town.</td>
<td></td>
</tr>
<tr>
<td>A local flight around a town that you are visiting.</td>
<td></td>
</tr>
<tr>
<td>A cross-country flight that you have been on many times before.</td>
<td></td>
</tr>
<tr>
<td>A cross-country flight. You have never flown this route before, although your passenger has. You have driven to this destination before.</td>
<td></td>
</tr>
<tr>
<td>A cross-country flight. You have never flown this route before, although your passenger has. You have never been to this destination.</td>
<td></td>
</tr>
</tbody>
</table>

Do you have any comments?
On the next 4 pages you will be presented with sets of situations, each situation varying in the amount of threat involved. You should rank the amount of threat from most to least threat. **Threat refers to the downside, or the possibility of loss, of a particular situation.** Any situation might contain a variety of threats, such as financial loss, danger, injury or property damage, loss of self-esteem, and disapproval of others. While all situations contain some probability of loss, some have a higher likelihood of loss, and therefore threat, than others.
Please rank the threat in these situations (1 = high threat, 6 = low threat). Assume that the pilot is about to go on a cross-country flight.

<table>
<thead>
<tr>
<th>The pilot arrived at the airfield</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 minutes before he was due to depart.</td>
<td></td>
</tr>
<tr>
<td>1 hour before he was due to depart.</td>
<td></td>
</tr>
<tr>
<td>15 minutes before he was due to depart.</td>
<td></td>
</tr>
<tr>
<td>45 minutes before he was due to depart.</td>
<td></td>
</tr>
<tr>
<td>1 hour 30 minutes before he was due to depart.</td>
<td></td>
</tr>
<tr>
<td>30 minutes before he was due to depart.</td>
<td></td>
</tr>
</tbody>
</table>

Do you have any comments?
Please rank the threat in these situations for a non-instrument rated pilot with a CPL (1 = high threat, 6 = low threat). In all situations the pilot is planning to leave on a cross-country flight at 10 am. The weather forecast is valid from 10 am to 11 pm.

<table>
<thead>
<tr>
<th>En-route weather forecast</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 km visibility reduced to 6000m visibility in showers. There is broken cloud at 2000 feet AMSL and the winds aloft at 3000 feet are 15 knots.</td>
<td>5</td>
</tr>
<tr>
<td>80km visibility with no precipitation. The skies are clear and the winds aloft at 3000 feet are 5 knots.</td>
<td>6</td>
</tr>
<tr>
<td>30 km visibility with no precipitation. There are scattered clouds at 5000 feet AMSL. The winds aloft at 3000 feet are 10 knots</td>
<td>4</td>
</tr>
<tr>
<td>10 km visibility reduced to 3000m visibility in showers. The cloud is overcast at 1000 feet AMSL. The winds aloft at 3000 feet are 20 knots.</td>
<td>5</td>
</tr>
<tr>
<td>40 km visibility, with no precipitation. There are few clouds at 10000 feet AMSL. The winds aloft at 3000 feet are 5 knots.</td>
<td>6</td>
</tr>
<tr>
<td>Isolated thunderstorms and rain with 1000m visibility. The cloud is overcast at 800 feet AMSL. The winds aloft at 3000 feet are 35 knots.</td>
<td>6</td>
</tr>
</tbody>
</table>

Do you have any comments?
Please rank the threat in these situations (1 = high threat, 6 = low threat). In all of the situations the pilot is planning on leaving on a flight at 10 am.

<table>
<thead>
<tr>
<th>Situation</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>The pilot had not drunk any alcohol for the last couple of days and had 8 hours sleep the previous night. She was on medication for a cold and had personal problems.</td>
<td></td>
</tr>
<tr>
<td>The pilot had not drunk any alcohol for the last week and had 8 hours sleep the previous night. She felt like she was coming down with a cold and had personal problems.</td>
<td></td>
</tr>
<tr>
<td>The pilot had been drinking moderately the night before and had 6 hours sleep the previous night. She was on medication for a cold and had personal problems.</td>
<td></td>
</tr>
<tr>
<td>The pilot had not drunk any alcohol for the last week and had 8 hours sleep the previous night. She felt healthy but had personal problems.</td>
<td></td>
</tr>
<tr>
<td>The pilot had been drinking heavily the night before and only had 5 hours sleep the previous night. She was on medication for a cold and had personal problems.</td>
<td></td>
</tr>
<tr>
<td>The pilot had not drunk any alcohol for the last week and had 8 hours sleep the previous night. She felt healthy and had no personal problems.</td>
<td></td>
</tr>
</tbody>
</table>

Do you have any comments?
Please rank the threat in these situations (1 = high threat, 6 = low threat). In each case, assume that the aircraft is a Cessna 172 and the pilot has his or her PPL.

<table>
<thead>
<tr>
<th>Situation</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>The aircraft had its last inspection 11 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of more than 200 rpm.</td>
<td></td>
</tr>
<tr>
<td>The aircraft had its last inspection 9 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 70 rpm.</td>
<td></td>
</tr>
<tr>
<td>The aircraft had its last inspection 9 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 150 rpm.</td>
<td></td>
</tr>
<tr>
<td>The aircraft had its last inspection 11 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 150 rpm.</td>
<td></td>
</tr>
<tr>
<td>The aircraft had its last inspection 2 months ago (annual or 100-hour). The run-up check showed a magneto drop of 50 rpm.</td>
<td></td>
</tr>
<tr>
<td>The aircraft had its last inspection 9 months ago (annual or 100-hour). The run-up check showed a magneto drop of 100 rpm.</td>
<td></td>
</tr>
</tbody>
</table>

Do you have any comments?
Appendix G: Flight Scenarios (Study Four, Phase Two)

Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs. Two tourists come into your business; if you take these passengers you will earn $100. Assume that you will be flying a Cessna 206. The en-route weather forecast is as follows: 80 km visibility with no precipitation. The skies are clear and the winds aloft at 3000 feet are 5 knots.

Please rate the likelihood that you would go on this flight VFR as described above.

1 2 3 4 5 6

Definitely no Definitely yes

Do you have any comments?
Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs.

Two tourists come into your business; if you take these passengers you will earn $100. Assume that you will be flying a Cessna 206.

The en-route weather forecast is as follows:

30 km visibility with no precipitation. There are scattered clouds at 5000 feet AMSL. The winds aloft at 3000 feet are 10 knots.

Please rate the likelihood that you would go on this flight as described above.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definitely no</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Definitely yes</td>
</tr>
</tbody>
</table>

Do you have any comments?
Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs. Two tourists come into your business; if you take these passengers you will earn $100. Assume that you will be flying a Cessna 206. The en-route weather forecast is as follows: Isolated thunderstorms and rain with 1000m visibility. The cloud is overcast at 800 feet AMSL. The winds aloft at 3000 feet are 35 knots.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs. Three tourists come into your business; if you take these passengers you will earn $350. Assume that you will be flying a Cessna 206. The en-route weather forecast is as follows: 80 km visibility with no precipitation. The skies are clear and winds aloft at 3000 feet are 5 knots.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no Definitely yes

Do you have any comments?
Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs. Three tourists come into your business; if you take these passengers you will earn $350. Assume that you will be flying a Cessna 206. The en-route weather forecast is as follows: 30 km visibility with no precipitation. There are scattered clouds at 5000 feet AMSL. The winds aloft at 3000 feet are 10 knots.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no Definitely yes

Do you have any comments?
Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs.

Three tourists come into your business; if you take these passengers you will earn $350. Assume that you will be flying a Cessna 206.

The en-route weather forecast is as follows:
Isolated thunderstorms and rain with 1000m visibility. The cloud is overcast at 800 feet AMSL. The winds aloft at 3000 feet are 35 knots.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no Definitely yes

Do you have any comments?
Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs.

Five tourists come into your business; if you take these passengers you will earn $850. Assume that you will be flying a Cessna 206.

The en-route weather forecast is as follows:
80 km visibility with no precipitation. The skies are clear and winds aloft at 3000 feet are 5 knots.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

| 1 2 | 3 4 | 5 6 |

Definitely no

Definitely yes

Do you have any comments?
Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs.
Five tourists come into your business; if you take these passengers you will earn $850.
Assume that you will be flying a Cessna 206.
The en-route weather forecast is as follows:
30 km visibility with no precipitation. There are scattered clouds at 5000 feet AMSL. The winds aloft at 3000 feet are 10 knots.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no Definitely yes

Do you have any comments?
Imagine that you are an owner of a tourist flight operating company, providing VFR charter flights from Timaru to Christchurch. Each passenger pays $250 for the flight but each time you fly, it costs $400 in operating costs. Five tourists come into your business; if you take these passengers you will earn $850. Assume that you will be flying a Cessna 206.
The en-route weather forecast is as follows:
Isolated thunderstorms and rain with 1000m visibility. The cloud is overcast at 800 feet AMSL. The winds aloft at 3000 feet are 35 knots.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Arrow (single-engine aircraft, constant speed propeller, and a retractable landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight.

You have 5 minutes to complete the pre-flight check and flight planning.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Arrow (single-engine aircraft, constant speed propeller, and a retractable landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight.

You have 30 minutes to complete the pre-flight check and the flight planning.

Please rate the likelihood that you would go on this flight as described above.

1  2  3  4  5  6

Definitely no  Definitely yes

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Arrow (single-engine aircraft, constant speed propeller, and a retractable landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight.

You have 1 hour to complete the pre-flight check and the flight planning.

Please rate the likelihood that you would go on this flight as described above.

1  2  3  4  5  6

|                | Definitely |    |
|----------------|------------|--|---|
| Definitely no  |            |    |   |
|                |            |    |   |

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Seneca (twin-engine aircraft, constant speed propellers, and a retractable landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight.

You have 5 minutes to complete the pre-flight check and flight planning.

Please rate the likelihood that you would go on this flight as described above.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definitely no</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Definitely yes</td>
</tr>
</tbody>
</table>

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Seneca (twin-engine aircraft, constant speed propellers, and a retractable landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight.

You have 30 minutes to complete the pre-flight check and the flight planning.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Seneca (twin-engine aircraft, constant speed propellers, and a retractable landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight. You have 1 hour to complete the pre-flight check and the flight planning.

Please rate the likelihood that you would go on this flight as described above.

1  2  3  4  5  6

Definitely no  Definitely yes

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Warrior (single-engine aircraft, fixed pitch propeller, and a fixed landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight.

You have 5 minutes to complete the pre-flight check and the flight planning.

Please rate the likelihood that you would go on this flight as described above.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definitely no</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Definitely yes</td>
</tr>
</tbody>
</table>

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Warrior (single-engine aircraft, fixed pitch propeller, and a fixed landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and flight instructor qualified on this type will be present on the flight.

You have 30 minutes to complete the pre-flight check and the flight planning.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you are planning on going on a cross-country flight in a Piper Warrior (single-engine aircraft, fixed pitch propeller, and a fixed landing gear). You usually fly a Cessna 172 (single-engine aircraft, fixed pitch propeller, and a fixed landing gear) and a flight instructor qualified on this type will be present on the flight.

You have 1 hour to complete the pre-flight check and the flight planning.

Please rate the likelihood that you would go on this flight as described above.

1  2  3  4  5  6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you have promised to take a friend on a flight today. Your friend has never been in a light aircraft before. You have not drunk any alcohol in the last couple of days and had 8 hours sleep the previous night. You are on medication for a cold and are having relationship problems.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you have promised to take a friend on a flight today. Your friend has never been in a light aircraft before.
You had been drinking heavily the night before and had only 5 hours sleep. You are on medication for a cold and have relationship problems.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you have promised to take a friend on a flight today. Your friend has never been in a light aircraft before.
You had been drinking heavily the night before and had only 5 hours sleep. You are on medication for a cold and have relationship problems.

Please rate the likelihood that you would go on this flight as described above.

1  2  3  4  5  6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you have promised to take a friend on a flight today. Your friend has never been in a light aircraft before. You have not drunk any alcohol the last week and had 8 hours sleep the previous night. You feel healthy and have no personal problems.

Please rate the likelihood that you would go on this flight as described above.

1  2  3  4  5  6

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Definitely no</td>
<td>Definitely yes</td>
</tr>
</tbody>
</table>

Do you have any comments?
Imagine that is a Friday afternoon and you decide to take some time off work to go on a flight.
You have not drunk any alcohol in the last couple of days and had 8 hours sleep the previous night. You are on medication for a cold and are having relationship problems.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no Definitely yes

Do you have any comments?
Imagine that is a Friday afternoon and you decide to take some time off work to go on a flight.
You had been drinking heavily the night before and had only 5 hours sleep. You are on medication for a cold and have relationship problems.

Please rate the likelihood that you would go on this flight as described above.

1  2  3  4  5  6

Definitely no  Definitely yes

Do you have any comments?
Imagine that it is a Friday afternoon and you decide to take some time off work to go on a flight.
You have not drunk any alcohol the last week and had 8 hours sleep the previous night. You feel healthy and have no personal problems.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no Definitely yes

Do you have any comments?
Imagine that you are a medical rescue pilot and get a phone call to transport a critically injured adult male.
You have not drunk any alcohol in the last couple of days and had 8 hours sleep the previous night. You are on medication for a cold and are having relationship problems.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you are a medical rescue pilot and get a phone call to transport a critically injured adult male. You had been drinking heavily the night before and had only 5 hours sleep. You are on medication for a cold and have relationship problems.

Please rate the likelihood that you would go on this flight as described above.

[1 2 3 4 5 6]

Definitely no

Definitely yes

Do you have any comments?
Imagine that you are a medical rescue pilot and get a phone call to transport a critically injured adult male.
You have not drunk any alcohol the last week and had 8 hours sleep the previous night. You feel healthy and have no personal problems.

Please rate the likelihood that you would go on this flight as described above.

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have flown this route many times before.
You will fly a Cessna 172 which had its last inspection 11 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of more than 200 rpm.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no Definitely yes

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have flown this route many times before.
You will fly a Cessna 172 which had its last inspection 2 months ago (annual or 100-hour). The run-up check showed a magneto drop of 50 rpm.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have flown this route many times before.
You will fly a Cessna 172 which had its last inspection 9 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 100 rpm.

Please rate the likelihood that you would go on this flight as described above.

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have never flown this route before, although your passenger has. You have driven to this destination before.

You will fly a Cessna 172 which had its last inspection 11 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of more than 200 rpm.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have never flown this route before, although your passenger has. You have driven to this destination before.

You will fly a Cessna 172 which had its last inspection 2 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 50 rpm.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6
Definitely no

Definitely yes

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have never flown this route before, although your passenger has. You have driven to this destination before.

You will fly a Cessna 172 which had its last inspection 9 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 100 rpm.

Please rate the likelihood that you would go on this flight as described above.

1  2  3  4  5  6

Definitely no  Definitely yes

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have never flown this route before, although your passenger has. You have never been to this destination before.

You will fly a Cessna 172 which had its last inspection 11 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of more than 200 rpm.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have never flown this route before, although your passenger has. You have never been to this destination before.

You will fly a Cessna 172 which had its last inspection 2 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 50 rpm.

Please rate the likelihood that you would go on this flight as described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Imagine that you and your passenger (also an experienced pilot) plan to go on a cross-country flight. You have never flown this route before, although your passenger has. You have never been to this destination before.

You will fly a Cessna 172 which had its last inspection 9 months ago (annual or 100-hour inspection). The run-up check showed a magneto drop of 100 rpm.

Please rate the likelihood that you would go on this flight under the conditions described above.

1 2 3 4 5 6

Definitely no

Definitely yes

Do you have any comments?
Appendix H: Flight Information (Study Five)

**FLIGHT INFORMATION**

**Scenario:**
It is 1700 hours Sunday. You are returning home to Tillamook (S47) with a friend after a trip to Quillayute (KUIL). You are expected back at work on Monday morning. The flight from Quillayute to your destination of Tillamook is approximately 154 nm. You will be flying a Cessna 172.

**Weather:**

*Flight Weather Forecast* for flight Quillayute - Tillamook  

<table>
<thead>
<tr>
<th>Synoptic Situation</th>
<th>A high is situated to the west of the west coast extending onto Washington State.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone</td>
<td>Quillayute – Tillamook</td>
</tr>
<tr>
<td><strong>Wind (Heights above M.S.L.)</strong></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>270 5 knots</td>
</tr>
<tr>
<td>3000</td>
<td>290 9 knots</td>
</tr>
<tr>
<td>5000</td>
<td>270 12 knots</td>
</tr>
<tr>
<td><strong>Weather</strong></td>
<td></td>
</tr>
<tr>
<td>Surface Visibility</td>
<td>15 to 20 km</td>
</tr>
<tr>
<td>Cloud</td>
<td>OVC - 2000 ft</td>
</tr>
<tr>
<td>Freezing Level</td>
<td>8000 ft</td>
</tr>
<tr>
<td>Ice Formation</td>
<td>-</td>
</tr>
</tbody>
</table>

**Task:**
Fly the pre-planned direct route between Quillayute and Tillamook using the laminated sectional chart and equipment provided. You are to fly this route with the aid of the moving-map GPS system installed in the aircraft.

**Aircraft:**
The aircraft has been fuelled with 73 litres of fuel – providing for a 77 minute flight time and a 45 minute reserve. You can assume it has already been subjected to an external pre-flight inspection.
You may request the amount of fuel to be changed before the flight begins.
You will be given a copy of the full aircraft specifications. A full set of checklists is available inside the simulator.

**Flight Plan (See Chart)**

**Quillayute → Tillamook**

**Distance:** 154 nm

**Estimated Fuel Burn:** 45 litres at 35 litres per hour

**Estimated time en route:** 1:17

<table>
<thead>
<tr>
<th>Waypoints</th>
<th>Route</th>
<th>Alt (ft)</th>
<th>Hdg (M)</th>
<th>Distance</th>
<th>GS(kts)</th>
<th>Fuel Time off</th>
<th>Track</th>
<th>TAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>KUIL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Leg</td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rem Est</td>
<td>Est</td>
<td>Est</td>
<td>ETE</td>
</tr>
<tr>
<td>S47</td>
<td>-D-+</td>
<td>2000</td>
<td>154</td>
<td>150</td>
<td>120</td>
<td>45</td>
<td>18:17</td>
<td>150</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Aerodrome Information**

**Quillayute (KUIL)**

Latitude: N47°56.23'  
Longitude: W124°33.67'  
Elevation: 194 FT

Runway 4 Length (ft) Surface  
4992 Concrete

**Tillamook**

Latitude: N45°25.09'  
Longitude: W123°48.86'  
Elevation: 36 FT

Runway 13 Length (ft) Surface  
4900 Asphalt  
31 4900 Asphalt  
1 2900 Asphalt  
19 2900 Asphalt
NOTAMS:
Evening Civil Time (ECT) at Tillamook is 1930 hours local time.
Tillamook (S47) ASOS is not available due to repairs.
Olympic A Military Operation Area (MOA) is not active today.
Olympic B MOA is not active today.
Appendix I: Colour Photos of IMC and VMC Pictures (Study Five)

VMC