

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Flights into Deteriorating Weather Conditions: Investigating Cognitive Biases
in Weather-Related Decision Making**

A thesis presented in partial fulfillment of the requirements for the degree of:

Doctorate of Philosophy

In

Aviation

At Massey University, Manawatu

New Zealand

Stephen Walmsley

2016

Abstract

In this thesis, the author's aim was to investigate whether the use of three cognitive heuristics may lead to systematic biases leading visual flight rules (VFR) qualified pilots to make inappropriate or ineffective decisions when faced with adverse weather and fly into instrument meteorological conditions (IMC). Although heuristics may reduce cognitive workload in weather-related decision making, they may lead VFR pilots to judge weather conditions as being better than they are in reality and continue flight into IMC conditions, when diverting or turning back would be the judicious choice.

Three cognitive biases that may potentially occur in pilot decisions to fly from VFR into IMC were identified: anchoring effect, confirmation bias and outcome bias. Three vignette-based studies found that pilots tended to anchor and under-adjust on initial information ($n = 201$), favour a confirmatory strategy when testing a hypothesis ($n = 278$) and evaluate judgments by the outcome rather than the decision process ($n = 300$).

Three intervention studies tested whether encouraging pilots to consider additional information rather than focusing on a narrow set of evidence when making judgments could reduce the impact of the three cognitive biases. Although a 'consider the alternative' strategy is sometimes effective, it was largely unsuccessful in reducing all three cognitive biases ($n = 101$). The perseverance of the biases in all six empirical studies is discussed in relation to the extant literature, as are the implications for flight-training and general aviation pilots generally.

Acknowledgments

I would like to express my appreciation to a number of people for their support and advice over the several years on my PhD journey.

I would like to express my sincere appreciation to my supervisor, Dr Andrew Gilbey, for the guidance and advice along the way. I would also like to express my thanks to my co-supervisor, Dr Jose Perezgonzalez.

I dedicate this thesis to my wife (Michaela) and my two sons (Matthew and Thomas) for their patience throughout the long PhD journey. Their support has always given me the impetus and motivation to complete this undertaking.

Table of Contents

ABSTRACT	ii
ACKNOWLEDGMENTS	iii
TABLE OF CONTENTS.....	iv
TABLES	ix
FIGURES	xi
GLOSSARY	xii
CHAPTER ONE.....	1
INTRODUCTION.....	1
1.1 The Thesis Context.....	1
1.2 Study Context	3
CHAPTER TWO.....	8
LITERATURE REVIEW.....	8
2.1 Chapter Overview.....	8
2.2 Pilot Licences and Ratings	8
2.3 VFR Meteorological Criteria.....	10
2.4 Historical Trends in VFR Flight into IMC Accidents.....	14
2.5 Primary Cause: Decision Error.....	16
2.5.1 Risk Perception	17
2.5.2 Social and Motivational	19
2.5.3 Situational Assessment.....	20
2.6 Cognitive Heuristics and Bias in Decision Making.....	22
2.6.1 Characteristics of Aeronautical Decision Making	22
2.6.2 Cognitive biases in Decision Making.....	26
2.7 Cognitive Bias in Weather-Related Decision Making	28
2.7.1 Decision Framing	28
2.7.2 The Sunk Cost Effect	29
2.7.3 Optimistic and Ability Bias.....	29
2.8 Cognitive Bias Explored in this Thesis	30
2.8.1 The Anchoring Effect.....	31
2.8.2 Confirmation Bias	34
2.8.3 Outcome Bias	38
2.9 Research Problem.....	42
2.10 Research Questions	43

2.10.1	Development of Research Question 1	43
2.10.2	Development of Research Question 2:.....	44
2.10.3	Development of Research Question 3:.....	45
2.10.4	Development of Research Questions 4–6: Debiasing.....	46
2.11	Study Overview	46
CHAPTER THREE		48
STUDY 1		48
THE ANCHORING EFFECT IN VFR FLIGHT INTO IMC.....		48
3.1	Introduction	48
3.2	Method.....	56
3.2.1	Participants	56
3.2.2	Material & Design.....	56
3.2.3	Procedure.....	63
3.3	Results	65
3.3.1	General findings	65
3.3.2	Cloud Height Assessment	67
3.3.3	Visibility Assessment.....	68
3.3.4	Safety Assessment For Continuing the VFR Flight.....	69
3.3.5	Confidence Assessment.....	70
3.3.6	Licence Type	70
3.3.7	Age	71
3.4	Discussion.....	73
CHAPTER FOUR		82
STUDY 2		82
CONFIRMATION BIAS AND SEEKING INFORMATION TO CHECK A HYPOTHESIS.....		82
4.1	Introduction	82
4.2	Method.....	87
4.2.1	Participants	87
4.2.2	Materials and Design.....	87
4.2.3	Procedure.....	93
4.3	Results	94
4.3.1	General findings	94
4.3.2	Novice vs Experienced Pilots.....	96
4.3.3	VFR Flight into IMC Experience.....	96
4.3.4	Instrument Rating.....	98
4.3.5	Licence Type	99
4.3.6	Age	100
4.4	Discussion.....	101

CHAPTER FIVE.....	106
STUDY 3	106
OUTCOME BIAS IN VFR FLIGHT INTO IMC AND THE INFLUENCE OF OUTCOME INFORMATION ON THE DECISION PROCESS	106
5.1 Introduction	106
5.2 Method.....	110
5.2.1 Participants	110
5.2.2 Materials and Design.....	111
5.2.3 Dependent Measures	115
5.2.4 Procedure.....	115
5.3 Results	117
5.3.1 General findings	117
5.3.2 Decision Judgment	118
5.3.3 Risk Assessment.....	119
5.3.4 Would Pilots Conduct the Same Flight?	120
5.3.5 Correlation.....	121
5.3.6 IR Status	122
5.3.7 Age Groups	122
5.3.8 Licence Groups	122
5.4 Discussion.....	123
CHAPTER SIX.....	128
DEBIASING WEATHER-RELATED DECISION MAKING.....	128
6.1 Introduction	128
6.2 Debiasing Techniques.....	129
6.3 Research Problem & Research Questions	133
6.4 Study 4 to 6 Overview	134
CHAPTER SEVEN.....	136
STUDY 4	136
DEBIASING THE ANCHORING EFFECT IN WEATHER-RELATED DECISION MAKING.....	136
7.1 Introduction	136
7.2 Method.....	137
7.2.1 Participants	137
7.2.2 Material & Design.....	138
7.2.3 Procedure.....	144
7.3 Results	145
7.3.1 General Findings	145
7.3.2 Cloud Height Assessment	146

7.3.3	Visibility Assessment	147
7.3.4	Comparison to Study 1	148
7.4	Discussion.....	149
CHAPTER EIGHT		152
STUDY 5		152
DEBIASING CONFIRMATION BIAS IN WEATHER-RELATED DECISION		
MAKING.....		
8.1	Introduction	152
8.2	Method.....	153
8.2.1	Participants	153
8.2.2	Material and Design	154
8.2.3	Procedure.....	159
8.3	Results	160
8.4	Discussion.....	161
CHAPTER NINE		165
STUDY 6		165
DEBIASING OUTCOME BIAS IN WEATHER-RELATED DECISION		
MAKING.....		
9.1	Introduction	165
9.2	Method.....	167
9.2.1	Participants	167
9.2.2	Materials and Design.....	167
9.2.3	Procedure.....	171
9.3	Results	173
9.3.1	General results.....	173
9.3.2	Decision Judgment	173
9.3.3	Risk Assessment.....	174
9.3.4	Would The Respondent Fly?.....	174
9.3.5	Comparison with Study 3	174
9.4	Discussion.....	175
CHAPTER TEN		178
GENERAL DISCUSSION.....		
10.1	Chapter Overview	178
10.2	Summary of Findings	178
10.2.1	Research Questions	179
10.3	Discussion.....	183
10.3.1	Cognitive Biases in Weather-Related Decision Making.....	183

10.3.2	Debiasing Weather-Related Decision Making.....	187
10.3.3	Implications of Cognitive Biases in Weather-Related Decision Making.....	189
10.4	Direction for Future Research	190
10.5	Limitations.....	196
CHAPTER ELEVEN.....		197
CONCLUSION.....		197
REFERENCES.....		201
APPENDICES		223
Appendix A	– Sample of Study 1 Questionnaire	223
Appendix B	– Low Risk Ethics Notification for Study 1.....	235
Appendix C	– Sample of Study 2 Questionnaire.....	236
Appendix D	– Low-Risk Ethics Notification for Study 2	243
Appendix E	– Sample of Study 3 Questionnaire.....	244
Appendix F	– Low-Risk Ethics Notification for Study 3.....	256
Appendix G	– Sample of Study 4–6 Questionnaire	257
Appendix H	–Human Ethics Approval for Studies 4–6	274
Appendix I	– Debiasing Lecture Slides	275

Tables

Table 1: <i>The five scenarios and their anchor pairings used in Study 1</i>	62
Table 2: <i>Mean assessment of cloud height and horizontal visibility distance by level of anchor</i>	67
Table 3: <i>Mean safety assessment scores about continuing the VFR flight (1 = not safe, 9 = very safe) by experience level (novice vs. expert) and anchor condition (high vs. low)</i>	69
Table 4: <i>Mean visibility assessment by licence type (student pilot [SP], private pilot licence [PPL], commercial pilot licence [CPL, air transport pilot [ATPL]]) and anchor condition (high vs. low).</i>	71
Table 5: <i>Evidence used for each scenario in Study 2</i>	90
Table 6: <i>Characteristics of pilots with and without past experience of VFR flight into IMC</i>	98
Table 7: <i>Demographic characteristics of pilots IR status</i>	99
Table 8: <i>Demographic characteristics of the pilot licence groups</i>	100
Table 9: <i>Negative outcome statements for each scenarios in Study 3</i>	114
Table 10: <i>Participants Decision Quality Assessment (1 = very good decision, 9 = very poor decision), Risk Assessment (1 =very low risk, 9 = very high risk) and Assessment of conducting the same VFR flight (1 = safe to conduct flight, 9 = not safe to conduct flight) between outcome condition.</i>	119
Table 11: <i>Participants Assessment of conducting the same VFR flight (1 = safe to conduct flight, 9 = not safe to conduct flight) between experience level (novice vs. expert) and outcome conditions.</i>	121
Table 12: <i>Pearson’s correlation between the three dependent variables (decision quality, risk assessment, safety assessment)</i>	121
Table 13: <i>Four anchor parings by scenario used in Study 4</i>	143

Table 14: <i>Cloud Height assessment after exposure to an anchor (high vs. low) and intervention group (debiased vs. control).</i>	147
Table 15: <i>Visibility assessment after exposure to an anchor (high vs. low) and intervention group (debiased vs. control).</i>	148
Table 16: <i>Evidence used in each scenario in Study 5</i>	158
Table 17: <i>Negative outcome statements for each scenarios in Study 6</i>	170

Figures

<i>Figure 1: New Zealand’s VFR Meteorological Minima (CAA NZ, 2015a)</i>	10
<i>Figure 2: Fatality rate for VFR flight into IMC accidents across a range of studies (countries) and time periods.</i>	14
<i>Figure 3: Simplified decision making model (Hutton & Klein, 1999; Madhavan & Lacson, 2006).</i>	24
<i>Figure 4: Characteristics of Type 1 and Type 2 processes in the Dual Process Theory [DPT] of decision making (Kahneman, 2011)</i>	28
<i>Figure 5: Comparison of Study 1 pilots to population totals by licence type</i>	66
<i>Figure 6: Mean assessment of horizontal visibility distance, by experience level (novice vs. expert) and by anchor condition (high vs. low).</i>	68
<i>Figure 7: Cloud assessment by age group and anchor condition (high vs. low).</i>	72
<i>Figure 8: Comparison of Study 2 pilots to population totals by licence type</i>	95
<i>Figure 9: Comparison of Study 3 pilots to population totals (Australia and USA).</i>	118

Glossary

The following terms and their corresponding definitions are used in the context of this thesis:

Term	Definition
ATPL	Air transport pilot licence: the highest level of aircraft pilot licence. Those certified are authorised to act as the pilot-in-command on larger aircraft that require two pilots to operate.
CAA NZ	Civil Aviation Authority of New Zealand: the regulatory authority of civil aviation in New Zealand.
CASA	Civil Aviation Safety Authority: the Australian national aviation authority (i.e., the government statutory authority responsible for the regulation of civil aviation).
CPL	Commercial pilot licence: a qualification that permits the holder to act as a pilot of an aircraft and be paid for his/her work. The pilot may also act as a co-pilot (first officer) of an aircraft that requires two pilots to operate.
FAA	Federal Aviation Administration: the regulatory authority of civil aviation in the United States.
GA	General aviation: aircraft operating on non-commercial flights. Aircraft of a variety of sizes can operate in GA, with four- to six-seater aircraft (e.g., a Cessna 172 with four seats) being a relatively common aircraft type
IFR	Instrument flight rules: regulations and procedures for flying aircraft by referring only to the aircraft instrument panel for navigation.
IR	Instrument rating: the qualifications that a pilot must have in order to fly under IFR.
IMC	Instrument meteorological conditions: meteorological conditions expressed in terms of visibility, distance from cloud and ceiling less than the minima specified for visual meteorological conditions.

NTSB	National Transportation Safety Board: an independent United States government investigative agency responsible for civil transportation accident investigation.
PIC	Pilot-in-command: in relation to any aircraft, means the pilot responsible for the operation and safety of the aircraft
PPL	Private pilot licence: a licence that permits the holder to act as the pilot-in-command of an aircraft privately (not for pay).
SP	Student pilot: someone who does not hold a pilot licence but is often in the training phase under supervision. They may fly solo without passengers provided they meet the required criteria (e.g., a valid medical certificate).
TSB	Transport Safety Board of Canada (officially the Canadian Transport Accident Investigation and Safety Board): the agency of the Government of Canada responsible for maintaining transportation safety in Canada.
VMC	Visual meteorological conditions: the meteorological conditions expressed in terms of visibility, distance from cloud and ceiling equal to or better than specified minima:
VFR	Visual flight rules: a set of aviation regulation under which a pilot may operate an aircraft in weather conditions that are sufficient to allow the pilot, by visual reference to the environment outside the cockpit, to control the aircraft's attitude, navigate and maintain safe separation from obstacles such as terrain, buildings and other aircraft.

CHAPTER ONE

INTRODUCTION

1.1 The Thesis Context

The role of a pilot is seldom passive; pilots are required to make a range of decisions, some of which may be complex, especially when made under conditions of uncertainty, when ambiguous information is involved or when there is limited time available. How visual flight rules (VFR) rated pilots make their decisions when approaching weather conditions unsuitable for visual flight (instrument meteorological conditions [IMC]) is of particular interest, because when a VFR pilot gets this decision wrong and flies into IMC, the consequences can be fatal (Wiggins, Hunter, O'Hare, & Martinussen, 2012).

VFR flight into IMC, is consistently the most common cause of general aviation (GA) weather-related accidents. In the United States (US) in 2011, 86% of accidents attributable to VFR flight into IMC were fatal; a rate well above that experienced in all GA accidents (Airline Owner and Pilots' Association [AOPA], 2014). Flying into IMC, when visibility may be marginal or even non-existent, the pilot must rely upon his/her instruments rather than a visual reference to the horizon to maintain control of the aircraft. Flying an aircraft with sole reference to the internal instruments requires

a higher level of training and skills, but, in principle, would provide pilots with the ability to avoid the illusions that can lead to loss of control of the aircraft or straightforward controlled flight into terrain. Despite the dangers of visual pilots flying into IMC, VFR flight into IMC remains a significant problem in general aviation (O'Hare, Hunter, Martinussen, & Wiggins, 2011).

The significant danger VFR flight into IMC poses was highlighted in a typical VFR flight into IMC accident in New Zealand in 2005. The pilot was conducting a cross-country flight from Nelson to Christchurch with a single passenger on board. Before departing, the pilot obtained the weather forecast for the flight, which indicated marginal weather conditions that were still suitable for a VFR flight. The initial phase of the flight was uneventful; however, as the flight approached Christchurch, the weather began to deteriorate significantly more than forecasted. The pilot began to fly lower and lower, until the aircraft inadvertently entered cloud (IMC). While flying in IMC, with little or no external reference, the pilot probably suffered a sensory illusion, leading to spatial disorientation. Essentially, this illusion results in a pilot not knowing which way is 'up', leading to the pilot inadvertently applying control input to counter the perceived adverse aircraft attitude. Even though the aircraft was equipped to fly in the cloud, the pilot was not qualified. Without any external references, the pilot crashed a short time later into the sea (Civil Aviation Authority of New Zealand [CAA NZ], 2005a).

With vision being the most powerful contributor to orientation, when a pilot enters conditions with limited visibility, orientation becomes difficult (Green, James,

Gradwell, & Green, 1996; Thom, 2003). When a pilot becomes spatially disorientated, he or she must act quickly. In one study, all participants lost control of the aircraft after entering simulated IMC and the time it took for the pilots to lose control averaged 178 seconds (cited in O'Hare & Roscoe, 1990). Once a VFR flight enters IMC, the final outcome will depend on a number of factors, with just a few seconds or, at most, a few minutes separating a safe outcome from a fatal one (Batt & O'Hare, 2005).

1.2 Study Context

Weather-related GA accidents have been a concern for some time (Hunter, Martinussen, Wiggins, & O'Hare, 2011). One of the most alarming aspects of this type of accident is that no significant improvement has been seen in this accident rate since the first major study into this type of accident was conducted by the National Transport Safety Board (NTSB) of the US (NTSB, 1968). The accident rate for VFR flight into IMC has remained considerably higher than all GA accidents (AOPA, 2014). Furthermore, this type of accident is not limited to a single country: the accident rate in other countries with varying meteorological and geographic conditions (e.g., United Kingdom, Australia, New Zealand) also indicates a high fatality rate (Batt & O'Hare, 2005; O'Hare & Smitheram, 1995).

No simple explanation can account for all VFR flight into IMC accidents and, as is the case for any aircraft accident, a range of factors have been identified as possible contributing factors. The initial focus in this area concentrated on pilot

demographics, operational and geographic factors (NTSB, 1989, 2005). More recently, the focus has been on the psychological aspects of pilots' weather-related behaviour, particularly, weather-related decision making. Inefficient or ineffective weather-related decision making has been cited as a significant causal factor in GA accidents (Hunter, Martinussen, & Wiggins, 2003; Madhavan & Lacson, 2006).

Recent studies have supported a 'situational assessment' explanation for a pilot's decisions to continue a flight into IMC. The situational assessment explanation suggests that a pilot's decision to continue flight into IMC may be caused by impaired awareness of the information used during weather-related decision making (O'Hare et al., 2011). If the pilot has accurately recognised that he/she is flying towards conditions that are unsuitable for VFR flight, he/she is likely to divert (Wiegmann, Goh, & O'Hare, 2002). The accident report into the VFR flight into IMC accident described earlier (CAA NZ, 2005a) suggested that impaired situational awareness may have played a part in this accident, with the pilot's perception of the conditions influenced by the 'good' weather forecast received prior to the flight.

Development of advanced technologies is an appealing response to VFR flight into IMC accidents. This approach has resulted in significant improvements in safety in other areas of aviation (e.g., ground proximity warning systems in commercial aviation). However, in GA, in weather-related decision making, it appears to worsen, rather than improve, the problem (Wiggins, 2007). For example, Johnson, Wiegmann, and Wickens, (2006) found that pilots who were operating an aircraft with advanced displays were more likely to continue flight into deteriorating weather

conditions compared to pilots without advanced displays. This is in part because pilots spend more time looking inside the cockpit at various displays rather than looking outside (Johnson et al., 2006; Williams, Yost, Holland, & Tyler, 2002).

In this thesis, the aim of the author was to gain further understanding as to why pilots qualified only to fly under VFR fly into IMC, using the framework of cognitive heuristics and biases when making judgments under conditions of uncertainty. Cognitive biases may help to explain impaired situational awareness whereby VFR pilots make inappropriate or ineffective decisions to continue into IMC. A substantial body of evidence has identified cognitive biases as being a significant factor in decision error in a range of fields (see, e.g., Croskerry, 2002; English, Mussweiler, & Strack, 2006). However, cognitive biases have received limited attention in the literature on pilots' weather-related decision making.

In principle, when pilots are considering the weather ahead, they should carefully consider all available information before making a judgment (i.e., to proceed, divert or turn back). However, in practice, not all evidence required to make a judicious decision is always used, the evidence may not be available, the time in which to make a decision may be too short or perhaps the evidence is too readily discounted or simply ignored. Indeed, there are a plethora of situations in which the evidence overwhelmingly suggests that humans do not always evaluate evidence impartially or fully, but use cognitive heuristics (shortcuts). These cognitive heuristics are probably the result of the limitations in people's decision making capacity (Epley & Gilovich, 2001; Kahneman, 2011). In other words, to ease cognitive workload, a number of

heuristics are used to find adequate, though potentially imperfect, answers to what may be difficult and complicated questions (Gigerenzer & Gaissmaier, 2011).

Heuristic processes are an important part of decision making. They serve well most of the time, enabling people to make a judgment swiftly (Croskerry, 2014; Dobelli, 2013; Gigerenzer & Gaissmaier, 2011). Nevertheless, there is a price to pay for using heuristics. Occasionally, they can lead to systematic errors (biases), which essentially deviate from what would reasonably be considered rational or good judgments. This is especially the case when faced with an uncertain situation with ambiguous information (Kahneman, 2011; Tversky & Kahneman, 1974), as often may be encountered in GA.

The secondary aim of this thesis was to explore the role that experience plays in the impact of cognitive biases in weather-related decision making. Because of the nature of weather-related decision making, it is often considered a skill that a pilot will gradually develop as he or she is exposed to different weather-related environments. If a pilot is to obtain these skills through experience, they may put themselves into danger in the process. Therefore, understanding how experienced pilots make decisions may be useful in training less experienced pilots (Wiggins & O'Hare, 1995).

Under some circumstances—for example, deciding what one's family should eat for dinner—a decision made using cognitive heuristics is quite reasonable, especially

when the consequences of a wrong choice are minimal (Tversky & Kahneman, 1974). In an aviation environment, especially when operating in a dynamic, time-pressured situation like a VFR pilot experiences when flying near adverse weather, cognitive heuristics may result in the pilot forming an incorrect perception of the weather condition. If a VFR pilot's judgment is incorrect and the wrong decision is made, the consequences can be extremely serious. In consideration of the serious consequences that can result from cognitive biases, it is important to get a better understanding of their impact on pilots' weather-related decision making, particularly that which occurs prior to VFR flight into IMC.

CHAPTER TWO

LITERATURE REVIEW

2.1 Chapter Overview

This chapter presents a review of the key concepts in the literature that are applicable to this research. First, pilot licences and ratings are discussed. Second, the meteorological conditions required for a pilot to remain in visual flight will be reviewed. Third, historical trends in VFR flight into IMC accidents will be discussed. Fourth, the current theories behind the causes of this type of accident are discussed. Fifth, the part cognitive biases play in decision error and the cognitive biases that may influence weather-related decision making are discussed. This chapter concludes with the research problem and the research questions that will be addressed in this thesis.

2.2 Pilot Licences and Ratings

To operate an aircraft, a pilot must first undergo considerable training to obtain a private pilot licence (PPL). This would typically involve approximately 50 hours of flight experience, which is heavily supervised. During this phase, the pilot would be classified as a student pilot (SP). On successful completion of a competency check, a pilot is permitted to carry passengers and operate unsupervised; this would be

comparable to a driver passing a full driver licence test. Approximately half of the currently active pilots only hold a PPL (Civil Aviation Safety Authority [CASA], 2013; Federal Aviation Authority [FAA], 2009); the rest have undergone further training to pursue a career as a pilot. The next step would be to obtain a Commercial Pilot Licence (CPL) and then an Air Transport Pilot Licence (ATPL). The CPL training consists of an additional 100–150 hours of flight training and would allow a pilot to operate for hire or reward, either as the pilot-in-command (PIC) of a small aircraft or first officer (co-pilot) of a larger aircraft. An ATPL allows the pilot to operate as the PIC of a large aircraft (e.g., Boeing 737, Airbus A320) that requires two pilots to operate and normally requires a pilot to obtain at least 1500¹ hours before sitting a competency check. Pilots that hold a higher licence (i.e., CPL or ATPL) may also exercise the privileges of a lower licence. For example, an ATPL holder may wish to fly privately during a day off from operating commercial flights. In the context of this thesis, this is an important aspect, as a pilot conducting a VFR flight could hold any of the above pilot licences.

A pilot may also wish to add extra endorsements to their pilot licence, such as a night rating (to operate at night) or an aerobatic rating. Each rating requires additional training and competency tests. The basic PPL or CPL limits the pilot to operating in visual conditions and adhering to VFR. If a pilot wishes to operate in areas of low visibility or in cloud, an instrument rating (IR) is required (this will be discussed in more detail in the next section).

¹ The flight hour requirements for each of the licence types described in this section are for pilots wishing to obtain a license in New Zealand. A pilot would also be required to pass various knowledge tests for each licence.

2.3 VFR Meteorological Criteria

A flight conducted under VFR requires the pilot to remain in visual conditions during the flight. The minimum meteorological criteria (horizontal visibility and cloud height limits) vary slightly between countries and also within different types of airspace. In general, they are designed so pilots can visually orientate themselves, navigate using visual features and maintain visual separation from other aircraft (O'Hare et al., 2011). Figure 1 provides a visual representation of New Zealand's VFR meteorological minima. For flight below 3000 ft, the metrological conditions in uncontrolled² airspace are a horizontal visibility of more than 5000 m (5 km), operate clear of cloud and being in sight of the ground. For flight above 3000 ft, the aircraft must also remain 1000 ft vertically from cloud (CAA NZ, 2015b).

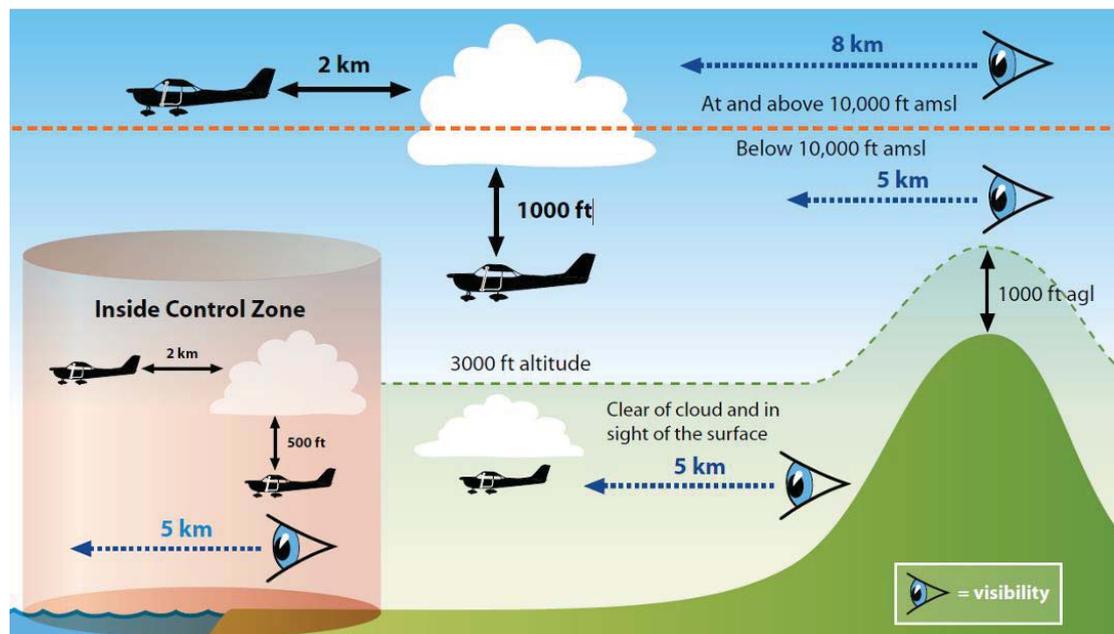


Figure 1: New Zealand's VFR Meteorological Minima (CAA NZ, 2015a)

² Uncontrolled airspace is where air traffic control (ATC) service is not deemed necessary. ATC does not exercise any authority but may provide basic assistance as a pilot requests. Flight in uncontrolled airspace will typically operate under VFR.

These weather conditions are referred to as visual metrological conditions (VMC). Pilots operating in these metrological conditions would be required to adhere to VFR. If the weather conditions do not meet these minimum requirements, the weather is referred to as IMC, these would include in cloud or areas of low visibility (e.g., a rain shower). For a pilot to conduct a flight in IMC advanced training is required to obtain an IR. This training is to ensure the pilot is able to orientate and navigate solely using the flight instruments. Furthermore, an IR needs to remain 'current', which requires regular practice of these advanced skills, which is particularly important because, without regular practice, IR skills degrade quickly (FAA, 2004; Wiggins et al., 2012).

Even if a pilot holds an IR, this does not automatically mean he or she can operate in IMC when ever he or she desires. Flights operating in IMC are normally conducted under specific conditions. For example, the aircraft needs to have specific instruments that are significantly beyond those which are required for a VFR flight. Furthermore, a flight in IMC is generally conducted along specific 'highways' in the sky, which ensure that the aircraft remains clear of terrain. Therefore, on a VFR flight, a pilot with a current IR may not safely fly into IMC.

If a VFR pilot flies into conditions that are unsuitable for visual flight (i.e., he or she flies into cloud or areas of low visibility), he or she may struggle to maintain spatial orientation. This is because vision is the most powerful contributor to orientation, so when a pilot enters conditions with limited visibility, orientation becomes difficult (Green et al., 1996). This situation is classified as VFR flight into IMC. Even if a pilot is able to remain in control of the aircraft, without visual reference, navigation

becomes a very difficult task, which can result in controlled flight into terrain. Not all VFR flight into IMC will end in an accident. Depending on a number of factors (e.g., local terrain and the extent of adverse weather), the pilot may safely regain visual conditions. In this case, the occurrence would be classified as an incident. However, if an accident does occur, the end result is more often than not fatal due to the high-velocity nature of the impact (Batt & O'Hare, 2005).

Another potential outcome is that the pilot may decide to conduct a 'precautionary landing' (also referred to as an 'off field' landing). In this situation, if a pilot recognises that he or she will shortly no longer be able to maintain VFR, he/she may elect to land in an area which is sufficiently large enough, such as a nearby field. Although, in principle, a precautionary landing is more ideal than entering IMC, the outcome can often be just as serious. The hazards of conducting a precautionary landing can be observed in the following accident report:

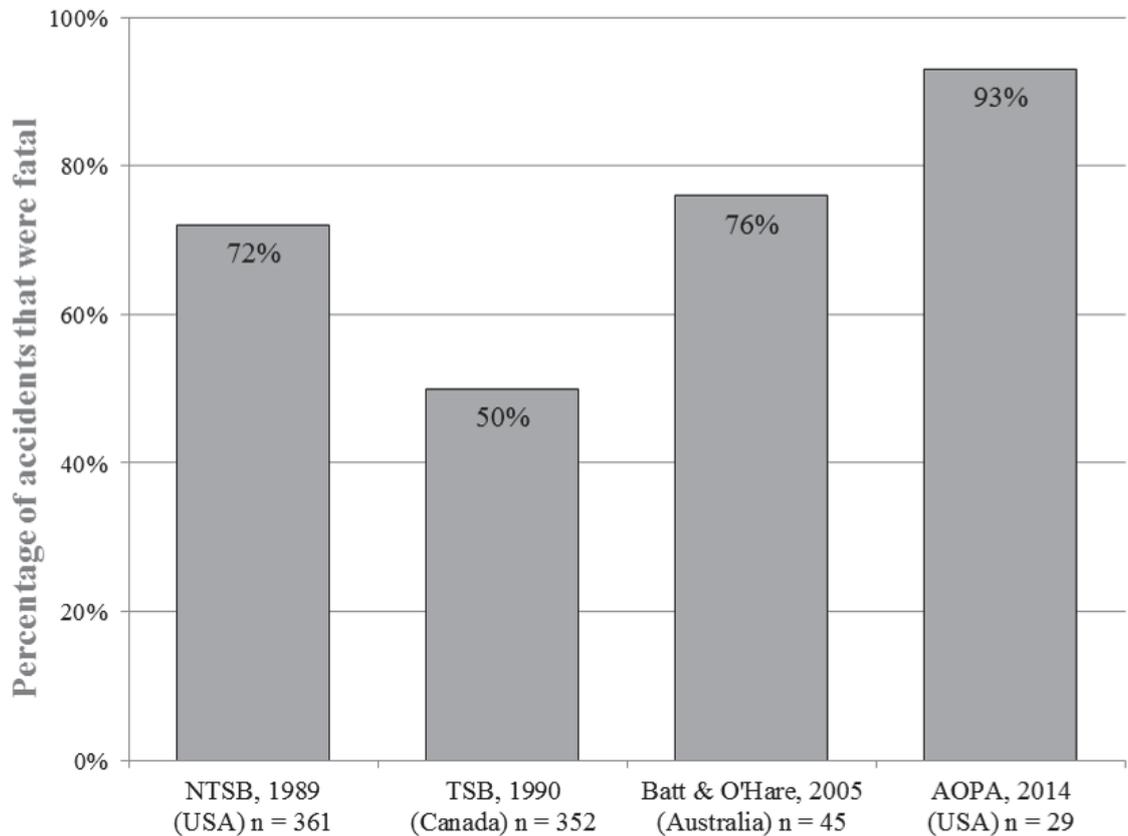
A pilot was conducting a local flight in a Cessna 150 with a single passenger in East Yorkshire, UK, in June 2014. The weather forecast was generally fine, with a 40% chance of reduced visibility (6000 m) in light drizzle. The initial phase of the flight was uneventful; however, on the return to the departure airport, the weather began to deteriorate. The pilot elected to make a precautionary landing in a field of crops about 10 nm from the airport. During the landing, the nose wheel dug into the ground and collapsed, pitching the aircraft forward onto its nose. The aircraft was substantially damaged but, fortunately, the pilot and passenger escaped serious injury (Air Accident Investigation Branch [AAIB], 2014).

VFR flight into IMC is generally a problem for pilots operating smaller aircraft in the GA environment. GA refers to aircraft operating on non-commercial flights, which can include aircraft of a range of sizes, with four- to six-seater aircraft (e.g., a Cessna 172 with four seats) being a relatively common aircraft type (AOPA, 2014). A pilot operating under a PPL is broadly categorised as flying in the GA environment (PPL holders are not able to operate for hire or reward); however, CPL or ATPL holders can still operate in the GA environment. The type of flight that a GA pilot would operate can vary from conducting a local flight lasting little more than half an hour to a cross-country flight over various terrain types over several hours.

VFR flight into IMC has also been highlighted as a concern for some small commercial operators. Small commercial operators typically carry fare-paying passengers in aircraft up to 10 seat on a mix of VFR and IFR operations; from sight-seeing to short charter flights. A report by the TSB (1990) found 27% of VFR flight into IMC accidents involved charter operations. Pilots operating larger aircraft, (e.g., a Boeing 747 carrying 300 passengers) primarily operate under IFR rules, even on a clear day, so therefore the pilots are qualified and the aircraft are equipped to operate in IMC.

2.4 Historical Trends in VFR Flight into IMC Accidents

A substantial body of research has highlighted the significant danger associated with VFR pilots flying into IMC. Figure 2 shows the fatality rate for VFR flight into IMC accidents across a range of studies and across a range of time periods.



TSB, Transport Safety Board

Figure 2: Fatality rate for VFR flight into IMC accidents across a range of studies (countries) and time periods.

One of the most alarming aspects of this type of accident is that no significant improvement has been seen in this accident rate since the first major VFR flight into IMC study by the NTSB (1989). The NTSB studied data for the 10-year period prior to 1986 and found that 72% of VFR flight into IMC accidents were fatal, compared

with 17% for all GA accidents. More recent studies (AOPA, 2014; Goh & Wiegmann, 2001; Transport Safety Board of Canada [TSB], 1990) have found similar and, in some cases, higher accident rates for VFR flight into IMC. Furthermore, the fatality rate of VFR flight into IMC is not limited to the US, with other countries (Australia, Canada, United Kingdom) with varying metrological and geographic conditions also showing high fatality and accident rates (Batt & O'Hare, 2005; O'Hare & Smitheram, 1995).

These descriptive studies have also provided an insight into the pilot characteristics and flight situations within which this type of accident typically occurs. The pilot involved is likely to have low flight experience, typically less than 500 hours of flight experience. The pilot is likely to hold a PPL and unlikely to hold an IR (Goh & Wiegmann, 2001; NTSB, 1989). Furthermore, these accidents are more likely to occur in the cruise phase of flight (Detwiler, Holcomb, Hackworth, & Shappell, 2008; Goh & Wiegmann, 2001) and in the second half of the flight, regardless of the length of the flight (Batt & O'Hare, 2005). These characteristics can be seen in the following fatal accident report:

The pilot, the sole occupant on board, set off in a Cessna 182 (single-engine, four-seater light aircraft) on a 90-min flight in New South Wales, Australia, in 2012. The pilot held a Private Pilot Licence, with no instrument rating and relatively low flight experience (~300 h). The flight departed in generally good weather (high cloud), but as the flight progressed, the pilot was forced lower with a lowering cloud base, down to about 1000 ft above terrain. Just after the halfway point in the flight, with the aircraft flying inside the cloud, the aircraft

impacted a rock face in mountainous terrain (Australian Transport Safety Bureau [ATSB], 2013).

More recent descriptive studies have been able to capture a wide range of personal and situational characteristics in an attempt to discover potential indicators of this type of accident. For example, Wiggins et al. (2012) explored the characteristics of pilots who deliberately versus inadvertently flew VFR into IMC. Pilots that had deliberately entered IMC tended to have experienced flying in these conditions previously and tended to have a greater tolerance of risk compared to pilots that inadvertently flew into IMC.

Furthermore, the current literature in this area (e.g., Hunter et al., 2011; Wiggin et al., 2012) have highlighted the importance of addressing the problem of VFR flight into IMC from a range of perspectives. The current approaches to addressing the VFR flight into IMC problem are discussed in the next section.

2.5 Primary Cause: Decision Error

If one considers the consequences that can arise from VFR flight into IMC, it is not surprising that research has focused on the underlying cause of this type of accident. Research exploring the possible causes of this type of accident has concentrated on issues relating to decision making. Ineffective or inappropriate pilot decision making has been highlighted as a significant factor in this type of accident, focusing on psychological factors in pilots' decision making strategies in adverse weather (Wiggins, Martinussen, & Hunter, 1999).

The focus on decision making stems from how faulty pilot decision making has been identified as the primary cause in this type of accident (Madhavan & Lacson, 2006). It is important to note that weather-related decision making often occurs simultaneously with time pressure (e.g., in GA, only so much time can pass before fuel may run low or the flight environment may change), few and ambiguous cues and limited experience (Sarter & Schroeder, 2001). It is reasonable to assume that decisions made in this environment are therefore both quantitatively and qualitatively different and more complex than, for example, those made by drivers who find themselves driving into foggy conditions but can pull over to the side of the road and await better conditions. The ability to make a sound decision can be further complicated by extra workload, motivational factors and stress, all of which can be present when flying near IMC, in addition to simply flying the aircraft and keeping a good lookout (Orasanu, Martin, & Davison, 2009; O'Hare, Wiggins, Batt, & Morrison, 1994).

Several factors have been proposed to explain why pilots make inappropriate decisions when approaching IMC, including risk perception, motivational factors, impaired situational assessment and cognitive biases. It should be noted that a VFR flight into IMC event is likely to be the result of an interaction of a number of these factors, rather than one of these factors in isolation (Goh & Wiegmann, 2001).

2.5.1 Risk Perception

Due to the uncertain nature of weather-related decision making, a level of risk is involved (Madhavan & Lacson, 2006). Studies into pilots' risk perception have

found that pilots who continue flight into adverse weather may have inaccurate perception of risk (Hunter, 2002). In other words, pilots are not necessarily willing to take greater risks in hazardous situations, but they may not accurately realise the risk of continuing the flight as their perception of the risk has been influenced (Madhavan & Lacson, 2006). A considerable amount of research into pilot risk perception (e.g., Hunter, 2006; O'Hare, 1990) has been conducted using various risk assessment tools (e.g., the aeronautical risk judgment questionnaires developed by O'Hare [1990]). These tools have explored a range of risk perception measures, such as the perceptions of the pilots' own ability to operate around adverse weather and hazard awareness. These studies have found that pilots exhibit a low level of risk awareness combined with high levels of self-appraised ability (Madhavan & Lacson, 2006). Pauley, O'Hare, and Wiggins (2008) further explored risk perception within the context of Lope's (1987) Theory. Lope's (1987) Theory suggested that risk may be governed by sensitivity to potential gains. Using a series of studies, Pauley et al. (2008) found evidence that some pilots make trade-offs between risk and returns.

It is also thought risk perception may change the more that someone is exposed to a risky situation that results in a successful outcome. This was highlighted in the following VFR flight in IMC accident:

A Twin Otter (a 19-seat, twin-engine aircraft) was conducting a VFR approach in adverse weather in British Columbia, Canada, in September 1995. It was common practice for pilots to attempt to find a way to the airport in bad weather by flying at low altitudes to 'take a look'. The pilot had successfully completed the manoeuvre on many occasions. The accident report highlighted that the more times such risky manoeuvres are completed, the more likely the

pilot is to believe that nothing bad will happen to him or her, leading to a level of comfort that is likely to reduce safety margins. Unfortunately in this flight, this type of manoeuvre resulted in the pilot flying into low visibility; the pilot progressively lost situational awareness and crashed into the side of a mountain. The pilot and seven of the passengers received fatal injuries (TSB, 1996).

2.5.2 Social and Motivational

Some social and motivational factors have been identified as contributing factors for VFR flight into IMC (Goh & Wiegmann, 2001). These factors are based around the concept that internal (personal) and external (social) pressures influence the decision of the pilot. For example 'Get-there-itis' is when the pilot places additional pressure on him- or herself to get home, even if the weather is not suitable (Goh & Wiegmann, 2001). Get-there-itis was highlighted in the following fatal VFR flight into IMC accident in New Zealand;

In April 2001, a pilot took off on a cross country flight to visit a friend who had taught him to fly. The weather forecast received prior to flight indicated fine and clear weather conditions, both enroute and at the planned destination. As the aircraft approached the halfway stage of the flight the weather began to deteriorate. The accident report suggested the pilot likely pressed on into deteriorating weather conditions rather than turn back or divert, with the pilots decision making influenced by the get-there-itis. The pilot was likely determined to continue with the planned flight despite changing conditions.

The further the pilot pressed on, the more likely the pilot felt pressured to continue, which results in a circular and vacuous cycle (CAA NZ, 2012).

Motivational factors have also been linked with some cognitive biases (e.g., the ‘sunk cost’ effect). The sunk cost effect describes the situation whereby the more an individual invests in a task, the more pressure he or she will feel to complete the task (O’Hare & Owen, 2001). ‘Get-there-it is’ is also referred to as plan continuation bias or plan continuous error (Wiggins, Azar, Hawken, Loveday, & Newman, 2014). Cognitive biases in weather-related decision making will be discussed in more detail shortly.

2.5.3 Situational Assessment

Recent studies have supported a situational assessment explanation for a pilot’s decision to continue flight towards IMC (e.g., Goh & Wiegmann, 2001; O’Hare et al., 2011). The situational assessment explanation suggests that a pilot’s decision to continue flight into IMC may be due to impaired awareness of the information used in the decision (O’Hare et al., 2011). If the pilot has accurately recognised he or she is flying towards conditions unsuitable for VFR flight, he or she is likely to divert (Wiegmann et al., 2002).

The importance of accurate perception of the conditions was highlighted by Goh and Wiegmann (2001), who found that accurate perceptions of visibility were among the most important factors that distinguished between pilots that would continue and pilots who would divert in a VFR flight into IMC scenario. Goh and Wiegmann

suggested that pilots who continued to fly below VFR meteorological criteria may have done so because of an inaccurate perception of the situation. Pilots who decided to continue the simulated flight overestimated the visibility (conditions were worse than the pilot believed) and pilots who chose to divert accurately assessed the visibility. This situation corresponds with Rasmussen's (1983) notion of rule-based behaviour, which describes a response based on a set of rules. In this situation, if a pilot correctly assesses the visibility as too low, they would be likely to divert, which was the case in the diversion group in the Goh and Wiegmann (2001) study. Coyne, Baldwin, and Latorella (2008) found further support for the situational awareness theory. They found that pilots have problems estimating weather conditions, with both IR and non-IR pilots represented among those with poor ability to assess weather conditions adequately.

Inaccurate perception of weather conditions was highlighted as a contributing factor in a VFR flight in IMC accident in Canada:

On 16 November 2008 a Grumman G-12 departed on a VFR flight with eight people on board, initially tracking over a lake. The accident report suggests that the pilot's initial decision to depart may have been influenced by inaccurate weather information, which indicated marginal VFR conditions. Once the aircraft was airborne, the visibility was, in fact, significantly below VFR conditions. However, the pilot may have had difficulty accurately assessing visibility over the featureless water surface. Several other factors were highlighted as influencing the pilot's decision to continue flight from VFR into IMC including: previous success in low visibility and a commitment to continue with the chosen course. When the pilot finally sighted terrain on the

other side of the lake, the aircraft was too close to take evasive action and crashed, killing seven of those on board (TSB, 2010a).

2.6 Cognitive Heuristics and Bias in Decision Making

Coyne et al. (2008) suggested that impaired situational awareness provides one of the most plausible explanations for poor decision making in VFR flight into IMC. However, only limited research has been conducted into understanding the underlying cause of this impaired situational awareness (Goh & Wiegmann, 2001). In this thesis, the author sought to further understand the impaired situational awareness involved in pilots' weather-related decision making within the framework of cognitive heuristics and biases. Cognitive biases have been identified as a significant factor in decision error in a range of fields (e.g., Croskerry, 2002; Englich et al., 2006) but appear to have received limited attention in pilots' decision making.

2.6.1 Characteristics of Aeronautical Decision Making

A considerable body of literature has explored human decision making, however; research specifically exploring aeronautical decision making is relatively new (Kochan, Jensen, Chubb, & Hunter, 1997). One of the first studies investigating aeronautical decision making was conducted by Thorpe, Martin, Edwards, and Eddows (1976), closely followed by Jensen and Benel (1977). These studies began the development of pilot judgment training, with an initial focus on applying literature from other areas to aviation. Investigating aeronautical decision making

became particularly important due to poor decision making being increasingly highlighted as a considerable factor in aviation accidents (Kochan et al., 1997).

The next stage in aeronautical decision making research focused on providing pilots with decision making tools. One of the first aeronautical decision making models was developed by Jensen, Adrion, and Maresh (1986) called 'DECIDE'. The DECIDE model provided a clear path to be used during aeronautical decision making (Kochan et al., 1997). Jensen (1995) further expanded on this early work by creating a framework for aeronautical decision making. Jensen (1995) judgment model, details the eight judgment process that should take place in aeronautical decision making;

- a) *Problem vigil*: pilot maintains a constant stage of vigil so that changes of the environment can be detected
- b) *Recognition* – pilot realises changes in the environment
- c) *Diagnose* – pilot attempts to understand the nature of the problem
- d) *Alternative identification* – pilot identifies various alternatives
- e) *Risk assessment* – pilot tries to determine risks associated with each alternative
- f) *Background factors* – pilot decision influenced by personal and social pressures
- g) *Decision-making* – the pilot chooses the final course of action
- h) *Action* – the pilot applies the decision

Wickens and Hollands' (2000) information processing model of decision making also provides a similar framework to describe aeronautical decision making. The focus of these models was to provide a path to support optimal decision making. These models have been simplified into three categories, which have been presented by Madhavan and Lacson (2006) as shown in Figure 3 (Hutton & Klein [1999] and Nagel [1988] simplified the analytical decision making model in a similar manner).

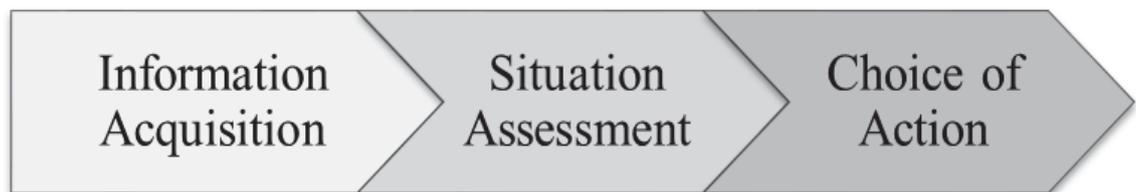


Figure 3: Simplified decision making model (Hutton & Klein, 1999; Madhavan & Lacson, 2006).

The first stage in achieving an optimal decision starts with *information acquisition*, which encompasses seeking and acquiring cues. The second stage, *situation assessment*, involves attempting to understand the problem, identifying various alternatives and determining the risk of each alternative. This is followed by *choice of action*: selecting and applying the most appropriate alternative.

There has also been particular interest in how experienced pilots process information during weather-related decision making. Because of the nature of weather-related decision making, it is often considered to be a skill that a pilot will gradually develop as he or she is exposed to different weather-related environments. If a pilot is to obtain these skills through experience, they may put themselves into danger in the

process. Giffin and Rockwell (1984) conducted early studies exploring the role of pilot characteristic on aeronautical decision making, they found less experienced pilots tended to use larger amounts of information. Wiggins and O'Hare (1995) explored weather related decision making between expert and novice pilots. Experts were significant more efficient than novice pilots in terms of the acquisition and integration of information used during weather-related decision making. They found expert decision makers have a particular advantage when exposed to the same cues, which enables them to assess the situation automatically.

Various aeronautical decision making teaching techniques have been explored. O'Hare, Mullen, and Arnold (2010) investigated case-based reflection on aeronautical decision making. Participants that reflected on cases involving pilots flying into adverse weather were more likely to avoid flying into such conditions themselves. Wiggins & O'Hare (2003b) developed a Computer Based Training program called 'weatherwise', which aided pilots to recognise key cues during deteriorating weather conditions.

Goh and Wiegmann (2001) suggested that pilots who continue VFR flight into IMC make errors in the early phase of the decision making process, and this is compounded by other factors such as risk perception and over-confidence in their flight skill. Some (but not all) cognitive biases occur in the first stage of decision making (e.g., anchoring during the information acquisition stage [Madhavan & Lacson, 2006]). Cognitive biases that occur later in the decision making process influence how and what information is used (e.g., decision framing and confirmation bias).

2.6.2 Cognitive biases in Decision Making

In principle, using the decision making model in Figure 3, when pilots are considering the weather ahead, they should carefully consider all the available information before making a judgment. In other words, they should, collect all the available information (*information acquisition*) from a range of sources (e.g., visual, text-based weather forecasts, observations by other pilots), assess the information (*situation assessment*) to determine what the best course of action is (e.g., to continue, divert or turn around) and then select the most appropriate action (*choice of action*) based on the information acquired.

In practice, not all evidence required to make a judicious decision is always used. The evidence may not be available, the time in which to make a decision may be too short or perhaps the evidence is too readily discounted or simply ignored. Indeed, Wiggins and O'Hare (1995) indicated that due to the nature of weather-related decision making, a full analytical strategy is less likely to be effective due to the environmental constraints placed on pilots, particularly time pressure. The evidence overwhelmingly suggests that humans do not always evaluate evidence impartially or fully, but use cognitive heuristics instead (Kahneman, 2011; Tversky & Kahneman, 1974; Gigerenzer & Gaissmaier, 2011). Heuristics are a useful decision making tool which work well most of the time, however if heuristics are used inappropriately cognitive biases can arise (Hutton & Klein, 1999).

These cognitive heuristics (shortcuts) are likely to be the result of the limitations in people's decision making capacity, which is best understood by considering the dual

processing theory [DPT] (Epley & Gilovich, 2001; Kahneman, 2011). The dual processing theory provides an insight into how the human brain deals with these limitations when processing information (Croskerry, Singhal, & Mamede, 2013). When making decisions, people use one of two modes (systems), which can be either Type 1 (intuitive) or Type 2 (analytical) processes (Croskerry et al., 2013; Kahneman, 2011). Type 2 processes are reliable and effective. These processes are rule-based and conducted under conscious control following a similar procedure to Figure 3. On the other hand, they take up a large amount of cognitive load.

To ease cognitive workload, many of our decisions are made using Type 1 processes. Type 1 processes are largely automatic, fast and *usually* effective. They use a number of shortcuts or heuristics to find adequate, though potentially imperfect, answers to what may be difficult and complicated questions (Gigerenzer & Gaissmaier, 2011), a notion captured by Simon's (1956) term, *satisficing*. It is thought that we spend about 95% of our time in this mode (Croskerry, 2014). However, inappropriate use of heuristics can lead to systematic errors (biases), which essentially deviate from what would reasonably be considered rational or good judgment (Kahneman, 2011; Tversky & Kahneman, 1974). Figure 4 provides a summary of the processes in the DPT. In the case of VFR flight into IMC, the 'rational judgment' would be to remain clear of IMC condition. Indeed, VFR pilots are trained and required to remain clear of IMC conditions. Moreover, the dangers of entering IMC conditions are well documented. Therefore, the 'deviation from rational judgment' would be for a VFR pilot to enter IMC conditions.

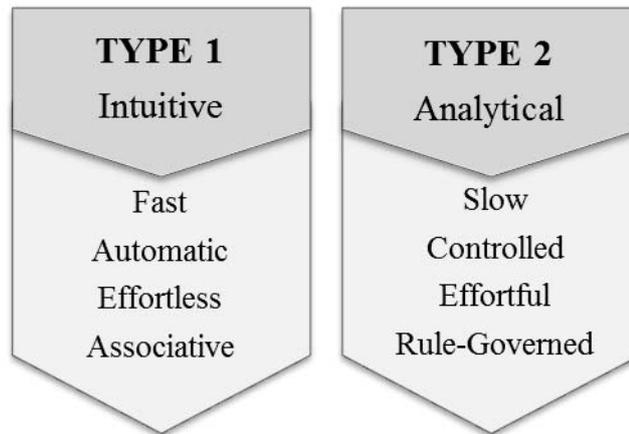


Figure 4: Characteristics of Type 1 and Type 2 processes in the Dual Process Theory [DPT] of decision making (Kahneman, 2011)

2.7 Cognitive Bias in Weather-Related Decision Making

Numerous cognitive biases and heuristics have been identified. For example, Dobelli (2013) reviewed over 100 biases in his recent book and at least 40 biases have been found to affect clinical reasoning (Mamede et al., 2010). Only a handful of cognitive biases have been explored in a pilot's decision to continue VFR flight into IMC. These include decision framing, the sunk cost effect, optimistic bias and ability bias.

2.7.1 Decision Framing

According to Tversky and Kahneman (1981), a person's decision can be influenced depending on whether an option is framed as a gain or a loss. In the case of VFR flight into IMC, if a person frames the decision to divert as a loss (i.e., a waste of money or time) the pilot may have a greater tendency to continue. If the pilot frames the decision as a gain (i.e., it is safer), then he or she may be more likely to divert.

Some support for decision framing was found by O'Hare and Smitheram (1995). In their study, pilots were less likely to continue a flight into IMC when they were encouraged to frame the VFR flight into IMC as a loss and diverting in terms of a gain. This study also suggested that decision frames may change when a pilot gets closer to their destination, with a natural shift towards the loss frame (Goh & Wiegmann, 2001).

2.7.2 The Sunk Cost Effect

The sunk cost effect suggests that a person is more likely to continue with an action if he or she has already made a considerable investment of time or money (Arkes & Blumer, 1985). Support for the sunk cost effect has come from examining the location of VFR flight into IMC accidents. VFR flight into IMC accidents are more likely to occur in the second half of a flight (i.e., closer to the destination), by which point the pilot has already invested a significant amount of time and resources into the flight. The effects of sunk costs are likely to have a greater influence the closer the pilot gets to his/her destination, similar to decision framing (Batt & O'Hare, 2005; O'Hare & Owen, 2001).

2.7.3 Optimistic and Ability Bias

Wilson and Fallshore (2001) explored optimistic bias and ability bias in pilots' decisions and perceptions of risk regarding continued VFR flight into IMC. Both these concepts are linked with risk perception, which was discussed earlier.

Optimistic bias refers to the tendency to believe that we have a better chance of success or a lower chance of failure than our peers (Weinstein, 1980). Ability bias refers to most people believing themselves to be superior to others in their skills and ability (Dunning, Meyerowitz, & Holzberg, 1989). Wilson and Fallshore (2001) found evidence to suggest that both may be present in pilots' decision making, with participants under-estimating their likelihood of experiencing an accident caused by VFR flight into IMC and over-estimated both their ability to avoid IMC and ability to successfully fly out of IMC.

2.8 Cognitive Bias Explored in this Thesis

A review of the literature and anecdotal evidence from aircraft accident reports highlighted a number of potential cognitive biases that could occur in pilots' decision making during a VFR into IMC event. Although there are potentially a large number of cognitive biases that could be investigated, in this thesis, the author aimed to explore three cognitive biases that meet the following criteria: (i) the cognitive bias could potentially lead a pilot to perceive the weather conditions to be better than they are in reality, (ii) the bias has been highlighted as a potential factor in aircraft accidents, and (iii) the bias has not received any significant attention in the VFR flight in IMC literature.

The three cognitive biases explored in this thesis were the *anchoring effect*, *confirmation bias* and *outcome bias*. The anchoring effect is defined as the tendency of relying too heavily on the first piece of information received when making decisions and under-adjusting when making the final judgment (Hastie & Dawes,

2010). Confirmation bias describes the tendency to seek out or interpret evidence in a way that favours an existing belief, expectation or hypothesis (Gilbey & Hill, 2012). Outcome bias involves assessing the quality of a decision by whether its outcome was good or bad, which can compromise what and how we learn from an event, potentially creating bias in subsequent decisions (Fischhoff, 1975; Kahneman, 2011).

2.8.1 The Anchoring Effect

When making decisions, people often make estimates based on a starting point or an anchor (Tversky & Kahneman, 1974). The initial piece of information (an anchor) may vary in the degree to which it is useful to the decision in hand; for example, it could be a weather forecast. It could also be information that is completely irrelevant to the decision (Englich et al., 2006). Although this heuristic could be useful to reduce cognitive workload, especially when making complex decisions, the evidence suggests that people often fail to make appropriate adjustments from the initial value of the anchor even when conditions are likely to have changed. Failure to reassess one's initial judgment leaves the final judgment biased towards the initial value (Epley & Gilovich, 2006; Tversky & Kahneman, 1974). In the context of pilots' decision making, placing too great an emphasis on earlier information may prove particularly hazardous, as weather conditions are highly dynamic (Wagtendonk, 2011).

One of the first studies on anchoring and adjustment was conducted by Tversky and Kahneman (1974). Tversky and Kahneman's (1974) seminal study asked students to estimate the percentage of African countries in the United Nations (35% at the time

of the study). Prior to making their judgment, students were asked to judge if the number (of African countries) was greater or less than a number determined by spinning a rigged wheel. Students who judged that the number was greater than or less than the 10 selected by the spinner made an average estimate of 25%; those who judged that the number was greater or less than 65 made an average estimate of 45%. In this study, 10 and 65 served as *anchors*. Even though these numbers were generated in an arbitrary manner, participants focused on the anchor value (the rigged number) at the start of their judgment process, and the final adjustment was insufficiently corrected away from the anchor (Hastie & Dawes, 2010; Kahneman, 2011).

Anchoring and adjustment has been demonstrated in a range of areas from estimates of risk and uncertainty (Plous, 1989; Wright & Anderson, 1989) to evaluations of gambles (Cervone & Peake, 1986; Chapman & Johnson, 1994) and many more (see Epley & Gilovich, 2001). Recent research (Epley & Gilovich, 2006) has focused on the different types of anchor and how this influences the adjustment process. Differences have been observed between two main types of anchor: experimenter-provided anchors (like the rigged wheel example) or self-generated anchors. When people are faced with an experimenter-provided (or externally sourced) anchor, they first assess whether the anchor's value might be correct before adjusting their judgment. Adjustment in this situation is thought to be produced by the increased accessibility of anchor-consistent information. People tend to search for confirming evidence to support the initial anchor, which then biases their absolute judgment (Epley & Gilovich, 2001).

Alternatively, when people use self-generated anchors, they generate an anchor that they know is wrong as a starting point and adjust away from it. A classic example of a self-generated anchor can be seen in a study that asked Americans when George Washington was elected president of the US. Most Americans in this study did not know the exact date but self-anchored on the date of the Declaration of Independence in 1776. In self-anchoring situations like this one, people know the anchor is not correct, so therefore people do not focus on the self-anchor but instead focus on how far to adjust from this anchor. In this situation, people often under-adjusted (Epley & Gilovich, 2006).

Importantly, the anchoring effect is not limited to arbitrary laboratory questions or people with little experience in a particular field. Northcraft and Neale (1987) asked professional real estate agents to assess the fair market values of properties, a task which they would have conducted hundreds of times. The real estate agents were provided with a 10-page summary of the house details and visited each property. Unknown to the participants, the study manipulated the originally listed value of the properties by $\pm 12\%$. This should not have had an effect on the fair value of the house. However, the study found that the manipulated value had a 'consistent and large effect on the professional real estate agent's appraisals'. Similarly, Englich et al. (2006) studied the anchoring effect in judicial sentencing and found that sentencing decisions were influenced by irrelevant sentencing demands even if they were blatantly determined at random (one of which was produced by the participants throwing a dice themselves).

The finding that even trained and experienced professionals in a particular field are subject to the anchoring effect has potentially serious implications in other fields, such as aviation (or medicine). Pilots receive considerable training and experienced pilots would have assessed the weather conditions countless times; however, it is possible that this may not protect them from the anchoring effect. Although this heuristic can be useful, if pilots anchor on information that is no longer valid and merely make small adjustments, this could result in pilots forming an inaccurate perception of the weather conditions.

2.8.2 Confirmation Bias

Confirmation bias is the tendency to seek out or interpret evidence in a way that favours an existing belief, expectation or hypothesis (Gilbey & Hill, 2012; Muthard & Wickens, 2001; Nickerson, 1998). In some situations, confirmation bias also results in unintentionally ignoring evidence which contradicts (disconfirms) the favoured hypothesis (Green et al., 1996). In the context of pilots' decision making, particularly which occurs prior to a VFR flight into IMC, focusing on confirmatory information may lead a pilot to interpret the conditions as being better than they really are. Confirmation bias could potentially support a more pessimistic assessment. For example, supporting the belief that the weather conditions are not suitable to fly, when in fact they are, resulting in the pilot diverting unnecessary. However, in the context of this thesis, exploring confirmation biases towards a more optimistic outcome is the primary concern as it could increase the risk of VFR flight into IMC.

Confirmation bias has been observed in a wide range of areas (Koslowski & Maqueda, 1993; Nickerson, 1998). One of the earliest demonstrations of the tendency to seek confirmatory evidence was conducted by Wason (1960). In this study, people were asked to find a rule that classified a set of three, with 2–4–6 being one example provided that conformed with the rule. Participants were instructed to try to work out the rule by suggesting further sets of triples to the researcher, who would provide feedback. This study found that in general, most participants tested their hypothesis using a confirmatory strategy (examples that confirmed with their hypothesis); for example, many participants first tested 8–10–12. This fitted a subset of the correct rule (which was increasing numbers) and therefore received positive feedback. However, this positive feedback often led participants to hypothesise the rule to be increasing consecutive even numbers. Positive feedback in this situation led participants to erroneously believe that their current hypothesis was correct.

Confirmation bias has also been shown to occur in applied areas of research such as medicine and criminal investigation. Studies by Pines (2006) and Croskerry (2002) highlighted the dangers confirmation bias can have in medicine, especially emergency medicine. Pines (2006) argued that emergency medicine often relies on heuristics, but this can lead to potentially misdiagnosing a patient and therefore providing inappropriate treatment. Criminal investigation is another field where confirmation bias has received considerable attention. Ask and Granhag (2005) found evidence of confirmation bias when processing crime-related information. Similarly, Kassin, Goldstein, and Savitsky (2003) found that the presumption of guilt that underlines the police interrogation process sets in motion the confirmation bias process.

Confirmation bias has been used to explain a wide range of phenomena. It has been argued (Evans, 2006) that some situations would be better explained by the type of hypothesis testing strategy a person uses. Klayman and Ha (1987) explored the *positive testing strategy*, which refers to the situation where people seek out information that confirms their initial hypothesis. The problem of using a *positive testing strategy* as a general default heuristic is that in some situations, this strategy may result in the confirmatory evidence supporting the proposed hypothesis when, in fact, it is incorrect. It may be possible to find several pieces of evidence to support one's initial hypothesis, even when it is wrong. It has been argued that a disconfirmatory strategy or *negative testing* is a better testing method (Wason, 1960, 1966). This strategy involves seeking out evidence which disproves one's chosen hypothesis, which is potentially more beneficial. Humans have a natural tendency to look for evidence that supports their hypothesis (Nickerson, 1998). However, this method of seeking out evidence to support one's hypothesis can lead to an error of judgment.

Despite confirmation bias having been identified as a causal factor in aviation accidents (see CAA NZ, 2005b; NTSB, 2008), very few experimental studies of confirmation bias in aviation have been reported. That said, Gilbey and Hill (2012) investigated confirmation bias in GA procedures involving lost pilots/aircraft. These authors reported that when participants were asked to establish their location when lost, as GA pilots do when they are unsure of their position on a chart, they exhibited a preference for using a confirmatory strategy rather than a disconfirmatory one. They found that pilots favoured evidence that was consistent with their 'best guess' as to their location over evidence that suggested their best guess may be wrong (in

each of Gilbey and Hill's tasks, all of the 'best guesses' provided by the researchers were actually wrong).

Weather-related decision making can be difficult due to the ambiguous and dynamic nature of the information available. Particularly in countries such as New Zealand, where the weather conditions are highly variable, for example, the weather can shift from fog to bright sunny skies and then to heavy rain all in short space of time (O'Hare & Waite, 2012). Indeed, the ambiguity of information has been reported to have been one of the starting points of many accidents (Green et al., 1996). When pilots assess weather conditions, which they are likely to do throughout all flights, they search for a range of weather-related cues to assist their judgment (Wiggins & O'Hare, 1995, 2003). When a pilot makes a judgment about the weather—for example, that it is suitable to continue to fly towards their destination—he or she must seek out evidence to test this judgment. The evidence they seek or believe to be useful will depend on what type of hypothesis testing strategy they apply: a positive testing strategy in which they seek confirmatory evidence to 'prove' that their judgment is correct (e.g., the cloud base is high enough), or a negative testing strategy, seeking disconfirming or contradictory evidence to prove that their judgment is incorrect (e.g., the visibility is poor). There is nothing inherently wrong in using the positive testing strategy. However, an over-reliance on the positive testing strategy can lead to serious consequences.

2.8.3 Outcome Bias

Outcome bias is a cognitive bias which refers to the tendency to judge a decision by its eventual outcome instead of judging it based on the quality of the decision at the time it was made (Baron & Hershey, 1988). A limitation of the brain is the challenge it faces when reconstructing past states of knowledge once they have changed. Once someone is exposed to outcome information, it can be very difficult to recall what he or she believed beforehand (Kahneman, 2011). Although the use of outcome information can be useful, outcome information does not necessarily reflect whether the decision process was sound. Bad outcomes can result from good decisions and vice versa. Some situations may justify the use of outcome information, such as when the full context of the decision is not known. However, it has been found in these situations that people may give outcome information more importance than it deserves (Baron & Hershey, 1988).

Outcome bias has been observed in a wide range of judgments. One of the first studies to observe the influence outcome information had on judgments was that of Fischhoff (1975). Fischhoff conducted a series of studies exploring the 'I knew it all along' effect or hindsight bias, in which after an event has occurred, there is a tendency to exaggerate the likelihood that would have been assessed for the event before it occurred. Fischhoff (1975) highlighted the challenges faced when dealing with outcome information, especially when the information available is limited. Baron and Hershey (1988) further explored the use of outcome information and labelled this tendency as outcome bias. Their research found that subjects rated a decision with a positive outcome as being superior to a decision with a negative

outcome, even when the information available to the decision maker was the same in both cases.

The medical profession has received significant attention in the area of outcome bias. Studies by Gruppen, Margolin, Wisdom, and Grum (1994) and Croskerry (2002) have highlighted the dangers outcome bias can have in medicine, especially emergency medicine. Like the aviation field, medical decisions may be made under time and resource limitations. Baron and Hershey (1988) demonstrated the outcome bias present in a surgical scenario. In a series of studies, they presented participants with a medical case in which a surgeon decided to do an operation with a known probability of success. Participants were provided with one of two outcomes: either the surgery was a success and the patient made a full recovery, or else a complication during surgery resulted in the patient dying. Logically, the outcome of a decision should not influence the quality of a decision. However, participants judged the surgeon to be a better decision maker when the operation succeeded (Baron, 2008). Similar results have been found in a number of studies (Alicke & Davis, 1989; Mazzocco, Alicke, & Davis, 2004) that have explored outcome bias in relation to attribution for harmful events.

Prior research on the effect of outcome information has focused on judgments about decision quality (Baron & Hershey, 1988; Gruppen et al., 1994). Nevertheless, outcome information can also compromise what may be learned from an event (Fischhoff, 1975). As in the case of medicine, it is important for the aviation field to learn from others; however, if perceptions of events are compromised by outcome information, one may perceive differently from reality. Outcome information can

give the feeling that one understands what the past is about, which may prevent a person from learning anything from it. A pilot may successfully fly through adverse weather but the pilot may have taken a significant risk. In this situation, the flight's safe outcome may have been more down to luck rather than good judgment. The outcome information of this flight may influence the decision process of other pilots. Outcome bias has a tendency to reward irresponsible risk seekers (Kahneman, 2011).

The influence of outcome information and the impact it has on pilots' decision making has been explored in the area of commercial aviation (Hyams, Fans, & Kuchar, 1998; Midkiff & Hansman, 1992). Unlike VFR flights, commercial airliners are equipped to operate in IMC and the pilots are qualified to enter IMC. Most of the time, the consequences are minimal, resulting in little more than an uncomfortable ride for the passengers and a washing for the aircraft. Occasionally, however, an aircraft flies through severe weather conditions (e.g., a thunderstorm), which can result in damage, incident and even catastrophe. Why flight crews fly through severe weather has been the focus of a number of studies (e.g., Rhoda & Pawlak, 1999a, 1999b), with outcome information from other flights highlighted as being a significant factor in the decision process to continue through or deviate around adverse weather.

Rhoda and Pawlak (1999a, 1999b) investigated the behaviour of flight crew around convective³ weather. They found that an arriving aircraft is more likely to fly into convective weather if it was a follower. That is, the aircraft was following another

³ Convective weather typically refers to thunderstorms. Thunderstorms are just one type of convective weather; other types include the stages of cloud build-up leading to the thunderstorm. The presence of convection currents in the air is the defining characteristic of this type of weather. Due to the rapidly changing nature of convective weather, the flight conditions can be highly variable.

one that had successfully flown through the convective weather in the preceding 10 minutes. This was particularly pronounced if the aircraft approached the convective weather close to the destination aerodrome. The dangers of following the behaviour of aircraft ahead while operating in a dynamic weather environment were highlighted in the Air France A340 accident in Toronto in 2005. The aircraft, with 309 people on board, approached Toronto as a severe thunderstorm moved across the airport, lashing the area with torrential rain, frequent lightning and strong shifting winds. With these cues present, the crew continued on the approach, like several aircraft in front. However, as the Air France aircraft came in to land, the wind suddenly shifted, resulting in the aircraft over-running the 9000-foot runway. Fortunately, the cabin crew acted quickly and evacuated everyone safely (TSB, 2007).

Several studies have found evidence that pilots place significant weight on reports from other aircraft that have just flown in the area. This type of report is sometimes referred to as a 'ride report' (Midkiff & Hansman, 1992). A ride report can be the result of direct communication (e.g., a pilot warning the aircraft behind of a bumpy ride), or indirect, such as a pilot overhearing on the radio that an aircraft has diverted around adverse weather. Ride reports may also result in social pressure, for example, a ride report may suggest an action (e.g., turn left for smoother flying conditions) which may further influence pilot decision making. This research highlighted the strong influence pilot reports have on other pilots' decision making. Considering the strong influence outcome information has on airline pilots' decision making, could outcome information about previous flights bias the decision process of VFR pilots approaching IMC?

2.9 Research Problem

The literature review of this thesis points to the existence of a research problem, namely that alongside other factors, impaired situational awareness has been identified as one of the most plausible reasons why VFR pilots continue flight into IMC. Despite this, little research has explored what can cause this impairment. Cognitive biases, which have been shown to exist in a range of fields, may help to explain why pilots make inappropriate or ineffective decisions when approaching IMC. A substantial body of evidence has identified cognitive biases as a significant factor in decision error in a range of fields (e.g., Croskerry, 2002; English et al., 2006) but this factor has received limited attention in pilots' weather-related decision making.

It is important to note that under some circumstances, a decision made using cognitive heuristics is quite reasonable, especially when the consequences of a wrong selection are minimal (Dobelli, 2013; Tversky & Kahneman, 1974). Heuristics work well most of the time and ease the cognitive load when making decisions (Hutton & Klein, 1999). However, because they are short cuts and result in the decision maker not carefully considering all the information available, they can lead to cognitive biases. Therefore, if heuristics are used inappropriately, pilots may be led into a false understanding of the weather environment they are operating in, potentially perceiving the conditions to be better than they are in reality (Wickens & Flach, 1988). In an aviation environment, especially when operating in a dynamic, time-pressured situation like the one VFR pilots experience when flying near IMC, if the wrong decision is made, the consequences can be extremely serious.

The author's aim in this thesis is therefore to identify if the heuristics anchoring effect, confirmation bias and outcome bias, occur in weather-related decision making, and hopes to provide guidance on possible interventions to reduce any effects. The contribution this thesis therefore makes is to apply existing theories of decision making to the setting of aviation in an attempt to improve safety by helping pilots avoid VFR flight into IMC.

2.10 Research Questions

A series of research questions were developed in order to examine the research problem. The hypotheses for each of these research questions are contained within each individual study.

2.10.1 Development of Research Question 1: The Anchoring Effect

Anecdotal evidence from aircraft accidents (e.g., CAA NZ, 2005a) suggests that the anchoring effect may influence a pilot's perception of the weather conditions. Although this heuristic can be useful to reduce cognitive workload, especially when making complex decisions, evidence suggests that people often fail to make sufficient adjustments from the initial value of the anchor even when conditions are likely to have changed (Epley & Gilovich, 2006; Tversky & Kahneman, 1974). Relying too greatly on outdated weather information was highlighted as a possible factor in a fatal accident in New Zealand (CAA NZ, 2005a). Prior to takeoff, the pilot obtained a weather forecast that indicated VFR conditions during the latter part of the planned flight. However, the actual conditions the pilot encountered were

significantly worse than forecasted, on which the pilot may have anchored and under-adjusted. This may have resulted in an incorrect perception of the actual conditions, leading the pilot to push on into unsuitable flight conditions. The pilot inadvertently flew into IMC, resulting in spatial disorientation. A short time later, the pilot crashed into the sea.

In weather-related decision making, particularly that which occurs prior to a VFR flight into IMC incident or accident, clinging to earlier information may prove particularly hazardous, as weather conditions are highly dynamic (Wagtendonk, 2011). Thus the first research question is as follows:

Research Question 1: Are pilots influenced by the anchoring effect when making weather-related decisions? Specifically, do participants anchor on outdated weather forecasts and under-adjust when assessing the actual weather conditions, and does the anchoring effect influence a pilot's decision to continue a VFR flight?

2.10.2 Development of Research Question 2: Confirmation Bias

Confirmation bias has been identified as a casual factor in aviation accidents (e.g., CAA NZ, 2005b; NTSB, 2008); however, very few experimental studies of confirmation bias in aviation have been reported. In the context of pilot decision making, particularly that which occurs prior to a VFR flight into IMC incident or accident, focusing on confirmatory information may lead a pilot to perceive the

conditions as being better than they are in reality. Thus the second research question is as follows:

Research Question 2: Does confirmation bias occur in weather-related decision making? Specifically, what type of hypothesis testing strategy do pilots favour when judging the weather conditions?

2.10.3 Development of Research Question 3: Outcome Bias

Outcome information can give the feeling that a person understands what the past is about, which may prevent him or her from learning anything from it. If someone's perception of events is compromised by outcome information, that person may perceive things as being different from reality. Are pilots' decisions to fly towards IMC influenced by the outcome information of other flights? This could be hazardous if a pilot's perception of the weather conditions is affected by outcome information, particularly if a safe outcome was largely the result of luck in a high-risk situation. Thus the third research question is as follows:

Research Question 3: Does outcome bias occur in weather-related decision making? Specifically, does outcome information influence how decision quality is rated? Does outcome information influence subsequent decisions?

2.10.4 Development of Research Questions 4–6: Debiasing

The findings from Research Questions 1–3 raised three additional research questions, which will be discussed in Chapter Six.

2.11 Study Overview

The research problem and associated research questions could be investigated within the framework of either positivist or interpretivist paradigms using either quantitative or qualitative methodology (e.g., questionnaires verse case-studies). As the author's aim was to investigate objectively the presence of cognitive biases in weather-related decision making, which can, in principle, be generalised beyond the samples under consideration, this research is more appropriately aligned with methodological assumptions of the positivist paradigm and a quantitative methodology.

A vignette-based methodology was adopted for each of the research questions, which is a common method used in aviation psychology research (e.g., Gilbey & Hill, 2012; Hunter et al., 2003; Wiggins & O'Hare, 2003a). Flight simulator studies have been conducted in this area (e.g., Wiegmann et al., 2002; Wiggins & O'Hare, 1995) but the cost of conducting the research by this means greatly exceeded the resources available for this thesis, especially in the quantity needed to obtain statistically significant sample sizes. Conducting this type of research in the real-world environment (i.e., placing participants in a VFR flight into IMC situation) would be unethical and unviable. Studies 1 to 3 were undertaken using an online questionnaire, which provided an ideal platform to investigate the research problem. The methods

used to explore each of the research questions will be described in more detail in each of the corresponding studies.

This thesis can be divided into two research lines: descriptive research and experimental research. Descriptive Studies 1 to 3 (Chapters Three, Four and Five) explored the existence of the three cognitive biases outlined in the research problem. The findings from these studies raised three additional research questions, which will be discussed in Chapter Six and which led to experimental Studies 4 to 6 (Chapters Seven, Eight and Nine).

CHAPTER THREE

STUDY 1

THE ANCHORING EFFECT IN VFR FLIGHT INTO IMC

3.1 Introduction

The aim of this first study was to investigate whether the anchoring effect occurs in weather-related decision making. If pilots do have a tendency to be influenced by the anchoring effect, placing undue weight on information that may have changed, this could influence their ability to assess the weather conditions accurately. Having an accurate assessment of the weather environment is an important aspect of determining if VFR flight is possible and therefore the anchoring effect could have serious implications on the safety of a flight.

Anchoring information may vary in the degree to which it is useful to the decision in hand; for example, it could be a weather forecast but it could also be information that is completely irrelevant to the decision. For example, English et al. (2006) found sentencing decisions were influenced by irrelevant information, one of which was produced by the participants throwing a dice themselves. Although this heuristic can be useful in reducing cognitive workload, especially when making complex decisions, the evidence suggests that people often fail to make sufficient adjustments

from the initial value of the anchor even when conditions are likely to have changed (Epley & Gilovich, 2006; Tversky & Kahneman, 1974). In the context of pilots' decision making, placing too great an emphasis on information that may no longer be valid may prove particularly hazardous, as weather conditions are highly dynamic (Wagtendonk, 2011).

Are pilots anchoring on weather information obtained early in a flight and under-adjusting when they consider the actual conditions? Relying too greatly on outdated weather information was highlighted as a possible factor in a fatal accident in New Zealand (CAA NZ, 2005a). Prior to takeoff, the pilot obtained the weather forecast, which indicated marginal VFR conditions during the latter part of the planned flight. However, the actual conditions encountered were significantly worse than forecasted.

‘This may have resulted in an incorrect perception of the actual condition that he would encounter en route and at his destination’ (CAA NZ, 2005a).

With the more favourable weather conditions potentially anchored in his mind, the pilot continued to fly into weather conditions that did not meet the meteorological minima for the intend VFR flight. The pilot inadvertently flew into IMC, resulting in spatial disorientation and a subsequent crash into the sea. A similar case occurred in a fatal VFR flight into IMC in Canada. The accident report suggests that the pilot's initial decision to depart may have been influenced by inaccurate weather information, which indicated weather conditions suitable for a VFR flight. Once the aircraft was airborne, the weather conditions were, in fact, significantly below VFR conditions. (TSB, 2010a).

Although there are potentially a large number of variables of interest, this present study focused upon cloud height and horizontal visibility distance, as these are two of the key cues that a pilot needs to consider when remaining in VFR flight. Indeed, cloud height and horizontal visibility have been identified as the cues experienced pilots use when assessing the weather (Wiggins & O'Hare, 2003). If the anchoring effect is a factor in this type of situation, pilots could judge their current conditions as being closer to their 'anchor' (the provided weather forecast) than to the actual conditions. This may result in a pilot provided with a 'poor' weather forecast diverting unnecessarily or vice versa: a pilot with a 'good' weather forecast continuing with the flight when the sensible option would be to divert or to alter course.

Pilots can obtain cloud height and visibility information from a number of sources. Prior to the flight, a common method to obtain this information is via an online aviation weather briefing (O'Hare & Waite, 2012). For example, in New Zealand, pilots are able to access aviation weather information via MetFlight (<http://metflight.metra.co.nz>). The aviation weather briefing would include a range of information, including current meteorological conditions at selected airports (METAR), forecasts for selected airports (TAF), a general area forecast and dangerous weather reports. A more advanced briefing may also include a number of charts, such as mean sea level pressure charts. Not all aerodromes will have weather information available. Typically, aerodromes that receive commercial traffic or high volumes of GA activity will have weather information available. Smaller aerodromes that receive low volumes of traffic are likely to have minimal weather information available. Some of these smaller aerodromes may have an automated weather

recording (similar to a METAR) but are unlikely to have a weather forecast (TAF) available.

How often this weather information is updated varies. For example, a METAR (actual weather conditions) is normally updated every 30-60 minutes, whereas a TAF (forecast) is updated every 8–12 hours. During a flight, a pilot can also receive updated weather information as highlighted above by requesting them from an appropriate ATC service. Some advanced cockpit displays are able to receive this information, although this is not currently common in GA aircraft. In flight, a pilot can also receive weather information on a dedicated radio frequency (Automatic Terminal Information Service [ATIS]); however, these are only at a limited number of aerodromes (primarily at airports with commercial traffic).

The nature of weather forecasting is such that the forecast conditions do not necessarily reflect the actual conditions. Generally, the difference between the forecast and actual weather conditions are minimal; however, under some dynamic weather conditions (e.g., convective or fog or low cloud⁴), the difference can be considerable. This leads to a variety of scenarios where natural weather-related anchors can transpire. For example, the following is a forecast (TAF) at Hamilton Airport (NZ) retrieved at 0930 local time on 4 June 2015⁵:

TAF NZHN 042021Z 0421/0512

VRB02KT 30KM FEW030

⁴ Convective weather essentially refers to thunderstorms. Along with fog or low cloud, these conditions can be unpredictable in regards to timing and intensity.

⁵ The full TAF forecast reads: on the 4th day of the month (June 2015), the forecast was issued at 2021Z (0821 New Zealand local time) for a valid period between 2100Z (on the 4th June) and 1200Z (on 5th June). Winds are forecasted to be variable at 2 kt, with a visibility of 30 km and few clouds (1-2 8th of he sky covered) at 3000 ft.

The forecast, which was issued at 0821 local time, indicates that along with light winds (variable 2-kt winds), the cloud height is forecasted to be 3000 ft above ground level and the visibility is 30 km. These conditions would likely be suitable for even a novice VFR pilot, with only a few patches of cloud and good visibility. A pilot intending to fly to Hamilton would expect conditions to be well above the VFR meteorological minima (5 km of visibility and a 500-ft cloud base). However, the actual conditions were significantly worse than forecasted. For example, the conditions at 0930 local time/2130UTC (1 hour after the forecast was issued) were:

METAR 042130Z 16002KT 13KM BKN002 08/08 Q1009⁶

The conditions read: 13 km visibility and a cloud height of 200 ft. These conditions are below the VFR meteorological minima and are significantly different from the forecast issued 1 h earlier. In this particular case, an updated weather forecast that indicated the lower cloud and poorer visibility was not issued for another hour. The unpredictable nature of weather forecasting was observed in the following fatal VFR flight into IMC accident:

On the evening of 17 November 2010, a Cessna M337B was conducting a VFR flight near Avon Park, Florida. The pilot received a weather briefing prior to flight that indicated generally fine conditions, with no indication of rain showers, thunderstorms or other hazardous weather. However, during the flight, the weather began to deteriorate, with periods of heavy rain and reduced visibility. The aircraft inadvertently entered a heavy rain shower and crashed a short time later, killing all three people on board (NTSB, 2012).

⁶ The full METAR (actual conditions) read: at 2130Z on the 4th day of the month, the wind is 160 degrees at 2 knots, visibility is 13 km, with cloud broken (5-7 8th of the sky covered) at 200 ft. Temperature, 8°C; dew point, 8°C; sea level pressure (QNH) at 1009 hPa.

Given the reliance that pilots place on information about cloud height and horizontal visibility (Wiggins & O'Hare, 2003a) and evidence of the pervasive nature of anchoring (e.g., Hastie & Dawes, 2010), the first hypothesis of this study is as follows:

Hypothesis 1a: An anchoring effect for earlier weather information will influence how current cloud height is assessed.

Hypothesis 1b: An anchoring effect for earlier weather information will influence how current horizontal visibility is assessed.

Hypothesis 1c: An anchoring effect for earlier weather information will influence how pilots assess the safety to continue a VFR flight.

The role that experience plays on weather-related decision making has been of particular interest. Because of the nature of weather-related decision making, it is often considered to be a skill that a pilot will gradually develop as he or she is exposed to different weather-related environments. If pilots are to obtain these skills through experience, they may put themselves into danger in the process. Understanding how experienced pilots make decisions may help train inexperienced pilots (Wiggins & O'Hare, 1995). Expert decision makers have a particular advantage when exposed to the same cues as encountered in previous weather-related decision making, which enables them to assess the situation automatically. Furthermore, experienced decision makers are able to retrieve experiences from their

long-term memory which have resulted in positive outcomes (Hutton & Klein, 1999; Wickens, Hollands, Banbury, & Parasuraman, 2013).

Weather-related decision making studies (Wiggins & O'Hare, 1995, 2003a) have found that experienced pilots generally make more efficient decisions, using fewer but more specific cues than inexperienced pilots. However, it has been found in some situations, experience does not always improve decisions (Wickens et al., 2013). Feedback plays a critical role in how one learns from a decision. For example, if the feedback from a decision is ambiguous or delayed, it can influence how people learn from an event. For example, when a pilot decides to go flying, it can be some time before the outcome of that decision is known. Research exploring how experience impacts the anchoring effect has found mixed results. Previous anchoring and adjustment research has found that experienced decision makers in a particular field still tend to be affected by the anchor effect. Therefore, consistent with other experience-based studies (Northcraft & Neale, 1987; Wiggins & O'Hare, 2003a, 1995), it is predicted that experienced pilots will be less affected by the anchor effect but will still show some bias.

Hypothesis 2: Experience will reduce the anchoring effect of cloud height, visibility and safety assessment compared to that of novice pilots.

Alongside an explanation of the role experience might have in the anchoring effect of weather information, it is also appropriate to consider some other variables that may provide an indicator that influences the anchoring effect. These include the effect of pilot age and the type of licence held. The type of pilot licence held is expected to

mirror the hypothesis of experience (Hypothesis 2), namely that the more advanced the pilot licence, the more experienced the pilot is likely to be. However, it is expected that age will not influence the anchoring effect (i.e., there will be no relationship between age and degree of anchoring), because age does not necessarily reflect experience or the licence held (NTSB, 2005). It is not uncommon in the GA environment that a pilot will conduct training at a later age. Pilots that conduct training at a later age may be more likely to pursue flying for pleasure rather than a career path; therefore, it is not uncommon for older pilots to have low experience levels and only hold a PPL.

Hypothesis 3: A more advanced pilot licence (i.e., CPL or ATPL) will result in pilots having a reduced anchor effect.

Hypothesis 4: Age will not influence the anchoring effect.

The primary aim of the current study was, therefore, to investigate whether the anchoring effect occurs in weather-related decision making. Additionally, this study investigated the effect of experience on the anchoring effect in weather-related decision making.

3.2 Method

3.2.1 Participants

Two hundred and one pilots participated in this study. The mean age of the participants was 34.79 years (standard deviation [*SD*] = 15.01) with a mean flying experience of 832 hours (*SD* = 1292, ranging from 2 to 8100 hours). Just over half (52.8%) of the participants held a PPL; the remainder held a CPL (22.8%) or ATPL (14.2%), or were SPs (10.2%).

Consistent with previous research (Wiggins & O'Hare, 1995, 2003a), participants were classified as either expert or novice based on flight experience, with those with more than 1000 hours⁷ being classified as experts and those with less than 1000 hours being classified as novices. This classification resulted in 56 experts and 140 novices (five participants did not provide flight experience).

3.2.2 Material & Design

3.2.2.1 Overview

The experiment employed a within-subjects design to evaluate the anchoring effect on participants' assessments of weather conditions. Specifically, using an online questionnaire platform, participants were asked to assess a series of in-flight

⁷ Total flight hours was used rather than 'cross-country' hours, which Wiggins and O'Hare (1995, 2003a) used, as only total flight hours data were collected.

scenarios after being exposed to a weather forecast, which acted as the anchor. A vignette-based methodology is an accepted and common method used in aviation psychology research (e.g., Gilbey & Hill, 2012; Hunter et al., 2003; Wiggins & O'Hare 2003a). The main benefit of using this methodology is it allowed accurate control over variables, which was an important element of this study. Furthermore, the use of the online platform enabled the material to be distributed to a wider audience, which ensured that the sample collected more closely reflected the target population (e.g., a range of experience levels).

The main shortcoming of this type of methodology requires additional explanation. This thesis explored pilot behaviour during weather-related decision making. Conducting this type of research in the real-world environment (i.e., placing participants in a VFR flight into IMC situation) would be unethical and unviable. For this reason, many studies explore pilot behaviour by using self-reported measures, like the scenario-based questionnaire seen in this study. It has been suggested that such methods may result in a substantial mismatch between people's intentions and actual behaviour. However, more recently, it has been argued that this type of methodology, at the very least, acts as an indicator of real-life behaviour (e.g., Armitage & Connor, 2001; Exum, Turner, & Hartman, 2012; Pogarsky, 2004): Therefore this methodology should provide a good tradeoff between identifying the presence of cognitive biases in weather-related decision making and not placing pilots into increased levels of danger.

3.2.2.2 Development of Scenarios

To explore the anchoring effect in weather-related decision making, five scenarios were developed. In each scenario, participants were asked to imagine that they were the PIC of a Cessna 172 on a VFR flight and that they were approaching their destination. The Cessna 172 was chosen, as this aircraft is a common GA aircraft that seats four people. Most pilots would easily recognise the aircraft type and its performance. Furthermore, this type of aircraft is consistent with the type of aircraft that is most likely to be involved in GA weather-related accidents: single-engine, fixed gear aircraft (AOPA, 2014).

Participants were provided with an in-flight static image of a simulated environment, looking through the cockpit window in their intended direction of travel while flying at 1000 ft above ground level. The static images were captured using a similar method to that of Coyne et al. (2008) using Microsoft Flight Simulator X. Each static image was developed to satisfy the following criteria: (i) it was over areas of relatively flat terrain⁸; (ii) it was at locations situated close to sea level (therefore alleviating any confusion between using mean sea level and above ground level (AGL) when estimating cloud heights), (iii) during daylight conditions, (iv) all five were over similar but not exactly identical types of terrain to reflect the varying flight environment pilots operate over.

⁸ This study was primarily interested in the potential anchoring effect of weather conditions, therefore 'flat terrain' was used in each image to maintain control over the study variables.

The use of simulated static images was chosen as it allowed the weather conditions to be accurately set, which was an important requirement of the study. Video would have been preferable; however, the quality of the video would have had to be significantly reduced to make it acceptable for most users to access via the online platform. The weather conditions in all the in-flight static image were a horizontal visibility of 16 km and a cloud base at 2500 ft above ground level. These conditions were chosen, as they sit in between good flying conditions and poor flying conditions for VFR. The 'average' pilot would be likely to consider the weather conditions carefully, rather than automatically continuing or diverting, which are likely to be the case with very good or very poor weather.

The participants were asked to imagine that they had obtained a plain language text form weather forecast for their destination prior to the flight. The weather forecast conditions acted as the anchor. To avoid confusion, participants were only informed that the weather conditions were received 'prior to flight'. Participants were not informed how old the forecast were; however, in general, a weather forecast received prior to flight would be at least few hours old by the time a flight was conducted.

Approximately half the participants received a 'poor' weather forecast (i.e., a low anchor with low cloud or low visibility). An example one of the of a poor weather forecast is:

Wind at 270°T at 5 knots; visibility, 7 km; cloud broken at 1200 ft; temperature, 9°C; dew point, 6°C; QNH, 1012

The other half of the participants were exposed to a 'good' weather forecast (i.e., a high anchor with high cloud or high visibility), such as:

Wind at 270°T at 5knots; visibility, 25 km; cloud broken at 3800 ft; temperature, 9°C; dew point, 6°C; QNH 1012

The weather conditions in the in-flight static image were exactly halfway between the two anchors (i.e., visibility of 16 km and a cloud base at 2500 ft). The anchor weather condition pairs were chosen to ensure that the weather presented remained realistic, (i.e., very low cloud or very low visibility would be unrealistic for the weather in the static images). The non-anchoring information (i.e., wind, temperature, dew point and pressure) were set to conditions that were unlikely to influence the pilots' assessments of cloud and visibility (e.g., light winds) and remained the same for both anchoring conditions. A sample of the full questionnaire used in Study 1 can be found in Appendix A.

Each scenario was reviewed by several aviation employees at a New Zealand-based international flight training organisation with GA flying experience. The feedback from the reviews suggested that each scenario was realistic and consistent with the static images.

To better reflect the varying real-life weather assessments that a pilot may encounter, even on a single flight, each scenario had a slightly different anchor pairing (i.e., a high and a low weather forecast anchor). However, it is important to note that the weather conditions in the static images always remained the same and was always

halfway between the two anchors (i.e., a cloud base at 2500 ft and a horizontal visibility of 16 km) and the anchor pairings were fixed for each scenario. A list of the anchor pairings for each of the five scenarios can be found in Table 1 on the following page.

Table 1: *The five scenarios and their anchor pairings used in Study 1*

Scenario 1	High Anchor	Wind at 270°T at 5 kt; visibility is 27 km; cloud broken at 4000 ft; temperature, 9°C; DP, 6°C; QNH, 1012
	Low Anchor	Wind at 270°T at 5 kt; visibility is 5 km; cloud broken at 1000 ft; temperature, 9°C; DP, 6°C; QNH, 1012
Scenario 2	High Anchor	Wind at 030°T at 3 kt; visibility is 24 km; cloud broken at 3700 ft; temperature, 22°C; DP, 16°C; QNH, 1002
	Low Anchor	Wind at 030°T at 3 kt; visibility is 8 km; cloud broken at 1300 ft; temperature, 22°C; DP, 16°C; QNH, 1002
Scenario 3	High Anchor	Wind at 080°T at 8 kt; visibility is 26 km; cloud broken at 3900 ft; temperature, 19°C; DP, 6°C; QNH, 1004
	Low Anchor	Wind at 080°T at 8 kt; visibility is 6 km; cloud broken at 1100 ft; temperature, 19°C; DP, 6°C; QNH, 1004
Scenario 4	High Anchor	Wind at 180°T at 6 kt; visibility is 25 km; cloud broken at 3800 ft; temperature, 12°C; DP, 6°C; QNH, 1001
	Low Anchor	Wind at 180°T at 6 kt; visibility is 7 km; cloud broken at 1200 ft; temperature, 12°C; DP, 6°C; QNH, 1001
Scenario 5	High Anchor	Wind at 210°T at 9 kt; visibility is 23 km; cloud broken at 3600 ft; temperature, 18°C; DP, 15°C; QNH, 1005
	Low Anchor	Wind at 210°T at 9 kt; visibility is 9 km; cloud broken at 1400 ft; temperature, 18°C; DP, 15°C; QNH, 1005

DP, dew point; QNH, sea level pressure (hPa)

3.2.3 Procedure

Participants were recruited either through an email invitation to students and staff at a New Zealand-based international flight training organisation or through two leading GA online forums between February and May 2013. The link directed participants to the online questionnaire (Survey Gizmo), where participants were first provided with a brief introduction to the questionnaire and informed that the study had undergone an ethical review. A review of the ethics of this study was undertaken using the guidelines provided by the Massey University Ethics Committee. The project was deemed to be of low risk and therefore suitable for a 'low-level notification' to Massey University Human Ethics Committee. A copy of the low-risk Notification can be found in Appendix B.

Next, participants received each of the five scenarios. Studies of anchoring and adjustment have been conducted in a range of areas; therefore, this study followed a similar procedure to previous studies (Plous, 1989; Strack & Mussweiler, 1997; Tversky & Kahneman, 1974). Participants were first asked to make a comparative judgment, comparing the weather conditions in the in-flight image with the weather forecast ("Is the weather forecast worse, the same as or better than the conditions in flight?"). The purpose of the comparative question is to encourage participants to actually process and consider the given anchor values. Participants were then asked to make an absolute judgment of the cloud height and visibility distance. After that, participants rated each of the following dimensions using a nine-point scale for each scenario: (i) a safety assessment to continue the VFR flight ("It is safe to continue towards your destination.") on a scale ranging from 1 = *strongly disagree* to 9 =

strongly agree, and (ii) a confidence rating of weather assessment (“How confident are you of these estimates?”) on a scale ranging from 1= *not confident* to 9 = *very confident*.

The participants evaluated a total of five scenarios; the online platform randomly assigned the order of the scenarios, as well as the anchor (high or low) that accompanied each scenario. As a result of the random assignment of the anchor, participants received a mix of high and low anchor scenarios. At the end of the questionnaire, participants were asked the following demographic information: age, flight hours and licence held (SP/PPL/CPL/ATPL).

No time limit was imposed on participants completing the questionnaire. The questionnaire was conducted confidentially and therefore, it was not possible to determine which recruitment method participants came from. A *priori* power analysis, using the software G*Power (Erdfelder, Faul, & Buchner, 1996), was used to determine that with $\alpha = .05$, a total sample size of $n = 46$ would be sufficient for an experimental power of .80 (assuming an effect size of .25) and a correlation between within-subjects measurements of $r = .3$, using a repeated-measures within-between interaction analysis of variance (ANOVA).

3.3 Results

3.3.1 General findings

The data were initially screened to identify any outliers and to establish that the data met the requirements for the application of parametric statistical analysis. This screening identified three outliers. These outliers were defined as z-scores greater than +3.5 (Iglewicz & Hoaglin, 1993) in the cloud height and visibility data (no outliers were identified in the other variables). Further investigation ruled out data entry errors. It is not unreasonable to expect outliers, especially at the higher end of weather assessments, where cloud height and visibility can be considerable (e.g., very high cloud or very good visibility). For this reason, the outlier data were retained⁹.

The level of statistical significance was set at $\alpha = 0.05$ for all statistical tests and all tests were conducted as two-tailed. (Although four dependent variables were analysed in this study, a decision was made not to adjust alpha (e.g., using the Bonferroni correction) because so doing would increase the risk of *Type II* error. In the context of the study reported here, it was decided that, as any studies would require replication prior to being used in real-world environments, an increased risk of *Type I* error was less important than an increase in *Type II* error. The same reasoning was applied throughout this thesis.

⁹ To ensure the outliers did not impact the findings, each analysis was also conducted without the outliers, which indicated the outliers had little impact on the final analysis.

The internal consistency (Cronbach's alpha) was calculated and found a high level of internal consistency across all five scenarios for both cloud height ($\alpha = 0.71$) and visibility assessment ($\alpha = 0.85$). The Cronbach's alpha would not have been improved by removing a scenario. Participants judged the material to be above average for realism (mean $[M] = 4.77$ ($SD = 2.1$)).

Overall, there were 201 pilots who completed the survey. The distribution of pilots sampled in this study, by licence type, was compared to the active pilots in Australia and the US (CASA, 2013; FAA, 2009) as shown in Figure 5. Australian and US pilot statistics have been chosen as a comparison, as these two countries make up almost 77% of GA flying activity¹⁰ (International Council of Aircraft Owner and Pilot Associations [ICAOPA], 2014).

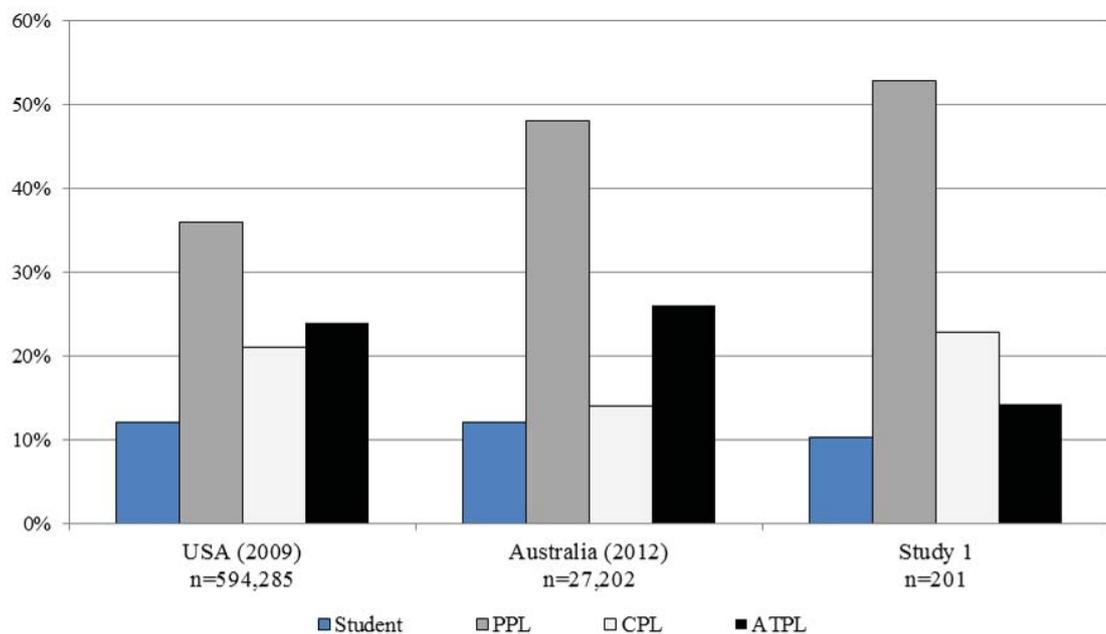


Figure 5: Comparison of Study 1 pilots to population totals by licence type

- A US (FAA) data also include recreational pilots, rotor craft pilots and glider pilots, who are not show in Figure 5 but make up the remaining 7% of pilots.
- B PPL, private pilot licence; CPL, commercial pilot licence; ATPL, air transport pilot licence

¹⁰ A comparison with New Zealand pilots was not conducted as the data on the number of student pilots was not available.

A chi-square test of independence was performed to examine the relationship between the distributions of Study 1 licence holders and the population licence holders. No significance was found, $\chi^2(3) = 5.91, p = .116$, which suggests that the pilots sampled in this study, by licence type, were similar to the active Australian and US pilots.

3.3.2 Cloud Height Assessment

A mixed model ANOVA was used to compare pilots' assessments of cloud height after being exposed to an anchor (high vs. low) and their level of experience (novice vs. expert). There was a significant main effect for cloud height assessment after having been exposed to an anchor, $F(1,184) = 100.9, p < .001, \eta^2 = .35$: pilots reported a higher assessment of cloud height after being exposed to the high anchor ($M = 2468$ ft, $SD = 813$), compared to when they were exposed to the low anchor ($M = 1756$ ft, $SD = 718$). There was no evidence of a main effect of experience (novice vs. expert), $F(1,184) = 0.51, p = .47, \eta^2 = .003$, nor any interaction between anchor and experience, $F(1,184) = 0.28, p = .60, \eta^2 = .001$. Table 2 provides the means and SDs of the anchors (high vs. low) for the cloud height and visibility assessments.

Table 2: *Mean assessment of cloud height and horizontal visibility distance by level of anchor*

	Low Anchor	High Anchor
Cloud Height (ft)	1756 ($SD = 718$)	2468 ($SD = 813$)
Visibility (km)	11.48 ($SD = 6.57$)	15.49 ($SD = 8.19$)

3.3.3 Visibility Assessment

A mixed model ANOVA was used to compare pilots' assessments of visibility after having been exposed to an anchor (high vs. low) and experience level (novice vs. expert). There was a significant main effect for visibility assessment after being exposed to an anchor, $F(1,183) = 84.7$, $p < .001$, $\eta^2 = .31$, with pilots reporting a higher assessment of visibility after being exposed to the high anchor ($M = 15.49$ km, $SD = 8.19$), compared to when they were exposed to the low anchor ($M = 11.48$ km, $SD = 6.57$). The main effect comparing the two experience groups (novice vs. expert) was significant, $F(1,183) = 5.50$, $p = .020$, $\eta^2 = .029$. There was a significant interaction of visibility assessment after exposure to the anchor and experience level, $F(1,183) = 7.38$, $p = .007$, $\eta^2 = .027$. Figure 6 shows the interaction of visibility assessment between anchors (high vs. low) and experience (novice vs. expert).

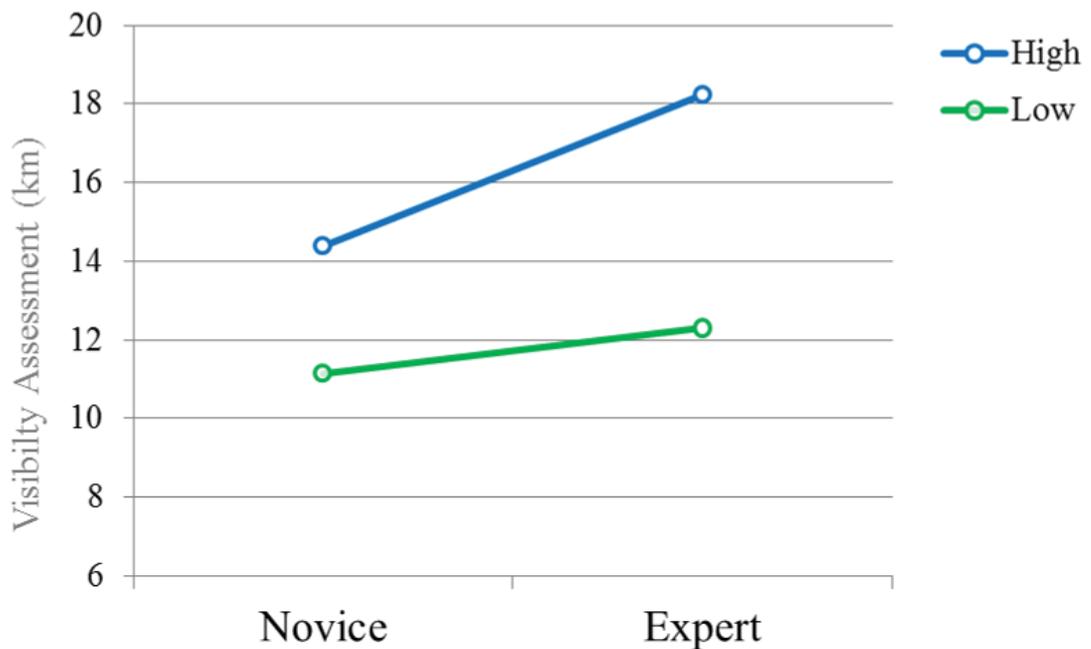


Figure 6: Mean assessment of horizontal visibility distance, by experience level (novice vs. expert) and by anchor condition (high vs. low).

3.3.4 Safety Assessment For Continuing the VFR Flight

A mixed model ANOVA was conducted to compare pilots' safety assessment to continue the VFR flight after having been exposed to an anchor (high vs. low) and by level of experience (novice vs. expert). There was a significant main effect for safety assessment after being exposed to an anchor, $F(1,183) = 8.51, p = .004, \eta^2 = .044$, with pilots reporting a higher safety assessment score after being exposed to the high anchor ($M = 6.62, SD = 1.66$), compared to when they were exposed to the low anchor ($M = 6.17, SD = 1.73$). The main effect comparing the two experience groups (novice vs. expert) was also significant, $F(1,183) = 17.2, p < .001, \eta^2 = .086$. Table 3 shows the safety rating between the anchor conditions and experience levels. Compared to the novice pilots, experienced pilots judged it to be safer to continue under both anchor conditions. There was no evidence of an interaction effect on the safety assessment scores for anchor and level of experience, $F(1,183) = 1.39, p = .24, \eta^2 = .009$.

Table 3: Mean safety assessment scores about continuing the VFR flight (1 = not safe, 9 = very safe) by experience level (novice vs. expert) and anchor condition (high vs. low)

	Low Anchor	High Anchor
Novice	5.85 ($SD = 1.69$)	6.39 ($SD = 1.64$)
Expert	6.95 ($SD = 1.56$)	7.19 ($SD = 1.57$)

3.3.5 Confidence Assessment

A mixed model ANOVA was conducted to compare pilots' confidence in their weather assessment after exposure to an anchor (high vs. low) and experience level (novice vs. expert). No significance was found in either the interaction or main effects, with pilots showing above average confidence for their weather assessment overall ($M = 5.74$, $SD = 1.56$).

3.3.6 Licence Type

A mixed model ANOVA was conducted to compare the four dependent variables (cloud assessment, visibility assessment, safety assessment and confidence rating) after exposure to an anchor (high vs. low) and licence type: SP ($n = 20$), PPL ($n = 104$), CPL ($n = 45$) and ATPL ($n = 26$).

The main effect comparing the four licence groups with visibility assessment was significant, $F(1,182) = 3.49$, $p = .017$, $\eta^2 = .054$. *Post hoc* tests (Tukey) revealed that the CPL holders assessed visibility as being significantly higher than did PPL holders for both anchor conditions ($p = .012$). Table 4 displays the visibility assessments for PPL and CPL holders. No other significant differences between groups were found. There was no evidence of an interaction effect on visibility assessments for anchor and licence group, $F(1,182) = 1.89$, $p = .13$, $\eta^2 = .033$.

Table 4: Mean visibility assessment by licence type (student pilot [SP], private pilot licence [PPL], commercial pilot licence [CPL, air transport pilot [ATPL]]) and anchor condition (high vs. low).

	Low Anchor	High Anchor
SP	11.9 km ($SD = 7.83$)	15.2 km ($SD = 8.46$)
PPL	10.5 km ($SD = 5.90$)	13.7 km ($SD = 6.79$)
CPL	13.1 km ($SD = 7.43$)	18.4 km ($SD = 7.23$)
ATPL	11.8 km ($SD = 6.02$)	17.4 km ($SD = 11.97$)

The main effect comparing the four licence groups with safety assessment was significant, $F(1,182) = 4.85$, $p = .003$, $\eta^2 = .074$. *Post hoc* tests (Tukey) revealed that the CPL holders assessed it to be safer to continue compared to PPL holders ($p = .050$) and SPs ($p = .002$). No significant differences were observed between any other licence groups. There was no evidence of an interaction effect on safety assessments for anchor and licence group, $F(1,182) = 1.03$, $p = .38$, $\eta^2 = .016$.

There was no significant main effect or interaction for pilot licence groups and the two remaining dependent variables (cloud assessment and confidence assessment).

3.3.7 Age

A mixed model ANOVA was conducted to compare the four dependent variables (cloud assessment, visibility assessment, safety assessment and confidence rating) after exposure to an anchor (high vs. low) and participants' age. Participants' ages

were collapsed into four age groups: 0–20 years ($n = 25$), 21–40 years ($n = 112$), 41–60 years ($n = 41$) and 61+ years ($n = 17$).

The main effect comparing the four age groups and cloud assessment was significant, $F(1,183) = 3.94$, $p = .009$, $\eta^2 = .061$. *Post hoc* tests (least significant difference [LSD]¹¹) revealed that younger age groups assessed the cloud height as being higher for both anchor conditions compared to the older age groups, as shown in Figure 7. The <20 age group was significantly different from the 41–60 ($p = .010$) and 61+ ($p = .007$) age groups. The 21–40 age group was also significantly different from the 41–60 ($p = .041$) and 61+ ($p = .027$) age groups. There was no significant difference between the two younger age groups (0–20 and 21–41) or between the two older age groups (41–60 and 61+). There was no evidence of an interaction effect on cloud height assessment of anchor and age groups, $F(1,183) = 0.27$, $p = .85$, $\eta^2 = .003$.

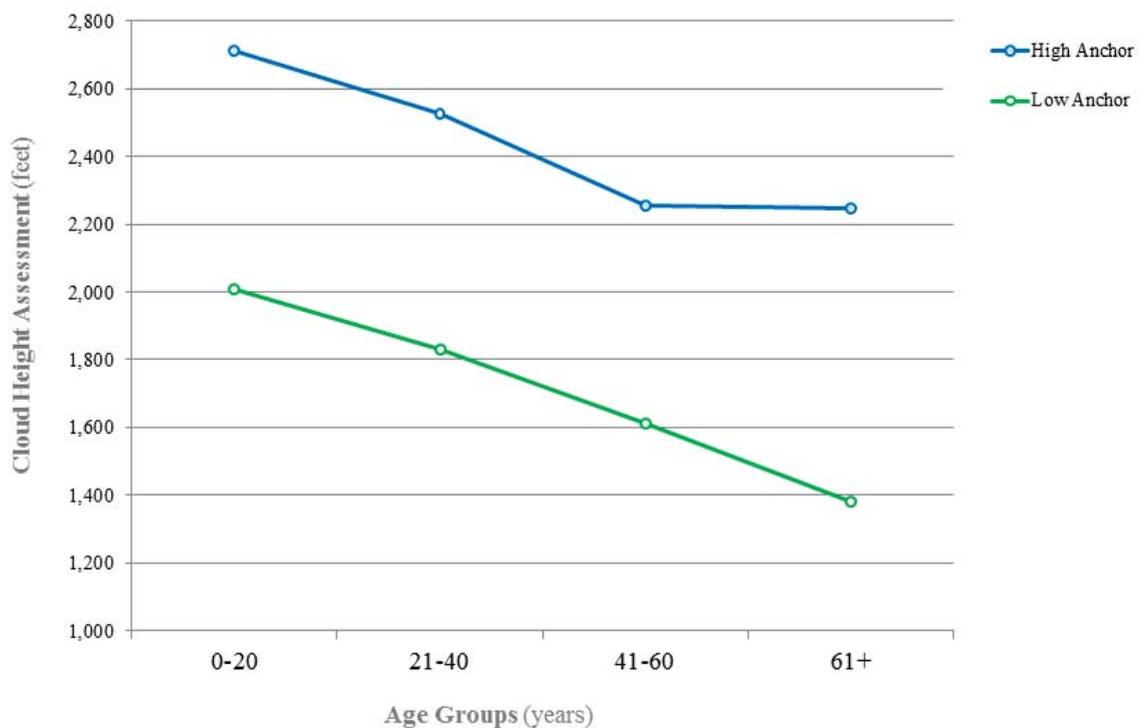


Figure 7: Cloud assessment by age group and anchor condition (high vs. low).

¹¹ LSD was chosen for the *post hoc* test rather than Tukey's, as Tukey's *post hoc* test was not sensitive enough to pick up the significance between some of the groups.

The main effect comparing the four age groups and confidence in their weather assessment was significant, $F(1,183) = 4.41$, $p = .005$, $\eta^2 = .067$. *Post hoc* tests (Tukey) revealed that there was a significant difference between the 41–60 age group and all other age groups: 0–20 ($p = .042$); 21–40 ($p = .020$) and 61+ ($p = .013$). The 41–60 age group had lower confidence overall ($M = 4.97$, $SD = 1.75$) about their weather assessments, whereas the younger age groups (>20 and 21–40) tended to have a higher level of confidence about their weather assessment ($M = 6.04$, $SD = 1.69$ and $M = 5.81$, $SD = 1.38$ respectively). There was no evidence of an interaction effect on confidence assessment of anchor and age group, $F(1,183) = 0.06$, $p = .98$, $\eta^2 = .001$.

There was no significant main effect or interaction for age groups and the two remaining dependent variables (visibility assessment and safety assessment).

3.4 Discussion

The findings presented here suggest that the anchoring effect can affect weather-related decision making. After exposure to a ‘good’ weather forecast, participants assessed the conditions as being better than they did if they were exposed to a ‘poor’ weather forecast, thus supporting Hypotheses 1a and 1b for this study. In other words, pilots appear to anchor on reports that may no longer be valid. A comparison of pilots’ assessments of cloud height and horizontal visibility distance found that when pilots were exposed to a high anchor (forecasted high cloud and good visibility), their assessment of the cloud and visibility was greater than when they were exposed to a low anchor (forecast low cloud and poor visibility). Even though

pilots were exposed to the same in-flight images, their judgment of the weather conditions was influenced by the ‘initial’ weather report acting as a high or low anchor.

The findings also suggest that, due to the anchoring effect, a pilot’s perception of the safety of a flight may be influenced. That is, after having been exposed to the high anchor, pilots felt it was safer to continue the flight compared to those who were exposed to the low anchor. These findings support Hypothesis 1c. Pilots provided with a ‘poor’ weather forecast may divert unnecessarily, when, in reality, the weather may have improved since the earlier forecast they had obtained. Pilots with a ‘good’ weather forecast may continue with the flight when, in reality, the weather may have worsened since the forecast they had originally obtained and the sensible option would be to divert or to alter their course. Participants in both anchor conditions scored towards the ‘safe to continue’ end of the scale. One explanation for this may be due to the scenarios located close to an airport. Scenarios which anchoring occurs closer towards the departure airport may be more inclined not to continue. The extent to which pilots were confident about their weather assessment did not appear to be affected by the anchor. This finding is interesting and possibly demonstrates that participants were completely unaware of the effect an anchor may have on their situational awareness. Such a possibility would be consistent with Pronin’s (2007) observation that “people tend to recognise (and even overestimate) the operation of bias in human judgment – except when that bias is their own” (p. 37).

The size of the anchoring effect was consistent with that reported in other studies of anchoring and adjustment (e.g., Englich et al., 2006; Northcraft & Neale, 1987), with

the difference between the high and low anchor assessment being 700 ft for cloud assessment and 2 km for visibility. The size of the anchoring effect, especially on assessments of the correct cloud height, could have a significant impact on the safety of a flight, particularly given the number of flights that encounter changing weather conditions every day. For example, a number of weather-related accidents have occurred shortly after takeoff (NTSB, 2005), where the pilot took off into conditions that he or she felt were suitable for flight but later found that was this not the case.

The dangers of departing on a VFR flight when the cloud height and/or visibility are not suitable can be observed in the following fatal accident:

A Cessna 150 (a two-seat aircraft) departed on a local flight in Ontario, Canada, in August 2009. The accident report indicated the pilot likely overestimated the height of the cloud, resulting in the aircraft flying into IMC shortly after takeoff. The pilot probably became spatially disoriented, resulting in a loss of control of the aircraft and uncontrolled flight into the terrain. The aircraft was destroyed and the pilot was fatally injured (TSB, 2010b).

In these types of accidents, just a small difference in the cloud height or forward visibility could make a significant difference between a safe flight and an unsafe flight.

It was particularly interesting that experienced pilots, who presumably would have assessed the weather many more times than novice pilots, were affected by the anchoring effect to the same degree (in some cases more so), therefore only partially supporting Hypothesis 2. These findings are broadly consistent with previous research (e.g., English et al., 2006; Northcraft & Neale, 1987), in which experts in a

particular field were also reported to fall prey to the anchoring effect. The findings would suggest that gaining flight experience does not necessarily protect a pilot from being influenced by the anchoring effect. It could be argued that the 1000-hour cut-off point between experts and novices used in this study was too low and more experienced pilots would be less likely to be influenced by the anchoring effect. Wiggins and O'Hare (1995, 2003a) used 1000 'cross country hours'¹² to distinguish between expert and novice pilots, a measure which was not collected in this study. However, the ATPL licence group of pilots, who had a mean flight experience of 2654 hours ($SD = 1781$) experienced almost the same anchoring effect for both cloud height and visibility assessment as the less experienced licence groups.

This points to the possibility that another measure to define 'experts' may be necessary for studies of weather-related cognitive biases. One possible tool, which has produced contrasting findings, is the role an instrument rating (IR) may have on weather-related decision making (Coyne et al., 2008). A holder of an IR would have received considerable training to obtain the skills necessary to operate in IMC. Moreover, only 25% of VFR flight into IMC accidents in the US occurred where the pilot held an IR, even though over half of active US pilots hold an IR (FAA, 2009), which would suggest that pilots with IR may possess better skills at judging the weather conditions to avoid IMC. However, Coyne et al. (2008) found that IR pilots were no better at judging cloud height and visibility than non-IR pilots. IR data were not collected in this study and therefore it was not possible to determine if IR pilots would be influenced differently by the anchoring effect compared to non-IR pilots.

¹² Cross country hours include flights which travel more than 25nm in a straight line from the departure airport. Pilots are not required to record cross country hours (unlike night flying or instrument flying hours), therefore determining ones cross country hours is challenging. 'Total' flight experiences is accurately recorded therefore much easier to collect.

Therefore, IR data were explored in the next study to further our understanding of this indicator of cognitive biases.

The findings that experience had little impact on the anchoring effect would imply that experienced pilots are just as likely to be involved in VFR flight into IMC accidents as the novice pilots. However, the general accident data do not support this (NTSB, 1989). This may be because the expert pilot is able to cope with the hazardous situation. Due to the anchoring effect, both novice and expert pilots may be drawn into a false sense of security, perceiving the weather conditions better than they are in reality. When a pilot recognises that VFR flight is no longer possible, it is possible that the advantages of experience provide the critical difference between a safe outcome and a fatal outcome. A more experienced pilot may have the necessary skills to enable him or her to negotiate a safe passage out of the dangerous situation, whereas the novice pilot may be less able to cope with the high workload and stressful situation.

The analysis of pilot licence groups broadly supports Hypothesis 3. Differences between pilot licence groups were similar to the general findings comparing experience levels (i.e., novices vs. experts), with small but significant differences observed in the areas of visibility assessment and safety assessment scores. These differences were particularly noticeable when comparing PPL and CPL holders. This was not unexpected, considering that the PPL holders broadly sat within the novice group (mean flying hours = 407) and CPL holders within the experienced group (mean flying hours = 1037). In visibility assessments and safety assessments, the anchoring effect was observed; however, CPL holders tended to anchor around a

higher visibility assessment (i.e., overall, they assessed the visibility as being greater) and tended to assess it to be safer to continue than PPL holders under both anchoring conditions.

The findings of pilot age partially supported Hypothesis 4, with small differences being observed in cloud assessment and confidence levels. Consistent with the primary findings of this study, the anchoring effect of cloud height was observed across all the age groups. Overall, the younger pilot groups tended to assess the cloud height as being higher, whereas the older pilot groups assessed the cloud height as being lower. No anchoring effect was observed in the confidence assessments; however, younger pilots tended to be more confident overall with their weather assessment, whereas older pilots were less confident. Combining these two findings would suggest that younger pilots may be putting themselves at greater risk by continuing flight towards IMC. Studies (Coyne et al., 2008) have shown that if pilots accurately assess the weather conditions and accurately recognise that they will not remain in VFR on their current heading, they are more likely to divert. The findings of this current study would suggest that a younger pilot is likely to assess the weather conditions as being better than they are (e.g., increased cloud height) and be more confident in this assessment. This may result in the younger pilot pushing on into deteriorating weather conditions, putting him/herself at a greater risk of VFR flight into IMC. These findings are also consistent with the general findings that younger pilots are more likely to be involved in VFR flight into IMC accidents (Goh & Wiegmann, 2001).

Considering the findings of this study, it is possible to speculate how the accident described at the start of this chapter could potentially have been influenced by anchoring and adjustment (CAA NZ, 2005a). As the pilot set off on the cross-country flight, the weather forecast obtained before takeoff anchored in the mind of the pilot would have suggested sufficient cloud height and visibility for the VFR flight. However, around halfway into the flight, the weather then began to deteriorate. The anchoring effect potentially biased the pilot's assessment of the weather conditions, with the pilot perceiving the conditions to be better than they were and choosing to press on with the flight. It is likely that in this situation, the pilot reached a stage he realised the conditions were, in fact, worse than forecast, possibly when he was forced to fly lower than the expected cloud height forecast. However, the pilot pushed on, travelling at speeds of over 200 km/h in limited visibility under high workload and stress make it all too easy to inadvertently enter IMC.

There are a few aspects of this study that potentially may limit the generalisation and validity of these results. First, the study involved the use of static images from a simulated environment; video would have been preferable but the quality of the video would have had to have been significantly reduced to make it acceptable for most users to access via the online platform. Whether similar results would be obtained in actual flight conditions cannot be determined. However, in real-life situations, the time and workload pressures associated with flying the aircraft may lead to increased stress levels, which have been shown to deteriorate decision making ability by inducing people to place greater reliance upon heuristics (Wickens, Stoke, Barnett, & Hyman, 1993). Nevertheless, Northcraft and Neale (1987) found evidence that laboratory research on decision biases, particularly anchoring and adjustment, is

applicable to ‘real-world’ information-rich decisions. Second, the study looked at a single weather scenario in which the weather conditions were marginal; therefore, caution should be used when applying these results to more extreme conditions (e.g., in very poor weather conditions or very good weather conditions). Third, caution also needs to be applied when generalising these results to the GA pilot population. Although this study captured a wide spectrum of pilots, as highlighted in Figure 5, the core sampling method for this study was through convenience sampling at a New Zealand-based international flight training organisation and via advertising through two major aviation forums. Fourth, a control condition (without an anchor) would have been a useful addition to the general analysis, as it would have been interesting to note if the control group’s assessments were closer to one particular anchor. Fifth, participants may have viewed the images on different displays (e.g., smartphone, large computer displays etc.). This could potentially influence how participants assessed the weather conditions. This is likely to have had minimal impact as the online platform is specifically designed to be displayed on various different display sizes. Furthermore the findings of Study 4 were consistent with the findings of this study, which used the same images but all printed on high quality paper.

This study suggests that weather-related decisions can be influenced by the anchoring effect; however, further research in this area is necessary to get a better understanding of the full extent. Three areas of interest could provide the next step. First, exploring different weather scenarios are likely to give a better understanding of the influences of anchoring over a wide spectrum of weather scenarios. For example, does the anchoring effect change in very good weather or very bad weather? Second, a simulator study would provide some insight into how the

pressure of flight demands influences the anchoring effect, particular during flight in highly demanding situations. Third, are pilots influenced by the anchoring effect depending on the source of the weather related information? For example, does the anchoring effect differ if the information is supplied from an official weather forecaster or from a fellow pilot?

The possibility that the anchoring effect can occur when pilots assess the current weather information using environmental cues highlights the importance of obtaining accurate and current weather information. This is particularly important when flying in dynamic weather conditions such as convective activity or frontal systems. In these situations, even a weather forecast that is a few hours old may no longer reflect the actual weather conditions.

In conclusion, the findings of the current study suggest that the anchoring effect may adversely affect weather-related decision making. When pilots are assessing the weather, they have a tendency to anchor and under-adjust on reports that may no longer be valid, thus potentially increasing the risk of a pilot flying into conditions unsuitable for VFR flight. Experience had little impact on reducing the rate at which the anchoring effect occurred.

CHAPTER FOUR

STUDY 2

CONFIRMATION BIAS AND SEEKING INFORMATION TO CHECK A HYPOTHESIS

4.1 Introduction

Study 2 explored confirmation bias, a cognitive bias that influences the type of information used when making decisions. In the context of pilot decision making, confirmation bias may result in a pilot seeking out or interpreting evidence in a way that favours an existing belief, expectation or hypothesis (Gilbey & Hill, 2012; Muthard & Wickens, 2001; Nickerson, 1998). In some situations, confirmation bias also results in unintentionally ignoring evidence that disconfirms the favoured hypothesis (Green et al., 1996).

Confirmation bias has been shown to occur in a number of applied areas of research such as medicine (e.g., Croskerry, 2002; Pines, 2006) and criminal investigation (Ask & Granhag, 2005). However, very few experimental studies of confirmation bias in aviation have been reported, despite confirmation bias having been identified as a casual factor in aviation accidents (e.g., CAA NZ, 2005b; NTSB, 2008). If a pilot is assessing the weather conditions, particularly those which occur prior to a

planned VFR flight, focusing on confirmatory information may lead a pilot to interpret the conditions as being better than they really they are, which could result in the pilot flying into hazardous conditions (i.e., VFR flight into IMC).

When pilots assess weather conditions, which they are likely to do throughout a flight, they will be searching for a range of weather-related cues to assess their judgment (Wiggins & O'Hare, 1995, 2003a). When a pilot makes a judgment about the weather conditions—for example, whether it is suitable to continue to fly towards the destination—he or she must seek out evidence to test that judgment. The evidence a pilot seeks or believes to be useful will depend on what type of hypothesis testing strategy he or she applies: a *positive testing strategy*, in which he/she seeks confirmatory evidence to prove that a judgment is correct, or a *negative testing strategy*, in which he/she seeks disconfirming evidence to prove that the judgment is incorrect (Klayman & Ha, 1987). There is nothing inherently wrong in using the positive testing strategy. However, like the general use of heuristics, an over-reliance on the positive testing strategy can lead to serious consequences.

Pilots are likely to have a natural desire to complete a flight. For example, during a cross-country flight a pilot is probably aiming to complete a set task, such as meet with someone or returning home from a few days away. This tendency can be further reinforced by various motivational and social pressures, such as the sunk cost effect and plan continuation bias (Goh & Wiegmann, 2001; O'Hare & Owen, 2001; Wiggins et al, 2014). The sunk cost effect suggests that a person is more likely to continue with an action (e.g., continue a VFR flight) if he or she has already made a considerable investment of time or money (Arkes & Blumer, 1985). The plan

continuation bias is a determination to reach your destination despite changing circumstances (CAA NZ, 2012). This strong motivation to complete a VFR flight may result in pilots seeking out evidence that would support the desire to complete the flight; however, this could lead to a false positive: judging the weather conditions to be suitable for VFR flight when, in fact, they are not (Nickerson, 1998).

The author's aim in Study 2 was to explore confirmation bias in weather-related decision making; specifically, the study aimed to determine what type of hypothesis testing strategy pilots use when judging the weather conditions. It is predicted that when judging the weather conditions, participants will favour confirmatory over disconfirmatory evidence, consistent with Gilbey and Hill's (2012) study of confirmation bias in aviation 'lost' procedures (procedures to be followed when a pilot has lost their way during a flight).

Hypothesis 1: Participants will have a tendency to favour a positive testing strategy by seeking confirmatory over disconfirmatory evidence.

Similar to Study 1, this study also investigated the role that experience plays in the type of information pilots use when making decisions. Consistent with the findings in Study 1 and other experience-based cognitive bias research (Englich et al., 2006; Northcraft & Neale, 1987), it is expected that there will be little difference between expert and novice pilots when it comes to the type of hypothesis testing strategy favoured.

Hypothesis 2: There will be no difference between novice and expert pilots in regards to the type of evidence they use.

This study also explored a number of pilot characteristics to determine if they have any influence on the hypothesis testing strategy pilots use. Pilot age and licence held (SP, PPL, CPL and ATPL) were explored, and it was expected that age and licence type will have little influence on the type of evidence that participants used.

Hypothesis 3: The licence type held (i.e., SP, PPL, CPL or ATPL) will mirror the experience hypothesis (Hypothesis 2), with no difference expected in regards to the type of evidence used.

Hypothesis 4: Age will not influence the type of evidence used.

The findings of Study 1 suggested that basic pilot characteristics (total flight experience and age) were unreliable indicators to help predict the anchoring effect; therefore, two additional characteristics were also explored in this study: IR status and experience of VFR flight into IMC. Studies exploring these two factors have found contrasting results. A holder of an IR would have received considerable training to obtain the skills necessary to operate in IMC. Only 25% of VFR flight into IMC accidents occurred where pilots held an IR (Coyne et al., 2008), which would suggest that pilots with IR may possess better skills at judging the weather conditions in order to avoid IMC or help them fly safely out of IMC.

However, Coyne et al. (2008) found that pilots with an IR were no better at judging cloud height and visibility than pilots without an IR. Furthermore, in this study, the author was interested in the group of pilots that have held an IR in the past but no longer keep the IR current. The requirement to remain current is designed to ensure the pilot maintains the skills necessary to operate safely in IMC. Keeping an IR current can be challenging for a pilot who is not frequently operating in IMC. The exact criteria vary slightly between countries; for example, for a pilot operating in New Zealand, keeping the IR current requires him or her to (i) complete at least 3 hours of instrument time (operating in IMC conditions) within the preceding 3 months, (ii) carry out at least three instrument approaches and (iii) pass a competency check every 12 months (CAA NZ, 2011). It has been suggested (FAA, 2004; Wiggins et al., 2012) that without practicing the skills necessary to keep an IR current, these IR skills degrade quickly. Therefore, holding a non-current IR may be of little value.

The study by Wiggins et al. (2012) also explored the differences between pilots who have experienced VFR flight into IMC and those who have not. Differences in a number of characteristics were found; for example, pilots who had deliberately entered IMC tended to have experienced similar events previously and had a comparatively greater tolerance for risk. However, pilots who had inadvertently entered IMC had no demographic differences compared to pilots who had not experienced VFR flight into IMC.

Because of these contrasting findings, no specific prediction for the influence of holding an IR or having previous VFR flight into IMC experience was made.

The primary aim of the current study was therefore to investigate the type of evidence used in weather-related decision making. Additionally, this study investigated the effect of experience on the type of evidence used in weather-related decision making.

4.2 Method

4.2.1 Participants

Two hundred and seventy-eight pilots participated in this study. The mean age of the participants was 39.05 years ($SD = 15.33$) with a mean flying experience of 1931 hours ($SD = 3834$, ranging from 2 to 23,500 hours). Participants were classified as either experts ($n = 86$) or novices ($n = 169$), based on the same criteria as Study 1 (22 did not provide flight experience). Over half (50.6%) the participants held a PPL while the rest were SPs (8.2%), CPL holders (23.3%) and ATPL holders (17.9%).

4.2.2 Materials and Design

4.2.2.1 Overview

A vignette-based study was designed to evaluate the type of evidence participants used when assessing the weather conditions. In each of the five scenarios, pilots were asked to imagine they were flying towards their intended destination where the weather conditions were such that their VFR flight may be compromised. Therefore, in each of the scenarios, inappropriate use of the positive testing strategy may result in confirmation bias and put the pilot at a greater risk of VFR flight into IMC. The

benefits and potential limitations of using this methodology were discussed in Study 1, namely that a scenario-based methodology facilitates wide distribution of the material while keeping close control over the variables being studied. Furthermore, this methodology is a common method used in aviation psychology research (i.e., Gilbey & Hill, 2012; Hunter et al., 2003; Wiggins & O'Hare 2003a).

4.2.2.2 Development of Scenarios

This study used a similar procedure to that of Gilbey and Hill (2012), which explored confirmation bias in aviation lost procedures. Five scenarios were developed. In each of these, participants were asked to imagine that they were the PIC flying a Cessna 172 on a VFR flight approaching their destination (as in Study 1, the C172 was chosen because it is a common GA aircraft and is consistent with the type of aircraft that is most commonly involved in VFR flight into IMC accidents). Participants were provided with an in-flight static image, looking through the cockpit window. All images were captured by the author in the Waikato and Manawatu regions of New Zealand (generally flat terrain). Each static image was selected using the following criteria: (i) it was over relatively flat terrain, (ii) the image was taken during daylight conditions, (iii) all five were over similar types of terrain and (iv) the weather conditions in each image were marginal, with at least one of the criteria to remain VFR compromised (i.e., cloud height too low or visibility too low, or both) within a reasonable distance from the aircraft.

In each scenario, participants were asked to imagine that a situation had arisen which required them to proceed towards the closest aerodrome where they could land and

assess the situation. The closest airport was indicated by inserting a red box on the image, which was in the general direction which the aircraft was travelling. Situations ranged from a passenger feeling sick to the aircraft engine making abnormal noises.

Participants were instructed to choose one of three pieces of evidence that they could see and use to determine if they could ‘safely continue’ the VFR flight in their direction of travel. Participants were asked if they can ‘safely continue’, rather ‘not continue’ as this is consistent with pilots having a natural tendency to complete a flight (Goh & Wiegmann, 2001; O’Hare & Owen, 2001; Wiggins et al, 2014). All three pieces of evidence that accompanied each scenario could be observed in the static image and at least two were selected to ensure they were consistent with key cues experienced pilots used when making weather-related decisions, as identified by Wiggin and O’Hare (2003a).

Two pieces of evidence supported the hypothesis that safe flight was possible. For example, the cloud base ahead was high enough. These were designated as the confirmatory choices and indicated the use of a positive testing strategy for the hypothesis that VFR flight was possible. A third piece of evidence did not support VFR flight. For example, the visibility was too low. This piece of evidence was designated as the disconfirmatory choice used for negative testing of the hypothesis that VFR flight was possible. Table 5 presents the evidence that participants received for each scenario.

Table 5: Evidence used for each scenario in Study 2

Scenario No.	Confirmatory evidence	Disconfirmatory evidence
1	<ul style="list-style-type: none"> • I can see through the rain shower ahead. • Another aircraft ahead safely flew in my direction 10 min earlier. 	<ul style="list-style-type: none"> • I will almost certainly end up having to fly through cloud.
2	<ul style="list-style-type: none"> • The cloud base remains above my current level. • You have just received the meteorological report (ATIS) for your aerodrome which indicates light winds. 	<ul style="list-style-type: none"> • The dark cloud above my flight path may produce a heavy rain shower at any time.
3	<ul style="list-style-type: none"> • I can see the horizon through the rain shower. • Another aircraft ahead has reported that the cloud height is above circuit altitude. 	<ul style="list-style-type: none"> • The cloud base inside the rain shower may be lower than my current flight level.
4	<ul style="list-style-type: none"> • There is a clear patch towards the aerodrome I am heading towards. • The heavy showers are remaining to the right of my track. 	<ul style="list-style-type: none"> • The cloud base is too close to the ground.
5	<ul style="list-style-type: none"> • The visibility is very good towards my destination. • An aircraft ahead has reported light turbulence when flying over the hills. 	<ul style="list-style-type: none"> • The cloud is lowering towards the hill tops to the left of my track.

Confirmation bias is not usually considered a cognitive bias of simply seeking for confirmation evidence; instead, it is an interaction of using a positive testing strategy and encountering unusual situations (Gilbey & Hill, 2012). In each of the five scenarios, the weather conditions were such that continuing the flight could result in the VFR meteorological minima no longer being met. As a result, inappropriate use of the positive testing strategy may put the pilot at risk of VFR flight into IMC and was interpreted as evidence of confirmation bias.

One of the scenarios is given here as an illustration:

Imagine you are flying a Cessna 172 on a VFR cross-country. Your destination aerodrome is just beyond the red box (note the static image displayed with the scenario). As you approach the aerodrome, your passenger begins to feel air sick and wishes to land as soon as possible. To avoid too much discomfort for your passenger, it is important that you proceed towards the aerodrome if the conditions permit. The nearest suitable alternative aerodrome is an hour away.



Listed below are three statements about the weather conditions in front of your aircraft.

Pick the *one* that you feel is most useful in deciding if you can **safely continue** towards this aerodrome.

1. I will almost certainly end up having to fly through cloud.
2. I can see through the rain shower ahead.
3. Another aircraft ahead safely flew in my direction 10 min earlier.

In this particular scenario, although Items 2 and 3 (confirmatory choices) are evidence that could lead a pilot to decide that he or she could continue towards the aerodrome, these items could also occur simultaneously in situations when VFR flight is not possible. The forward visibility might be fine ('I can see through the rain shower ahead') but the cloud base might be too low. Another aircraft may have safely flown the route earlier but could have done at great risk (flying dangerously close to the ground), or the conditions could have changed in the intervening 10 min. Item 1 (the negative test), if correct, would be enough by itself to disprove the assessment that it is safe to continue the flight. The other four scenarios (five in total) were directly analogous to this scenario. The questionnaire used in Study 2 can be found in Appendix C.

Each scenario was reviewed by several aviation employees at a New Zealand-based international flight training organisation with GA flying experience. The feedback from these sessions confirmed that each scenario was realistic.

Participants were also asked for the following demographic information: age, flight hours, pilot licence held (SP, PPL, CPL or ATPL), IR status (holding a current IR,

holding an IR but not a current one or never held an IR) and whether or not the respondent had ever flown into IMC areas when he/she was not on an instrument flight plan.

4.2.3 Procedure

Participants were recruited either through an email invitation to students and staff at a New Zealand-based international flight training organisation, or through two leading GA internet forums between August and November 2013. The link directed participants to the online questionnaire (Survey Gizmo), where participants were first provided with a brief introduction to the questionnaire and informed that the study had undergone an ethical review. A review of the ethics of this study was undertaken using the guidelines provided by the Massey University Ethics Committee. The project was deemed to be of low risk and therefore suitable for a low-level notification to the Massey University Human Ethics Committee. A copy of the low-risk notification can be found in Appendix D.

Next, the participants evaluated the five scenarios. The online platform randomised the order in which the five scenarios appeared and the order in which the three pieces of evidence appeared with each scenario to minimise the potential for any order effects. Finally, participants completed the demographic section. No time limit was imposed on participants for completing the questionnaire.

The questionnaire was conducted anonymously and therefore it was not possible to determine which recruitment method had solicited the participants. *A priori* power

analysis using the software G*Power (Erdfelder et al., 1996) was used to determine that with $\alpha = .05$, a total sample size of $n = 90$ would be sufficient for an experimental power of .80 (two-tailed), using an effect size of $d = .3$ based on the mean effect size reported in the five studies by Gilbey and Hill (2012).

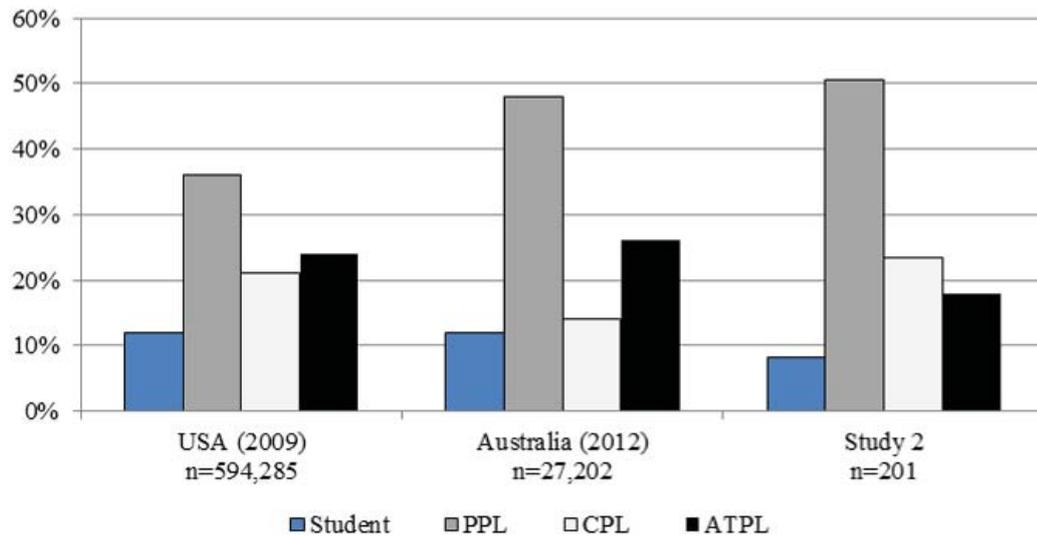
4.3 Results

4.3.1 General findings

The data were initially screened to identify any outliers and to establish that the data met the requirements for the application of parametric statistical analysis using the criteria of Iglewicz and Hoaglin (1993). No outliers were identified and the skewness and kurtosis of the dependent variable were within the limits recommended by Cameron (2004). The level of statistical significance was set at $\alpha = .05$ for all statistical tests and all tests were conducted as two-tailed. The internal consistency across all five scenarios, as measured by Cronbach's alpha, were $\alpha = 0.61$. Cronbach's alpha would not have been improved by removing a scenario.

Overall, 278 participants completed the survey. As was determined in Study 1, the distribution of pilots sampled in this study, by licence type, was compared to the distributions of active Australian and US pilots (CASA, 2013; FAA, 2009) as shown in Figure 8. A chi-square test of independence was performed to examine the relationship between the distributions of Study 2 licence holders and the population licence holders. No significance was found, $\chi^2 (3) = 4.09$, $p = .25$, which suggests

that the distribution of the pilots sampled in this study, by licence type, was similar to that of active Australian and US pilots.



PPL, private pilot licence; CPL, commercial pilot licence; ATPL, air transport pilot licence.

Figure 8: Comparison of Study 2 pilots to population totals by licence type

Participants made either 0 ($n = 69$), 1 ($n = 65$), 2 ($n = 60$), 3 ($n = 39$), 4 ($n = 31$) or 5 ($n = 14$) disconfirming choices out of a possible five. Participants selected the disconfirming choice less than half the time in 194 (69.8%) cases, whereas 69 (24.8%) participants made no disconfirming choice in all five scenarios. Only 14 (5.0%) made a disconfirming choice in all five scenarios. The mean number of disconfirming choices across all five scenarios was 1.78 ($SD = 1.49$). A single-sample t -test (two-tailed), using the total number of disconfirming choices for each participant as data and a test value of 1.66 (the mean number of disconfirming choices that would be expected if participants selected their responses by chance alone), was not significant, $t(277) = 1.39$, $p = .17$, indicating that participants performed no differently from what would be expected by chance. That is, there was

no evidence that participants were seeking disconfirming evidence when testing their hypothesis.

4.3.2 Novice vs Experienced Pilots

An independent-sample t -test was conducted to compare the mean number of disconfirmatory choices between experience groups: novice pilots ($n = 169$) and expert pilots ($n = 86$). Levene's test for equality of variances was not significant, confirming that the data met the assumption underlying the independent sample t -test. There was evidence of a significant difference between the two groups, $t(253) = 2.00$, $p = .046$, $d = .26$, with novice pilots selecting the disconfirmatory choice more often ($M = 1.86$, $SD = 1.46$) than experienced pilots ($M = 1.48$, $SD = 1.47$).

Two single-sample t -tests (two-tailed), using the total number of disconfirming choices for each experience group (novice vs. expert) as data and a test value of 1.66 (the mean number of disconfirming choices that would be expected if participants selected their responses by chance alone) was conducted. There was no evidence of a difference between the test value and the response of the novice group, $t(168) = 1.82$, $p = .070$ or of the expert group, $t(168) = 1.16$, $p = .25$, indicating that both groups performed no differently from what would be expected by chance.

4.3.3 VFR Flight into IMC Experience

An independent-sample t -test was conducted to compare the mean number of disconfirmatory choices between the pilots that had experienced VFR flight into IMC

($n = 112$) and pilots that had not experienced VFR flight into IMC ($n = 146$). Levene's test for equality of variances was not significant, confirming that the data met the assumption underlying the independent sample t -test. There was evidence of a significant difference between the two groups, $t(256) = 2.71$, $p = .007$, $d = .35$: pilots who had not experienced VFR flight into IMC selected the disconfirming choice more frequently ($M = 1.96$, $SD = 1.48$) than pilots who had experienced VFR flight into IMC ($M = 1.46$, $SD = 1.42$).

Two single-sample t -tests (two-tailed), using the total number of disconfirming choices for the two separate VFR flight into IMC experience groups (past experience of VFR flight into IMC vs. no experience) as data and a test value of 1.66 (the mean number of disconfirming choices that would be expected if participants selected their responses by chance alone) were conducted. There was no evidence of a difference between the test value and the response of pilots who had experienced VFR flight into IMC, $t(111) = 1.46$, $p = .15$, showing that the group performed no differently from what would be expected by chance. There was evidence of a difference from the test value for pilots who had not experienced VFR flight into IMC, $t(145) = 2.50$, $p = .016$, showing that this group of pilots performed better than chance.

The characteristics of the pilots in these two groups were explored. Table 6 shows that pilots who had experienced VFR flight into IMC were significantly older, $t(250) = 4.39$, $p < .001$, $d = .55$, and more likely to hold a current or non-current IR compared to the group of pilots that had not experienced VFR flight into IMC, $\chi^2(2) = 31.74$, $p < .001$. Pilots who had not experienced VFR flight into IMC were more likely to be SPs or PPL holders, $\chi^2(3) = 13.12$, $p = .004$. Nonetheless, there was no

significant difference in flight experience between the two groups, $t(253) = 1.51$, $p = .130$, $d = .19$.

Table 6: *Characteristics of pilots with and without past experience of VFR flight into IMC*

	Pilots who have experienced VFR flight into IMC	Pilots who have not experienced VFR flight into IMC
Mean Hours	2350 ($SD = 3824$)	1616 ($SD = 3834$)
Mean Age	43.78 ($SD = 16.0$)	35.50 ($SD = 13.9$)
<u>IR Status</u>		
Current IR	54 (21%)	43 (17%),
Non-current IR	26 (10%)	12 (5%),
No IR	32 (12%)	91 (35%)
<u>Licence Held</u>		
Student	2 (1%)	19 (7%)
PPL	53 (21%)	76 (30%)
CPL	31 (12%)	29 (11%)
ATPL	24 (9%)	22 (7%)

4.3.4 Instrument Rating

A one-way ANOVA was conducted to compare the mean number of disconfirmatory choices among the three IR groups: pilots who held a current IR ($n = 97$), those with a non-current IR ($n = 38$) and those who never held an IR ($n = 138$). There was no evidence of a difference between the responses of the three groups, $F(2,254) = 2.46$, $p = .088$, $\eta^2 = .019$.

Table 7 provides the demographic data of the three groups. There was a significant difference among the three groups for flight experience, $F(2,253) = 29.3, p < .001, \eta^2 = .19$. *Post hoc* (Tukey) tests revealed a significant difference among all three group ($p < .05$). There was no significant difference in age among the three groups, $F(2,251) = 2.76, p = .065, \eta^2 = .02$.

Table 7: *Demographic characteristics of pilots IR status*

	Current IR	Non-Current IR	No IR
Mean Hours	3919 ($SD = 5388$)	2213 ($SD = 2636$)	281 ($SD = 395$)
Mean Age	37 ($SD = 13$)	44 ($SD = 17$)	39 ($SD = 16$)

4.3.5 Licence Type

A one-way ANOVA was conducted to compare the mean number of disconfirmatory choices among pilots based on the type of licence they held: SP ($n = 21$), PPL ($n = 130$), CPL ($n = 60$) and ATPL ($n = 46$). A significant difference was found, $F(3,256) = 4.12, p = .008, \eta^2 = .047$. *Post hoc* (Tukey) tests revealed that the SP group selected the disconfirmatory evidence more often than all other licence groups ($M = 2.76, SD = 1.41, p < .05$). No significant difference was found among the PPL group ($M = 1.70, SD = 1.51$), the CPL group ($M = 1.72, SD = 1.24$) and the ATPL group ($M = 1.46, SD = 1.52$).

Table 8 provides the demographic data of these four groups. As expected, the more advanced pilot licence groups (i.e., CPL and ATPL) had the more flight experience.

A one-way ANOVA was conducted to compare experience levels (flight hours) among the four licence groups. There was a significant difference among the groups, $F(3,253) = 80.1, p < .001, \eta^2 = .49$. *Post hoc* (Tukey) tests revealed ATPL holders were significantly ($p < .001$) more experienced than all other groups but no difference was seen among the other groups. A one-way ANOVA was conducted to compare age between licence groups. A significant difference was apparent, $F(3,251) = 10.1, p < .001, \eta^2 = .11$. *Post hoc* tests (Tukey) revealed the SPs and CPL holders were significantly younger than the PPL and ATPL groups ($p = .005$).

Table 8: *Demographic characteristics of the pilot licence groups*

	Student	Private	Commercial	Air Transport
Mean				
Hours Flown	56 (<i>SD</i> = 48)	577 (<i>SD</i> = 1276)	1149 (<i>SD</i> = 1191)	7625 (<i>SD</i> = 5996)
Mean Age	32 (<i>SD</i> = 13)	43 (<i>SD</i> = 16)	32 (<i>SD</i> = 12)	41 (<i>SD</i> = 13)

4.3.6 Age

A one-way ANOVA was conducted to compare the mean number of disconfirmatory choices among the four age groups: 0–20 years ($n = 18$), 21–40 years ($n = 132$), 41–60 years ($n = 75$) and 61+ years ($n = 27$). No significant difference was found, $F(3,251) = 5.53, p = .053, \eta^2 = .040$.

4.4 Discussion

The findings of this study suggest that participants had an overwhelming preference to seek confirmatory evidence when assessing the weather conditions. In each of the five presented scenarios, this strategy could have provided a false positive. These findings suggest that when pilots are assessing the weather conditions to determine if they can maintain VFR flight, they may inadvertently be putting themselves into danger by relying on evidence which is consistent with their hypothesis. Over-reliance on the positive testing strategy, as demonstrated in this study, can lead to serious consequences. The findings in this study provide evidence to support Hypothesis 1 and are consistent with Gilbey and Hill's (2012) study of confirmation bias in GA lost procedures.

In addition to investigating the general decision strategy pilots used when assessing weather conditions, this study explored a number of factors that may be associated with the decision strategy chosen by participants. The first finding of interest was the role experience played in a pilot's decision making strategy. The findings of this study suggest experienced pilots have a greater preference for using the confirmatory strategy compared with novice pilots. These findings are consistent with previous research (e.g., English et al., 2006; Northcraft & Neale, 1987) on the effect of experience and cognitive biases. It must be noted that the significance level and the effect size were relatively small ($p = .046$, $d = .26$) suggesting that the difference between the two experience groups is relatively small at best. These findings provide some support for Hypothesis 2, in that experience level did not provide a reliable indicator of the type of hypothesis testing strategy participants were likely to prefer.

These findings suggest that even as pilots gain experience, they do not develop or retain cognitive tools to avoid or reduce the effects of confirmation bias. This also parallels findings in the area of risk perception; for example, Hunter's (2002) study of pilots' risk perceptions in adverse weather found that as pilots gain experience, their perception of risk reduces. One explanation for this is that as pilots gain experience flying in and around adverse weather with no significant adverse consequences, their perception of the risk reduces (Transport Safety Board, 1996). A similar process may occur when pilots use a confirmatory strategy to test a hypothesis; each time a pilot uses this confirmatory strategy with no adverse consequences, this reinforces the use of this strategy. These findings are also consistent with the findings of Study 1, in which experience levels had little impact on reducing the anchoring effect.

The analysis of pilot age and licence group broadly supports Hypotheses 3 and 4; in other words, neither provided a reliable indicator of the type of hypothesis testing strategy that was used. An interesting finding was that SPs outperformed other licence groups. The results suggest SPs have a far greater preference to select disconfirming evidence than any other licence group. However, as the number of SPs captured was very small, these findings may not necessarily provide a reliable indicator of SPs' true hypothesis preference. A larger SP sample size would be necessary to validate these findings. One possible explanation is SP were not comfortable flying in these conditions, therefore sought information to confirm it is 'not safe' to fly.

A difference was observed between groups that had experienced VFR flight into IMC, and those that had not. The results suggest that participants who had previously experienced VFR flight into IMC were more likely to seek confirmatory information compared to those that had not flown into IMC. A potential explanation for this finding is that pilots who are more likely to seek out confirmatory evidence are more likely to find false positive evidence that it is possible to maintain flight under VFR, resulting in their flying into IMC. A closer look at the differences in the two groups indicated pilots who have experienced VFR flight into IMC tended to be older and more likely to hold an IR (current or non-current) but the groups were not significantly different in regards to experience level. The relevance of the IR may be an important factor between these two groups. The higher level of training involved to obtain the skills necessary for the IR may have given these pilots greater confidence in their ability to operate safely in and around adverse weather.

Of particular interest is the group of pilots who have held an IR but no longer keep it current. The majority of the non-current IR pilots (70%) were in the group who had experienced VFR flight into IMC. The cost of remaining current can be challenging for a private pilot, so it may only be viable for pilots flying commercially to keep their IRs current. However, the initial skills acquired for gaining an IR may give a pilot unwarranted confidence in his/her ability to operate around adverse weather. IR skills degrade quickly without regular practice and therefore may be of little use if the pilot does enter IMC conditions (FAA, 2004; Wiggins et al., 2012). Pilots who hold a non-current IR may be at greater risk of putting themselves into danger by operating in and around IMC during a VFR flight. IR flight skills have been highlighted as a contributing factor in a VFR flight into IMC accident (TSB, 1996).

The accident report suggested the pilot's confidence about operating in adverse weather may have been enhanced due to holding an IR, which led to the VFR flight entering IMC and crashing a short time later (TSB, 1996). Further research is necessary to understand the role that holding a non-current IR and current IR has on VFR flight in IMC events.

There were a few aspects of this study that potentially may limit the generalisation and validity of these results. As discussed previously, the study involved the use of static images; whether similar results would be obtained in actual flight conditions cannot be determined. In real-life situations, the time and workload pressures associated with flying the aircraft may lead to increased stress levels, which have been shown to deteriorate decision making ability as people place greater reliance upon heuristics (Wickens et al., 1993). Caution also needs to be applied when generalising these results to the GA pilot population. Although this study captured a wide spectrum of pilots, as highlighted in Figure 8, the core sampling method for this study used convenience sampling at a New Zealand-based international flight training organisation and advertising through two major aviation internet forums and may not necessarily reflect the wider pilot population.

This study suggests that weather-related decisions can be influenced by confirmation bias; however, further research in this area is necessary to get a better understanding of the full extent of this factor. Two areas of interest could provide the next step for confirmation bias. First, exploring different weather scenarios would give a better understand of the influence of confirmation bias over a wide spectrum of weather

scenarios (e.g., in very poor weather). Second, a simulator study would provide some insight into how flight demands influence a pilot's hypothesis testing strategy.

In conclusion, the findings of the current study suggest that confirmation bias may occur in weather-related decision making. When pilots are assessing the weather conditions, they showed a greater preference to use the positive testing strategy by seeking confirmatory evidence when determining whether or not they could remain in VFR flight. Similar to Study 1, experience had little impact on the preferred hypothesis testing strategy. Finally, there is some evidence to suggest that previous experience of VFR flight into IMC may provide an indicator of the type of hypothesis strategy a pilot is likely to use: pilots who had experienced VFR flight into IMC were more likely to seek confirmatory evidence than those who had not.

CHAPTER FIVE

STUDY 3

OUTCOME BIAS IN VFR FLIGHT INTO IMC AND THE INFLUENCE OF OUTCOME INFORMATION ON THE DECISION PROCESS

5.1 Introduction

Outcome bias refers to the tendency to judge the quality of a decision by its eventual outcome instead of judging it based on the quality of the decision at the time it was made (Baron & Hershey, 1988). Although such a strategy is common—for example, it is used in court when considering sentencing—Mazzocco et al. (2004) argued that logic dictates that two identical decisions or behaviours, each of which is based on the same information and made under the exact same set of circumstances, one of which has a relatively innocuous outcome and the other has an unfavourable outcome, should be evaluated equivalently, as both decisions were made while the outcome was unknown. Nevertheless, Gino, Moore, and Bazerman (2009) reported that decisions which led to more severe outcomes were more likely to be ethically condemned, and Gino, Shu, and Bazerman (2010) found that negative outcomes were more likely to be punished.

It has consistently been found that people tend to overlook bad decisions where the outcomes have been relatively good (e.g., near misses or situations which went awry

but turned out well in the end) but focus on decisions where the outcomes have been bad (e.g., in fatal accidents; Tinsley, Dillon, & Madsen, 2011). This bias may mean that lessons that may be learned from near misses may be lost to all except those directly responsible (Dillman, Voges, & Robertson, 2007). What seems to be overlooked is that it may be just luck or good fortune that decides whether the outcome of a decision is good or bad; that is, 'if conditions shift slightly, or if luck does not intervene,' a disaster may result (Tinsley et al., 2011, p. 90).

Wrongly interpreting outcome information may potentially compromise what may be learned from an event (Fischhoff, 1975). For example, good outcomes resulting from bad decisions may [wrongly] lead people to believe that systems are resilient to threats (Tinsley et al., 2011). A pilot may successfully fly through adverse weather but in doing so, he or she may have taken significant risk, and the flight's safe outcome may have been more down to luck rather than good judgment. Some situations may justify the use of outcome information, such as when the full context of the decision is not known. Nevertheless, in these situations people may give outcome information more importance than it deserves (Baron & Hershey, 1988).

Of particular interest in this present study is the possibility that the outcome information of third-party flights may vicariously influence the decision process of other pilots. The influence of outcome information and the impact it has on pilots' decision making has been explored in the area of commercial aviation. Studies in commercial aviation (e.g., Rhoda & Pawlak, 1999a, 1999b) suggest that outcome information from other flights was a significant cue pilots used when deciding whether to fly through or deviate around adverse weather, especially if the aircraft

was close to its destination airport. A similar influence of outcome information may occur when a pilot is conducting a VFR flight. Does the outcome information of previous flights bias subsequent decisions to continue a VFR flight?

This study explored pilots' use of outcome information in weather-related decision making. It is predicted that outcome information will influence how participants rate decision makers' judgments, consistent with similar studies on outcome bias (e.g., Baron & Hershey, 1988):

Hypothesis 1: Outcome information will influence how participants rate the decision maker's judgment, with positive outcomes being rated more favourably than negative outcomes.

This study also explored the role outcome information plays on risk assessment and a participant's own decision to complete a VFR flight. The role that outcome information may have on risk assessment is of particular interest. Risk perception and risk tolerance have received considerable attention in regards to VFR flight in IMC (Hunter, 2002, 2006; Pauley, O'Hare, & Wiggins, 2008). Consistent with Hypothesis 1, it was predicted that outcome information will influence pilots' assessments of risk:

Hypothesis 2: Outcome information will influence how participants rate the amount of risk undertaken, with positive outcomes being rated as being less risky than negative outcomes.

Consistent with previous studies (Rhoda & Pawlak, 1999a, 1999b), it is predicted that outcome information will influence a pilot's own decision to conduct a VFR flight:

Hypothesis 3: Outcome information will influence a participant's own decision to conduct a VFR flight, with a positive outcome being more likely to lead participants to conduct the same flight compared to a negative outcome.

Similar to Studies 1 and 2, a number of pilot characteristics were explored, with the influence of experience being the primary interest. As outlined in Study 1, understanding how experienced pilots make decisions may help in training inexperienced pilots (Wiggins & O'Hare, 1995). Consistent with the findings in the first two studies and other experienced-based research (Northcraft & Neale, 1987), it is expected that there will be little difference between expert and novice pilots in regards to outcome bias.

Hypothesis 4: No difference will be seen between novice and expert pilots in regards to the effect of outcome bias in assessments of decision quality, risk assessment and participants' own decisions to conduct a VFR flight into IMC.

The type of pilot licence held is expected to mirror the hypothesis of experience level (i.e., the type of licence held will not influence the impact of outcome bias). Age is also not expected to influence the effect of outcome bias (i.e., there will be no relationship between age and the degree of outcome bias). No specific hypothesis was made for the impact of IR status. Unlike general pilot licence groups (e.g., PPL,

CPL), IR status does not necessary reflect experience level but the advanced training the pilot has chosen to undertake.

Hypothesis 5: The type of licence held (i.e., PPL, CPL or ATPL) will not influence outcome bias.

Hypothesis 6: Age will not influence outcome bias.

The primary aim of this study was therefore to investigate whether outcome bias occurs in judging weather-related decision making. Additionally, this study explored (i) if outcome information influences subsequent weather-related decisions and (ii) the effect of experience on outcome bias in weather-related decision making.

5.2 Method

5.2.1 Participants

Three hundred pilots participated in this study. The mean age of the participants was 37.48 years ($SD = 13.96$) with a mean flying experience of 3595 hours ($SD = 4845$, ranging from 12 to 22,000 hours). Participants were classified as either experts ($n = 176$) or novices ($n = 104$) based on the same criteria as Study 1 (20 did not provide flight experience). PPL holders made up 20.3% of the participants, with the rest made up of SPs (2.7%), CPL holders (38.7%) and ATPL holders (32.7%).

5.2.2 Materials and Design

5.2.2.1 Overview

A vignette-based study was designed to evaluate the effect of outcome information on weather-related decision making. Specifically, this study followed a similar methodology to that of Baron and Hershey (1988) and Mazzocco et al. (2004), whereby participants are asked to rate a series of scenarios on a number of relevant dimensions. The benefits and potential shortcomings of using this methodology were discussed in Study 1; that is, a scenario-based methodology facilitates a wide distribution of the material while keeping close control over the study variable. As before, this methodology is a common method used in aviation psychology research (e.g., Gilbey & Hill, 2012; Hunter et al., 2003; Wiggins & O'Hare 2003a).

5.2.2.2 Development of Scenarios

To explore outcome bias in weather-related decision making, five scenarios were developed. In each scenario, participants were asked to review a pilot's VFR flight in a Cessna 172. Participants were provided with all the information a pilot would use to conduct the flight: weather information, route map and general information about the time of day and the length of the flight. Each scenario was developed to satisfy the following criteria: (i) the route map was over flat terrain with no significant terrain or airspace, (ii) the weather information was above VFR metrological minima in uncontrolled airspace (at least 5 km visibility and a cloud base at least 500 ft¹³), (iii) the general information contained within the scenario was designed to be

¹³ This satisfies the minimum safe height of 500 ft.

favourable for conducting a VFR flight, (i.e., daylight hours, route conducted beforehand), and (iv) all five scenarios were similar.

For example, one of the scenarios read:

A pilot is conducting a VFR cross-county flight in a Cessna 172. The pilot has flown the same route several times before. The flight should take about an hour, conducted in daylight hours. The weather information available to the pilot for the flight is as follows.

Current Weather

Departure (A): wind variable, 3 knots; visibility, 16 km; light rain with a cloud base at 1800 ft.

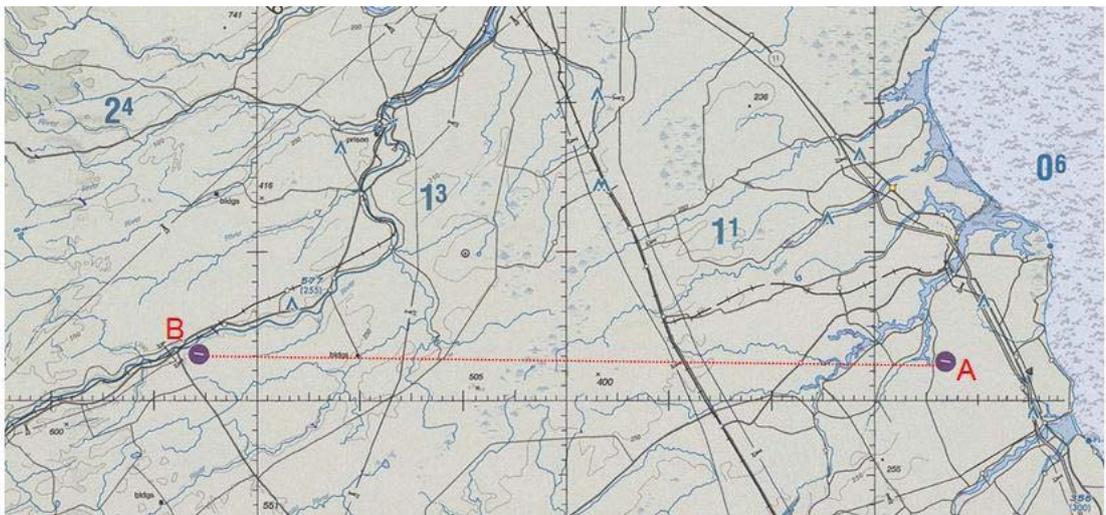
Destination (B): wind variable, 2 knots; visibility, 16 km; light drizzle with a cloud base at 1500 ft.

Area Forecast

Visibility: 16 km with scattered light drizzle.

Cloud broken to overcast at 1500 ft, with cloud tops at 24,000 ft.

Turbulence: nil.



The weather information provided in each scenario was based on actual weather reports at the time of real VFR flight into IMC accidents. For example, the weather information in the example scenario is from a fatal flight of a Cessna 350 Corvalis which inadvertently flew into IMC, resulting in spatial disorientation, loss of control and flight into terrain (NTSB, 2013).

After reviewing the information, the online platform randomly assigned participants with outcome information. The type of outcome information (i.e., negative, positive or no outcome) remained the same across all five scenarios once the initial outcome condition was randomly assigned. The between-subjects design adopted in this study (i.e., participants received the same outcome condition for all five tasks) was to ensure that participants were not influenced by a range of outcome conditions and is a consistent method in outcome bias studies (e.g., Mazzoco et al., 2004).

Participants in the positive outcome condition group read the following ending to all the scenarios:

‘You later find out that the flight safely landed at its intended destination.’

Participants in the no information/control group read only the common information; no outcome information was provided.

Participants in the negative outcome condition group read the statements outlined in Table 9.

Table 9: *Negative outcome statements for each scenarios in Study 3*

Scenario 1	En route, the weather deteriorated and the pilot decided to conduct a precautionary landing (an off-field landing). On landing in a paddock, the aircraft encountered a soft area, resulting in the nose gear breaking. The pilot escaped injury but the aircraft sustained substantial damage.
Scenario 2	En route, the pilot made a diversion to remain in VMC and landed at an airfield halfway along the intended route. The pilot reported that the weather was as per forecast.
Scenario 3	The flight inadvertently flew into cloud, resulting in the pilot losing control of the aircraft and crashing. The pilot received serious injuries.
Scenario 4	The pilot contacted the air traffic service and advised them that they had inadvertently flown into IMC conditions. Air traffic services were able to identify the aircraft on radar and provide navigational assistance allowing, the pilot to regain VFR and return to the departure aerodrome.
Scenario 5	The flight inadvertently flew into cloud, resulting in the pilot losing control of the aircraft and crashing. The pilot received minor injuries.

Each scenario was reviewed by several aviation employees at a New Zealand-based international flight training organisation with GA flying experience. The feedback from these reviews confirmed each scenario was consistent with real-world scenarios.

5.2.3 Dependent Measures

After reading each scenario, participants were instructed to rate each of the following dimensions using a nine-point scale:

- Decision quality ('How do you rate the pilot's decision to conduct the VFR flight?') on a scale ranging from 1 = *very good decision* to 9 = *very poor decision*.
- Risk rating ('How much risk do you think the pilot took in attempting this VFR flight?'), where 1 = *very low risk* and 9 = *very high risk*.
- Whether or not the respondent would conduct the same flight ('You wish to conduct the same flight. Considering the information available to the pilot before he/she departed, indicate your level of agreement with the following statement: I would be able to safely conduct the flight VFR. '), where 1 = *strongly agree* to 9 = *strongly disagree*.

Participants were also asked to provide the following demographic information: age, flight hours, pilot licence held (SP, PPL, CPL or ATPL) and IR status (holding a current IR, holding an IR that was not current or never held an instrument rating). A sample of the questionnaire is provided in Appendix E.

5.2.4 Procedure

Participants were recruited either through an email invitation to students and staff at a New Zealand-based international flight training organisation, or through two leading GA internet forums between March and May 2014. The invitation included a

link that directed participants to the online questionnaire (Survey Gizmo). Participants were first provided with a brief introduction to the questionnaire and informed that the study had undergone an ethical review. A review of the ethics of this study was undertaken using the guidelines provided by the Massey University Ethics Committee. The project was deemed to be of low risk and therefore suitable for a low-level notification to Massey University Human Ethics Committee. A copy of the low-risk notification can be found in Appendix F.

The online platform randomly assigned each participant to one of three outcome conditions: positive outcome ($n = 91$), negative outcome ($n = 100$) and no outcome/control ($n = 109$), with the outcome condition remaining the same for each participant across the five scenarios once the first outcome condition was assigned. Next, the participants evaluated each of the five scenarios. The order in which the scenarios appeared was also randomised by the online platform. Finally, participants completed the demographic section.

No time limit was imposed on participants completing the questionnaire. The questionnaire was conducted anonymously and therefore it was not possible to determine which recruitment method participants were recruited from. A *priori* power analysis using the software G*Power (Erdfelder et al., 1996) was used to determine that with $\alpha = .05$, a total sample size of $n = 150$ (split evenly among cells) would be sufficient for an experimental power of at least .80, assuming a mean effect size of $f = .3$ (based on that reported in studies by Mazzocco et al. (2004) and Gilbey, Tani, & Tsui (2015)) for a between-subjects ANOVA.

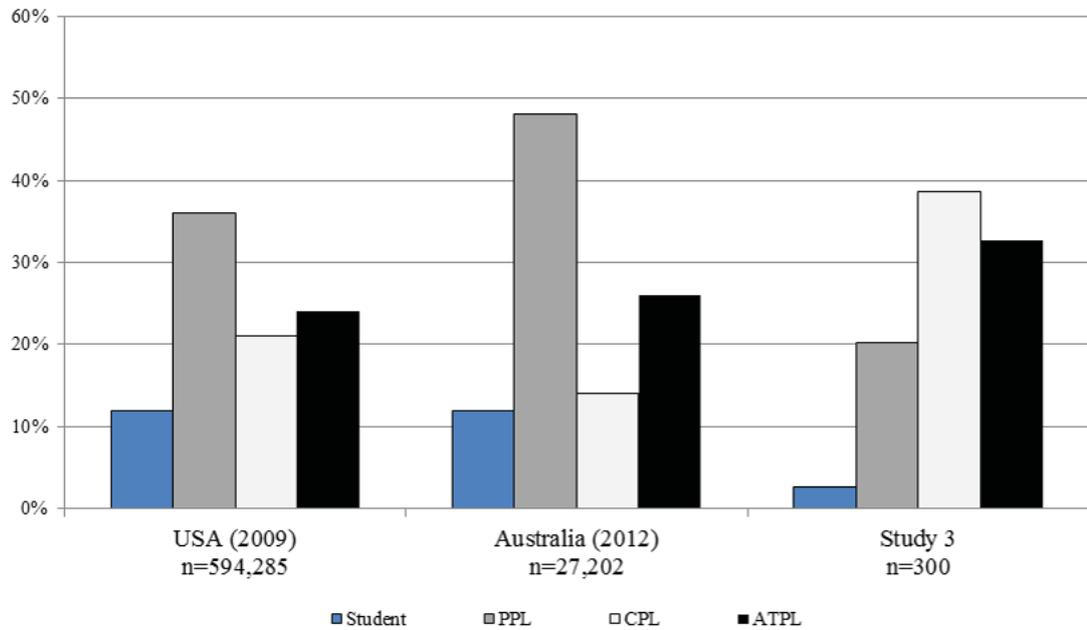
5.3 Results

5.3.1 General findings

As was carried out in previous studies, the data were initially screened to identify any outliers and to establish that the data met the requirements for the application of parametric statistical analysis. No outliers were identified, and the skewness and kurtosis of the dependent variable was within the limits recommended by Cameron (2004). The level of statistical significance was set at $\alpha = .05$ for all statistical tests and all tests were conducted as two-tailed. The internal consistency (Cronbach's alpha) was calculated for all three dependent variables across all five scenarios for decision quality ($\alpha = 0.72$), risk assessment ($\alpha = 0.74$) and safety assessment ($\alpha = 0.78$). Cronbach's alpha would not have been improved by removing a scenario.

Overall, 300 pilots completed the questionnaire. As in Studies 1 and 2, the distribution of pilots sampled in this study, by licence type, was compared to the distribution of active Australian and US pilots (CASA, 2013; FAA, 2009), as shown in Figure 9. A chi-square test of independence was performed to examine the relationship between the distributions of Study 3 licence holders and the population licence holders. There was evidence of a significant difference, $\chi^2(3) = 18.5$, $p < .001$, which suggests that pilots sampled in this study, by licence type, was different from the wider population characteristics. Considering that this study captured a larger sample size using the same distribution channels as the previous studies, this finding was somewhat surprising. Study 3 contained a greater proportion of CPL and ATPL holders, which correlated with an overall higher mean flight experience level

and age compared to Studies 1 and 2. Although this deviates considerably from the population distribution, the benefit of having a generally more experienced sample is that it allowed a more even comparison between the expert and novice pilot groups.



PPL, private pilot licence; CPL, commercial pilot licence; ATPL, air transport pilot licence.

Figure 9: Comparison of Study 3 pilots to population totals (Australia and USA)

5.3.2 Decision Judgment

A 3 (outcome condition) \times 2 (experience) between-subjects ANOVA revealed a significant main effect between the outcome condition and the quality of the decision, $F(2,280) = 10.43$, $p < .001$, $\eta^2 = .071$. *Post hoc* tests (Tukey) revealed that scenarios in the negative outcome condition were judged as being significantly worse than scenarios in the positive outcome condition ($p < .001$). In the control condition, the decision quality for each scenario was judged as being significantly better than it was in the negative outcome condition ($p = .002$) but no differently from how it was

judged under the positive outcome condition ($p = .28$). Table 10 displays the decision quality assessments for each outcome condition. There was no evidence of a main effect between experience groups (novice vs. expert), $F(1,280) = 0.095$, $p = .79$, $\eta^2 < .001$ and no evidence of an interaction between outcome condition and experience, $F(2,280) = 0.43$, $p = .65$, $\eta^2 = .003$.

Table 10: *Participants Decision Quality Assessment (1 = very good decision, 9 = very poor decision), Risk Assessment (1 = very low risk, 9 = very high risk) and Assessment of conducting the same VFR flight (1 = safe to conduct flight, 9 = not safe to conduct flight) between outcome condition.*

	Positive	Control	Negative
Decision Quality	5.01 ($SD = 1.22$)	5.33 ($SD = 1.51$)	6.02 ($SD = 1.43$)
Risk Assessment	5.03 ($SD = 1.17$)	5.40 ($SD = 1.50$)	5.96 ($SD = 1.43$)
Safety Assessment	4.16 ($SD = 1.33$)	4.76 ($SD = 1.77$)	5.36 ($SD = 1.69$)

5.3.3 Risk Assessment

A 3 (outcome condition) \times 2 (experience) ANOVA revealed a significant main effect between the outcome condition and risk assessment of the flight, $F(2,280) = 7.87$, $p < .001$, $\eta^2 = .054$. *Post hoc* tests (Tukey) revealed that scenarios in the negative outcome condition were perceived as being significantly riskier than those in the positive outcome condition ($p < .001$) and the control condition ($p = .016$). No significant difference was observed between the positive outcome condition and the

control condition ($p = .16$). Table 10 displays the decision quality assessments for each outcome condition. There was no evidence of a main effect of experience groups (novice vs. expert), $F(1,280) = 0.21$, $p = .65$, $\eta^2 = .001$ and no evidence of an interaction between outcome condition and experience, $F(2,280) = 0.85$, $p = .43$, $\eta^2 = .06$.

5.3.4 Would Pilots Conduct the Same Flight?

A 3 (outcome condition) \times 2 (experience) ANOVA revealed a significant main effect between the outcome condition and a decision to conduct the same VFR flight, $F(2,280) = 7.70$, $p = .001$, $\eta^2 = .052$. *Post hoc* tests (Tukey) revealed that scenarios with the negative outcome condition were judged as being significantly less safe for continuing than those with the positive outcome condition ($p < .001$). Comparisons with the control condition revealed a significant difference from both the negative outcome condition ($p = .027$) and positive outcome condition ($p = .03$). Table 10 displays participants safety assessments for each outcome condition. There was a significant main effect of experience group (novice vs. expert), $F(1,280) = 4.59$, $p = .03$, $\eta^2 = .016$. Overall, experienced pilots assessed it to be safer to conduct the flight compared to novice pilots. Table 11 displays participants safety assessments for each experience level and outcome condition. There was no statistically significant interaction between outcome condition and experience, $F(2, 280) = 1.03$, $p = .36$, $\eta^2 = .07$.

Table 11: *Participants Assessment of conducting the same VFR flight (1 = safe to conduct flight, 9 = not safe to conduct flight) between experience level (novice vs. expert) and outcome conditions.*

	Positive	Control	Negative
Novice	4.76 (<i>SD</i> = 1.29)	4.80 (<i>SD</i> = 1.46)	5.60 (<i>SD</i> = 1.51)
Expert	3.96 (<i>SD</i> = 1.29)	4.74 (<i>SD</i> = 1.93)	5.13 (<i>SD</i> = 1.82)

5.3.5 Correlation

A Pearson's product-moment correlation was used to determine the relationships among each of the dependent variables (decision quality, risk assessment and safety assessment). There were strong, positive correlations among all dependent variables. The correlations among the three dependent variables are presented in Table 12.

Table 12: *Pearson's correlation between the three dependent variables (decision quality, risk assessment, safety assessment).*

	Decision Quality	Risk Assessment	Safety Assessment
Decision Quality		$r = .93^*$	$r = .84^*$
Risk Assessment	$r = .93^*$		$r = .84^*$
Safety Assessment	$r = .83^*$	$r = .84^*$	

* Significance level, $p < .001$; $n = 300$

5.3.6 IR Status

A 3×3 ANOVA was conducted to compare pilots' IR Status (current IR, non-current IR or no IR) with the three outcome conditions. There was no evidence of a significant difference in either the interaction or main effect for all three dependent variables: IR status did not influence how pilots assessed decision quality, risk level and safety.

5.3.7 Age Groups

A 3×4 ANOVA was conducted to compare pilots' age groups (0–20 years, 21–40 years, 41–60 years, 61+ years old) with the three outcome conditions. There was no evidence of a significant difference in either the interaction or main effect for all three dependent variables. In other words, age did not influence how pilots assessed decision quality, risk level and safety.

5.3.8 Licence Groups

A 3×4 ANOVA was conducted to compare pilots' licence groups (SP, PPL, CPL or ATPL) with the three outcome conditions. No significance was found in either the interaction or main effect, for decision quality and risk assessment. However, there was evidence of a significant main effect between the four groups for safety assessment, $F(2,280) = 3.32$, $p = .020$, $\eta^2 = .033$. *Post hoc* tests (LSD) revealed that PPL holders were significantly different from CPL ($p = .018$) and ATPL holders ($p = .015$). The PPL holders rated the scenarios with the positive outcome condition as being less safe ($M = 5.07$, $SD = 1.27$) than did the CPL holders ($M = 3.95$, $SD =$

1.23) and ATPL holders ($M = 3.93$, $SD = 1.32$). There was no evidence of an interaction between pilot licence and outcome condition.

5.4 Discussion

The findings of Study 3 suggest that outcome bias can affect pilots' judgments of weather-related decisions, thus supporting Hypothesis 1 for this study. A comparison of the participants' assessments of the decision to conduct a VFR flight suggests that participants who were exposed to the negative outcome rated the decision to conduct the flight as being poor compared to participants who were exposed to the positive outcome. Even though the information available to the pilot was the same for both outcome scenarios, the participant's judgment of the decision to conduct the flight was influenced by an outcome assigned by the researcher.

Outcome bias was also observed in the evaluation of pilots' assessments of risk, with scenarios with positive outcomes assessed to be less risky than negative outcome situations, thus supporting Hypothesis 2 of this study. These findings may provide some insight on the impact outcome information may have on a pilot's perception of risk, particularly the influence of positive outcome information. Chapter 2 highlighted pilots' risk perceptions as one area that has been explored to help understand VFR flight into IMC accidents (Hunter, 2002, 2006). Repeated exposure to positive outcomes, whether it is the pilot's own flight experience or information from other's flights in deteriorating weather conditions, may lower and reinforce pilot's perceptions of risk when operating in deteriorating weather conditions. This cycle could have disastrous consequences: more positive outcomes may lead to a

reduced perception of risk, leading to a pilot taking increasing amounts of risk near deteriorating weather. Unfortunately, luck may not always be on the pilot's side to assist in a positive outcome each time.

The outcome bias observed in the participants' assessments to continue the flight was particularly interesting. The findings indicate that outcome information may influence how pilots make subsequent decisions when conducting a flight. When participants were exposed to a positive outcome, they were more likely to state they would conduct the same flight themselves than if they were exposed to a negative outcome. Considering the limited information that a pilot has available when conducting a VFR flight (especially in the cruise phase of a flight; often with limited weather reporting stations and only a general regional forecast), it is not unreasonable to expect pilots to use outcome information. However, the findings of this study suggest that pilots may be placing too much weight on outcome information. These findings are consistent with the findings of Rhoda and Pawlak (1999a, 1999b), which found that outcome information influenced the decision making process of commercial airline pilots about diverting or flying through adverse weather, especially close to their destination airport. The size of the correlation between all three dependent variables (decision quality, risk assessment, own assessment) suggests that all three are measuring the same (single) construct.

Outcome information could play an important role in supporting confirmation bias. Study 2 found evidence that pilots had a tendency to favour confirmatory evidence when deciding whether or not it is safe to continue a VFR flight. Positive outcome information from other flights, like that used in this current study, could potentially

reinforce confirmation bias in VFR flight, even though outcome information does not necessarily reflect if the conditions are suitable for safe flight. For example, if a pilot is conducting a cross-country flight, he or she is likely to be seeking information to determine if the weather conditions are suitable to continue. Positive outcome information from other flights is a possible cue a pilot could use to reinforce the hypothesis that safe flight is possible.

As was found in the previous two studies, little differences were observed between experienced and novice pilots, thus supporting Hypothesis 4. These findings are broadly consistent with previous research (e.g., English et al., 2006; Northcraft & Neale, 1987), in which experts and novices in a particular field were also reported to fall prey to cognitive biases. The only area in which a difference was observed was how the two experience groups assessed whether they would conduct the same VFR flight. Both groups were influenced by outcome bias; however, overall the more experienced pilots indicated they were more likely to conduct the same VFR flight compared to the novice pilots. These findings were not unexpected; as a pilot gains experience, he or she is likely to become more comfortable in poorer weather conditions and therefore, it is expected that the more experienced pilots would rate the risk involved as being lower compared to novice pilots. But this was not the case in these results, as no difference was observed in how the two experience groups assessed the risk of another pilot's VFR flight. These findings provide support for Wilson and Fallshore's (2001) study of ability bias in pilots' decision making and perceptions of risk. In general, people believe themselves to be superior to others in their skills and ability (Dunning et al., 1989). In the context of VFR flight into IMC, Wilson and Fallshore (2001) found that pilots overestimated both their own ability to

avoid and ability to successfully fly out of IMC. The findings of this present study suggest that experienced pilots assess their own risk in conducting a flight (i.e., they are more likely to conduct the same flight) differently from how they would rate the risk if someone else was conducting the flight (i.e., risk assessment of another flight).

The findings regarding age, licence type and IR status broadly support Hypotheses 5 and 6, in that neither provided a reliable indicator of reduced outcome bias. The only finding of note was that the more experienced licence groups (i.e., CPL and ATPL holders) tended to assess the scenarios with positive outcome situations as being more safe than the less experienced groups; however, overall, they still exhibited outcome bias. These findings were consistent with the experiencing findings of this study.

There are a few aspects of this study that potentially may limit the generalisation and validity of these results. First, the study used a scenario-based methodology in which participants self-reported on various dependent variables. Whether similar findings would be found in a real-world situation is difficult to determine. Nevertheless, it has been suggested that this type of methodology, at the very least, acts as an indicator of real-life behaviour (Exum et al., 2012, for a more complete review). Therefore, these findings should provide at least some indication of the impact of outcome bias in the real world. Caution also needs to be applied when generalising these results to the GA pilot population. Although this study captured a wide spectrum of pilots, as highlighted in Figure 9, the core sampling method for this study was through convenience sampling at a New Zealand-based international flight training organisation and advertising through two major aviation internet forums.

This study suggests that weather-related decisions can be influenced by outcome bias. Despite this, further research in this area is necessary to get a better understanding of the full extent. Two areas of interest could provide the next step for outcome bias. First, exploring different types of outcome information will give a better understanding of the influence of outcome bias over a wide spectrum of situations. For example, are VFR pilots influenced by outcome information differently depending on whether this information comes from a direct report (e.g., talking over the radio to another aircraft) or an indirect report (e.g., overhearing the outcome of another aircraft). Second, a simulator study would provide some insight into the effect of flight demands, especially if the outcome information is received in flight during adverse weather.

In conclusion, the findings of the current study suggest that outcome bias may occur in weather-related decision making. When pilots are assessing the decision quality and risk assessment of conducting a VFR flight, they may be placing too much emphasis on the outcome information of the flight. Furthermore, the findings of this study suggest that outcome information influences a pilot's own decision about conducting a VFR flight. This could be particularly hazardous for a VFR pilot flying towards IMC. Similar to Study 1 and 2, experience had little impact on reducing the rate at which the cognitive bias occurred.

CHAPTER SIX

DEBIASING WEATHER-RELATED DECISION MAKING

6.1 Introduction

This chapter in the thesis introduces a shift in the research methodology; from descriptive studies to experimental work. Studies 1 to 3 explored the presence of three cognitive biases in weather-related decision making: the anchoring effect, confirmation bias and outcome bias. Each study found evidence to suggest that when pilots were assessing the weather conditions, their decision making process was influenced by these cognitive biases. Each of these cognitive biases has the potential to impair a pilot's situational assessment when approaching IMC. Considering the serious consequences that can arise from decision error and the part cognitive biases can play in influencing how we process information when making a decision, it is important to explore methods to improve decision making. Aiding pilots to avoid cognitive biases will enable them to make more informed decisions by forming an accurate assessment of the dynamic weather environment. These findings raised an additional research problem; can the cognitive biases observed in these previous studies be reduced through debiasing?

Research into cognitive biases has received significant attention (Lilienfeld, Ammirati, & Landfield, 2009). The presence and impact of various biases has been observed in a wide range of fields, from law enforcement to emergency medicine

(Ask & Granhag, 2005; Croskerry, 2002). Research exploring methods that can be used to correct or prevent these cognitive biases (debiasing) has started to receive increasing attention with the recognition that debiasing is a critical area for improving decision making. Exploring debiasing techniques has been a challenge, producing mixed results (Larrick, 2004).

6.2 Debiasing Techniques

The techniques suggested for use in debiasing have been varied. Most debiasing techniques have focused on encouraging a shift from intuitive decision making processes to analytical decision making processes (Clarkson, Emby, & Watt, 2002). As outlined in Chapter 2, it has been suggested that when making decisions individuals use one of two modes, either Type 1 (intuitive) or Type 2 (analytical) processes (Croskerry et al., 2013; Kahneman, 2011). Many of the decisions a person makes use Type 1 processes. Type 1 processes use heuristics and are largely automatic, fast and *usually* effective. Croskerry et al. (2013) suggested that a person spends about 95% of the time in this mode. The main shortcoming of Type 1 processes is that this mode can lead to errors. Type 2 processes, which use rules-based processes and are conducted under conscious control are reliable and effective. However, Type 2 processes take up a large amount of cognitive load. Debiasing techniques are aimed at shifting from Type 1 processing (automatic, heuristic) to Type 2 (controlled, rule-based), allowing for a more careful analysis of the information (Arkes, 1991).

One of the simplest techniques of debiasing is basic education or instruction (Larrick, 2004). One of the advantages of debiasing through instruction is that this method is less intrusive than other methods and provides the necessary tools for potentially avoiding biases in the future. Clarkson et al. (2002) explored debiasing of the outcome effect in an auditor review using instruction. The findings of this study found the use of strong debiasing instruction (e.g., encouraging participants to reflect more deeply on the evidence) was successful in mitigating the outcome effect. However, intervention which only advised participants on potential biases did not have a statistically significant effect. Similar findings were found in an attempt to reduce confirmation bias in aviation lost procedures by Gilbey and Hill (2012), and Gilbey et al. (2015).

Another debiasing technique is ‘training in rules’ (Larrick, 2004). It is assumed that most people have a basic understanding of some of the principles that lead to cognitive bias but have trouble knowing when to apply appropriate techniques to avoid them (Nisbett, Krantz, Jepson, & Kunda, 1983). Providing basic training has provided some successes in debiasing. For example, Larrick, Morgan, and Nisbett’s (1990) research on economic principles demonstrated that students could be trained to ignore the sunk costs in financial domains. Fong, Krantz, and Nisbett (1986) found that training in the principles of sampling and sample validity during short sessions was effective. Overall, this area of debiasing has demonstrated that brief, specific training with the aid of examples can have some effect. The main limitation of ‘training in rules’ is that only limited research has been carried out in highly complex, unfamiliar situations (Larrick, 2004; Larrick et al., 1990).

One of the more effective techniques of debiasing is the strategy of ‘consider the alternative’ (Fischhoff, 1982). This simple strategy has been shown to reduce overconfidence bias, hindsight biases, confirmation bias and the anchoring effect (Larrick, 2004; Mussweiler, Strack, & Pfeiffer, 2000). This strategy requires a person to think of alternative reasons why their initial judgment might be wrong. A number of cognitive biases (including the three cognitive biases explored in this thesis) are partly caused by a person focusing on a narrow set of information. The ‘consider the alternative’ strategy encourages a person to open up to a wider sample of information, which may counteract the effect of cognitive biases (Larrick, 2004). Arkes, Faust, Guilmette, & Hart (1988) explored debiasing of hindsight bias using the ‘consider the alternative’ technique. Arkes et al. (1988) used a similar technique to previous debiasing studies (e.g., Koriat, Lichtenstein, & Fischhoff, 1980), which required participants to write down reasons why alternative outcomes might be correct. The findings of these studies indicated that the use of this technique reduced but did not eliminate hindsight bias. Similarly, Kennedy (1995) found that outcome bias could be reduced by requiring participants to explain why a particular outcome might not occur.

Importantly, debiasing using the ‘consider the alternative’ technique has had some success in the real-world environment. For example, Arkes et al.’s (1988) study discussed earlier used the ‘consider the alternative’ technique to reduce hindsight bias among neuropsychologists. Similar, Mussweiler et al. (2000) required experts to assess the value of a used car after being exposed to an anchor. Participants were prompted to list anchor-inconsistent arguments before making their judgment. This study found that the bias was reduced but not eliminated.

Other debiasing techniques have received mixed results. Motivational techniques have been studied using incentives and accountability. The idea behind these techniques is that an individual will reflect more carefully if the stakes are higher (Larrick, 2004). These techniques have been found to be generally ineffective, reducing bias in only a handful of cases (Arkes, 1991). Larrick (2004) suggests that one of the reasons why incentives and motivational techniques have been less successful is because one of the assumptions that are required for this approach may not necessarily have been achieved. In other words, for this approach to improve decision making, decision makers need to possess effective decision making strategies which they have previously failed to apply. Camerer and Hogarth (1999) refer to this, stating that a decision maker must possess the necessary 'cognitive capital'. Therefore, a decision maker who does not have the right decision making strategy, may apply an inferior strategy when faced with motivational factors, resulting in a less than ideal outcome (Larrick, 2004).

Although there has been some success in debiasing, a range of studies have found debiasing to be predominately unsuccessful. Fischhoff (2002) noted that most debiasing techniques made little or no difference in hindsight bias and overconfidence bias. A study by Weinstein and Klein (1995) explored debiasing of optimistic bias. Over a series of studies, various types of debiasing techniques were used, from informing participants about relevant factors to having participants generate factors themselves. Overall, the debiasing techniques were unsuccessful in reducing optimistic bias.

An interesting finding from one of the debiasing techniques which required participants to focus on risk-increasing factors is that this technique exaggerated optimistic bias, indicating that some techniques of debiasing may actually make the impact worse. This has also been observed in a study by Roese (2004), where participants were asked to generate a long list of contrary reasons. If participants struggled to generate a large number of contrary reasons, they were more likely to convince themselves that their initial judgment must have been correct (Roese, 2004). Similar findings have been found with the effect of reasoning on judgment. Halberstadt and Levine (1999) found that participants who were asked to analyse and list reasons for predicating an outcome performed poorer than participants that were explicitly told not to analyse their reasons.

6.3 Research Problem & Research Questions

The presence and potential serious consequences of cognitive biases in weather-related decision making points to the existence of an additional research problem; specifically, can the effects of cognitive biases be corrected or prevented through debiasing? Three additional research questions were developed to explore debiasing of the cognitive biases identified in the three descriptive studies.

Research Question 4: Can debiasing reduce the rate at which the anchoring effect occurs in weather-related decision making? Specifically, after having been exposed to the debiasing intervention ‘consider the alternative’ technique, do pilots have a reduced tendency to anchor on outdated weather forecasts?

Research Question 5: Can debiasing reduce the rate at which confirmation bias occurs in weather-related decision making? Specifically, after having been exposed to the debiasing intervention ‘consider the alternative’ technique, do pilots have a reduced tendency to seek confirmatory evidence?

Research Question 6: Can debiasing reduce the rate at which outcome bias occurs in weather-related decision making? Specifically, after having been exposed to the debiasing intervention ‘consider the alternative’ technique, do pilots have a reduced tendency to use outcome information?

6.4 Study 4 to 6 Overview

To attempt to answer these research questions, practical debiasing techniques were applied to investigate how decision making may be improved in real-life situations. Although there has been some success in debiasing, these results have often been achieved using intrusive techniques (e.g., instructing participants to write down an alternative argument; Mussweiler et al., 2000). Intrusive methods of debiasing do not necessarily reflect how decision making can be improved in a real-world environment, particularly in a high workload and dynamic situation such as aviation. Therefore, the debiasing method used in these studies used nonintrusive techniques.

These debiasing studies will be targeted towards the group of pilots most at risk of VFR flight into IMC: relatively young and inexperienced pilots (AOPA, 2014). These pilots will receive training in the debiasing technique during the *ab initio* stage of pilot training. The distribution method of the material in these studies was paper-

based. Therefore, obtaining a sufficient sample size for all three debiasing studies was a concern. To overcome this challenge, participants completed the tasks for all three debiasing studies during the same session. The implications of and any potential impact this procedure may have had will be discussed with each study.

CHAPTER SEVEN

STUDY 4

DEBIASING THE ANCHORING EFFECT IN WEATHER-RELATED DECISION MAKING

7.1 Introduction

The aim of Study 4 was to explore debiasing of the anchoring effect in weather-related decision making. If the debiasing technique was successful, pilots would demonstrate a reduced tendency to anchor towards the initial piece of information received. As the anchoring effect can have serious implications on the safety of a flight, exploring practical methods to support pilots in weather-related decision making is an important step to improve safety.

Study 1 presented evidence that pilots anchored and under-adjusted when assessing both cloud height and visibility during weather-related decision making. Exploring methods that reduce the anchoring effect may aid pilots to assess the weather conditions accurately. The ‘consider the alternative’ debiasing technique was chosen for this study, as evidence suggests that it may mitigate the effects of a range of biases, including the anchoring effect (Larrick, 2004). For example, over a series of studies, Mussweiler et al. (2000) reducing the anchoring effect by increasing the accessibility of anchor-inconsistent knowledge. Therefore, consistent with a range of

debiasing studies using the ‘consider the alternative’ technique, the following hypotheses are derived:

Hypothesis 1: The anchoring effect on cloud assessments in weather-related decision will be reduced after receiving the debiasing technique ‘consider the alternative’.

Hypothesis 2: The anchoring effect on visibility assessments in weather-related decision will be reduced after receiving the debiasing technique, ‘consider the alternative’.

7.2 Method

7.2.1 Participants

One hundred and one pilots participated in this study. Participants were all enrolled in a flight training programme in which they learn to fly (to CPL level) and complete an academic programme of study. As the participants were enrolled in a structured training programme, when tested, they would have had 20–200 hours of flying experience (and, in some cases, more, due to the relatively common occurrence of participants entering the programme with previous flight experience). To protect confidentiality, neither demographic information nor flying hours were collected, thus ensuring that no participant could be identified. Although the lack of demographic and flight experience data limited the scope of the data analysis, this was an important element of the study design needed to obtain ethical approval.

7.2.2 Material & Design

7.2.2.1 Overview

Participants were assigned to one of two experimental groups, each of which completed the same tasks. The random assignment of participants in the groups was achieved using a randomising program available on Microsoft Excel. One group was exposed to the debiasing intervention but the other group had no debiasing intervention (the control group). The debiasing intervention was delivered as a short lecture, which was designed to provide participants with the tools necessary to reduce the anchoring effect on an outdated weather forecast, primarily concentrating on the two key cues experienced pilots use during weather-related decision making: cloud height and visibility (Wiggins & O'Hare, 2003a). The paper-based questionnaire, which both groups received, required participants to assess a series of in-flight images after being exposed to a weather forecast, which acted as the anchor. The benefits and shortcomings of using this methodology have already been discussed in detail in Chapter 3.

7.2.2.2 Development of the Debiasing Intervention

This study used the cognitive debiasing model suggested by Croskerry et al. (2013) (adapted from Wilson & Brekke, 1994). Croskerry et al.'s model contains four important steps, each of which are required to be completed for successful debiasing (and therefore optimal decision making): (i) awareness of the bias, (ii) motivation to correct the bias, (iii) awareness of direction and magnitude of the bias, and (iv) ability to apply an appropriate debiasing technique. Wilson and Brekke (1994)

suggested that the varying results in debiasing studies may be the result of the debiasing intervention not satisfying one or more of these steps.

Using the debiasing model suggested by Croskerry et al. (2013), a short lecture was developed which ensured that each of the four steps would be covered. The outline of the short lecture is as followed:

- i) *Awareness*: What is anchoring and adjustment and why does it happen?
- ii) *Motivation*: Examples of how the anchoring effect may influence assessment of the weather conditions and therefore increase the risk of VFR flight into IMC.
- iii) *Awareness of the direction and magnitude of the bias*: Examples of how to recognise anchoring and the magnitude of the bias.
- iv) *Debiasing technique*: An introduction to the ‘consider the alternative’ debiasing technique, providing examples of how to use it to reduce the anchoring effect.

After the short lecture, participants in the debiasing group received the same questionnaire material as the control group. Some studies have shown this method (a less intrusive strategy) to be successful in mitigating biases (Clarkson et al., 2002).

The data for all three debiasing studies (Studies 4–6) were collected at the same time. As a result, the short lecture which the intervention group received contained elements for all three debiasing studies. The impact of conducting all three debiasing interventions at the same time was largely mitigated by all three debiasing studies

using the same debiasing technique: ‘consider the alternative’. The lecture slides used in the debiasing intervention can be seen in Appendix I.

7.2.2.3 Development of the Questionnaire

The questionnaire material was based on Study 1 but the participants received a total of four scenarios rather than five. The number of scenarios was reduced to improve participation rate by reducing survey fatigue. The data from the tasks for Studies 5 and 6 were also collected at the same time, which resulted in participants receiving a total of 10 scenarios in the combined data collection. This combined data collection was necessary to overcome the considerable challenge of obtaining a sufficient sample size for all three debiasing studies. Four scenarios were retained for Study 4 based on the internal consistency (Cronbach’s alpha) of Study 1. Cronbach’s alpha for the original five scenarios was calculated for both visibility ($\alpha = 0.85$) and cloud height ($\alpha = 0.71$) assessment; in both cases, the removal of Scenario 5 produced the highest internal consistency score (visibility, $\alpha = 0.84$; cloud assessment, $\alpha = 0.69$) after Scenario 5 was removed.

Each scenario contained the same static images as Study 1 (the criteria for the images have been outlined previously). The key difference was in the presentation of these images. In Study 1, the high quality of the images was able to be retained by the use of an online platform. As this current study was paper-based, it was critical that the quality of the images was not compromised. To ensure that the high quality of the image was retained, each image was printed on high-quality photo print paper before being attached to the paper-based questionnaire.

The remainder of the study was the same as described in Study 1. The participants were asked to imagine that they were the PIC of a Cessna 172 on a VFR flight approaching their destination. Participants were provided with an in-flight static image of a simulated environment, looking through the cockpit window in their intended direction of travel, flying at 1000 ft above ground level. Participants had obtained a plain language text weather forecast for their destination prior to flight (provided by the researcher). The weather forecast conditions acted as the anchor. Approximately half of the participants received a 'poor' weather forecast (i.e., a low anchor with low cloud and low visibility). An example of a poor forecast is:

Wind at 270°T at 5 knots; visibility, 7 km; cloud broken at 1200 ft; temperature, 9°C; dew point, 6°C; QNH, 1012.

The other half of the participants were exposed to a good weather forecast (a high anchor with high cloud and high visibility), such as the following:

Wind at 270°T at 5 knots; visibility, 25 km; cloud broken at 3800 ft; temperature, 9°C; dew point, 6°C; QNH, 1012.

The weather conditions in the in-flight static image were exactly halfway between these two anchors (i.e., the visibility was 16 km and a cloud base at 2500 ft). The anchor weather condition pairs were chosen to ensure the weather presented remained realistic (e.g., very low cloud or very low visibility would be unrealistic for the weather in the static images) and the non-anchoring information (i.e., wind,

temperature etc) was set to conditions that were unlikely to influence the pilots' assessment of cloud and visibility (e.g., light winds). Each participant was randomly assigned to ensure they received a total of two high anchor scenarios and two low anchor scenarios (a within-subjects design).

To better reflect the varying real-life weather assessments that a pilot encounters, even on a single flight, each scenario had a slightly different anchor pairing (i.e., a high and a low weather forecast anchor for each). It is important to note that the weather conditions in the static images always remained the same and were always halfway between the two anchors (i.e., a cloud base at 2500 ft and a horizontal visibility of 16 km) and the anchor pairings were fixed for each scenario. A list of the anchor pairings for each of the four scenarios can be found in Table 13.

After reviewing the weather forecast (the anchor), participants were first asked to make a comparative judgment comparing the weather conditions in the in-flight image with the weather forecast ('Is the weather forecast worse, the same as or better than the conditions in flight?'). The purpose of the comparative question was to encourage participants to actually process and consider the given anchor values. Participants were then asked to make an absolute judgment of the cloud height and visibility distance. The questionnaire used in Study 4 can be inspected in Appendix G.

Table 13: *Four anchor pairings by scenario used in Study 4*

Scenario 1	High Anchor	Wind at 270°T at 5 kt; visibility is 25 km; cloud broken at 3800 ft; temp, 9°C; DP, 6°C; QNH, 1012
	Low Anchor	Wind at 270°T at 5 kt; visibility is 7 km; cloud broken at 1200 ft; temp, 9°C; DP, 6°C; QNH, 1012
Scenario 2	High Anchor	Wind at 030°T at 3 kt; visibility is 24 km; cloud broken at 3700 ft; temp, 22°C; DP, 16°C; QNH, 1002
	Low Anchor	Wind at 030°T at 3 kt; visibility is 8 km; cloud broken at 1300 ft; temp, 22°C; DP, 16°C; QNH, 1002
Scenario 3	High Anchor	Wind at 080°T at 8 kt; visibility is 26 km; cloud broken at 3900 ft; temp, 19°C; DP, 6°C; QNH, 1004
	Low Anchor	Wind at 080°T at 8 kt; visibility is 6 km; cloud broken at 1100 ft; temp, 19°C; DP, 6°C; QNH, 1004
Scenario 4	High Anchor	Wind at 180°T at 6 kt; visibility is 25 km; cloud broken at 3800 ft; temp, 12°C; DP, 6°C; QNH, 1001
	Low Anchor	Wind at 180°T at 6 kt; visibility is 7 km; cloud broken at 1200 ft; temp, 12°C; DP, 6°C; QNH, 1001

temp, temperature; DP, dew point; QNH, sea level pressure (hPa)

7.2.3 Procedure

Participants completed the task at the end of a shorter than normal lecture during a 7-month period between September 2014 and March 2015. The task was conducted in groups, with approximately half the groups receiving the debiasing intervention, then completing the questionnaire immediately afterwards. The other half of the participants did not receive any debiasing intervention (the control group) and only received the questionnaire material. On each occasion the task was completed, the group size varied from 7 to 20 participants, with a randomisation program on Microsoft Excel being used to determine which type of intervention each group received (debiased vs. control). No time limit was imposed and participants were asked not to confer during participation or discuss the study outside of class.

The questionnaire distribution method was slightly different from that used for Study 1. Study 1 was conducted online; Study 4 was paper-based. However, the design of the questionnaire followed a similar process to Study 1. First, participants were provided with a brief introduction to the questionnaire and informed that the study had undergone an ethical review. A review of the ethics of this study was undertaken using the guidelines provided by the Massey University Ethics Committee. The project was deemed to require a full ethical review as the participants were students of the researcher. The Massey University Human Ethics Committee reviewed and approved the study, with the requirement of confidentiality. A copy of the ethical approval can be found in Appendix H.

Next, the participants evaluated the four scenarios. The order of the scenarios was randomly assigned using a randomised program on Microsoft Excel for each participant. Furthermore, as the participants completed the tasks for all three debiasing studies during the same session, it was also important to randomise the order in which they completed each task. It is important to note that participants completed all the tasks for a single study before moving on to the next. For example, if a participant was randomly selected to complete Study 5 first, they would first complete all the tasks for Study 5 before moving onto the next study.

A priori power analysis using the software G*Power (Erdfelder et al., 1996) was used to determine that, with $\alpha = .05$, a total sample size of $n = 68$ would be sufficient for an experimental power of .80 (assuming an effect size of .25) and a correlation between within-subjects measurements of $r = .3$, using a repeated-measures within-between interaction ANOVA. It was determined that at least a medium effect size would be a meaningful starting point in the design of any intervention.

7.3 Results

7.3.1 General Findings

The data were initially screened to identify any outliers and to establish that the data met the requirements for the application of parametric statistical analysis. The skewness and kurtosis of the dependent variable were within the limits recommended by Cameron (2004). This screening identified two outliers, the outliers were defined as z-scores greater than +3.5 (Iglewicz & Hoaglin, 1993). One outlier was in the

cloud height assessment debiased group and one was in the visibility assessment control group. Further investigation ruled out data entry errors. It is not unreasonable to expect outliers, especially at the higher end of weather assessments, where cloud height and visibility can be considerable. For this reason, it was decided to retain the outlier data¹⁴.

The level of statistical significance was set at $\alpha = .05$ for all statistical tests and all tests were conducted as two-tailed. Participants were exposed to a total of four scenarios, two high anchor scenarios and two low anchor scenarios. For each of the anchor conditions, the mean cloud height and visibility assessment was calculated.

7.3.2 Cloud Height Assessment

A mixed model ANOVA was used to compare the pilots' assessments of cloud height after being exposed to an anchor (high vs. low) and their intervention group (debiased vs. control). There was no evidence of a main effect for intervention group (debiased vs. control), $F(1,100) = 0.19$, $p = .67$, $\eta^2 = .001$ or an interaction between the anchor and the intervention, $F(1,100) = 0.25$, $p = .62$, $\eta^2 = .002$. There was a significant main effect for cloud assessment after exposure to a high or low anchor, $F(1,100) = 76.95$, $p < .001$, $\eta^2 = .43$, with pilots reporting a higher assessment of cloud height after exposure to the high anchor ($M = 2374$ ft, $SD = 814$), compared to when they were exposed to the low anchor ($M = 1602$ ft, $SD = 572$). Table 14 provides the means and SDs of the anchors (high vs. low) and intervention groups (debiased vs. control) for the cloud height.

¹⁴ Similar to Study 1, the data was also analysed without the outliers which confirmed the outliers did not have an impact on the findings

Table 14: *Cloud Height assessment after exposure to an anchor (high vs. low) and intervention group (debiased vs. control).*

	Low Anchor	High Anchor
Debiased	1600 ft ($SD = 619$)	2326 ft ($SD = 788$)
Control	1604 ft ($SD = 533$)	2417 ft ($SD = 842$)

7.3.3 Visibility Assessment

A mixed model ANOVA was used to compare the pilots' assessments of visibility after being exposed to an anchor (high vs. low) and their intervention group (debiased vs. control). There was no evidence of a main effect of intervention groups (debiased vs. control), $F(1,100) = 0.96$, $p = .33$, $\eta^2 = .010$, or an interaction between the anchor and intervention, $F(1,100) = 0.25$, $p = .62$, $\eta^2 = .002$. There was a significant main effect for visibility assessment after exposure to a high or low anchor, $F(1,100) = 52.15$, $p < .001$, $\eta^2 = .34$, with pilots reporting a higher assessment of visibility after exposure to the high anchor ($M = 14.37$ km, $SD = 7.52$), compared to when they were exposed to the low anchor ($M = 10.48$ km, $SD = 5.62$). Table 15 provides the means and SDs of the anchors (high vs. low) and intervention groups (debiased vs. control) for the visibility assessment.

Table 15: *Visibility assessment after exposure to an anchor (high vs. low) and intervention group (debiased vs. control).*

	Low Anchor	High Anchor
Debiased	11.26 km ($SD = 5.21$)	14.86 km ($SD = 7.41$)
Control	9.81 km ($SD = 5.97$)	13.94 km ($SD = 7.66$)

7.3.4 Comparison to Study 1

The findings of the current study were compared to those of Study 1 to explore if the anchoring effect found in the original study was replicated in the control group of this study. A 2 (anchor) \times 2 (Study 1 vs. Study 4 cloud assessment) ANOVA was not significant, $F(1,243) = 1.33$, $p = .25$, $\eta^2 = .005$. A 2 (anchor) \times 2 (Study 1 vs. Study 4 visibility assessment) ANOVA was not significant, $F(1,242) = 2.27$, $p = .13$, $\eta^2 = .010$. These results suggest that there is no significant difference in the two studies and that therefore, the anchoring effect observed in Study 1 was replicated within the control (non-intervention) group of this study.

It must be noted at least three differences between the two studies existed: the delivery method (paper vs. online), the sample size between the two studies (Study 4, $n = 54$ ¹⁵; Study 1, $n = 201$) and the characteristics of participants in each study (Study 1 = wide spectrum of pilots, Study 4 = young and inexperienced pilots).

¹⁵ Note that for the purposes of comparison with Study 1, only participants within the control of Study 4 were ($n = 54$).

7.4 Discussion

The findings presented here suggest that the debiasing technique used in this study was unsuccessful in reducing the phenomenon of the anchoring effect in weather-related decision making. The findings do not support Hypotheses 1 and 2 for this study. In other words, encouraging pilots to consider additional information and thus not focus on a narrow set of information when assessing weather conditions by ‘considering the alternative’ had little impact on reducing the anchoring effect during cloud height assessment and visibility assessment.

A comparison of the pilots’ assessments of cloud height and visibility distance for both experimental groups (debiased and control group) found that when pilots were exposed to a high anchor (a forecast of high cloud and good visibility), both groups assessed the cloud height and visibility as being greater than when they were exposed to a low anchor (forecast low cloud and poor visibility). Even though pilots were exposed to the same in-flight weather conditions, their judgment of the weather conditions were influenced by the initial weather report, which acted as a high or low anchor, thus supporting the findings of Study 1.

Why this particular debiasing intervention was unsuccessful may be due to the challenge of achieving one of the criteria in the cognitive debiasing model suggested by Croskerry et al. (2013). Step (iii) of this model, ‘awareness of the direction and magnitude of bias’ can be challenging to achieve in real-world scenarios when an individual is confronted with an anchor. When presented with the weather forecast, which acted as the anchor in this study, without explicitly informing participants

whether their forecast was the high or low anchor (which would not be known in real life), each individual had to judge which way to apply the debiasing technique. In situations where pilots encounter an anchor that is considerably different from reality, this judgment might be easier to achieve. In the marginal weather conditions presented in this study—which are more likely to be encountered in VFR flight into IMC events—this judgment is much more difficult.

Studies in which the anchoring effect has been reduced using the ‘consider the alternative’ technique tended to provide an indication of what type of anchor (i.e., high or low) the participants were dealing with. For example, Mussweiler et al. (2000) asked participants to list ‘reasons why a *high* price is inadequate’ and therefore participants had a better understanding of which direction to apply the debiasing technique. In the study reported here, participants were not provided with any indication of the type of anchor they had been exposed to. The method adopted in this study better reflects how real-world decision making, particularly weather-related decision making.

The finding that debiasing of the anchor effect was unsuccessful in this study would suggest that other techniques to avoid the anchoring effect and aid pilots to assess the weather conditions accurately need to be explored. Wickens et al. (2013) suggested that to address cognitive biases a range of measures may be required. These measures will be discussed in detail in the General Discussion (Chapter 10). One of the measures that could be used to aid pilots to avoid the anchoring effect is ‘displays and information’. Anchoring, as presented in this study (and Study 1), is caused by an individual relying on outdated information. When a pilot is operating in a

dynamic weather environment, especially in the likes of convective or frontal situations, a weather forecast a few hours old can quickly become obsolete. Forecasts are regularly updated or amended, but it is often necessary for the pilot to seek out these amendments, especially when operating in uncontrolled airspace where he or she has little contact with an air traffic controller. Providing simple, easy methods to ensure that pilots remain up to date with the latest information may limit the anchoring effect of outdated weather forecasts.

The limitations of this study are similar to those identified in the previous studies. This study used static images, focusing on similar weather scenarios and therefore, care should be taken in generalising these results to a range of situations. Caution also needs to be applied when generalising these results to the GA pilot population. This study sampled pilots with limited flight experience at a New Zealand-based international flight training organisation.

In conclusion, the findings of Study 4 suggest that using the ‘consider the alternative’ debiasing technique had little effect on reducing the anchoring effect in weather-related decision making in this group of pilots.

CHAPTER EIGHT

STUDY 5

DEBIASING CONFIRMATION BIAS IN WEATHER-RELATED DECISION MAKING

8.1 Introduction

The aim of Study 5 was to explore debiasing of confirmation bias in weather-related decision making. If the debiasing was successful, pilots would have a reduced tendency to favour confirmatory evidence when making weather-related decisions compared to when no debiasing was present, such as in Study 2. Humans have a natural tendency to look for evidence that supports their own hypotheses (Nickerson, 1998); however, this method of seeking out confirmatory evidence can lead to errors of judgment. There is nothing inherently wrong in using confirmatory evidence (a positive testing strategy) but an over-reliance on the positive testing strategy can lead to serious consequences. For example, if a pilot falsely judges the weather conditions as being safer than they actually are, this may result in the pilot continuing a flight towards IMC.

Study 2 found that when assessing whether to continue a VFR flight, pilots have a strong tendency to use a positive testing strategy by seeking confirmatory evidence. Exploring methods to reduce the rate of confirmation bias may assist pilots to assess the weather conditions more accurately. The ‘consider the alternative’ debiasing

technique was chosen as this technique has been shown to mitigate the effects of a range of biases, including confirmation bias (Larrick, 2004).

Debiasing has received mixed results in a range of studies (Fischhoff, 2002) and this has been particularly the case with confirmation bias (Larrick, 2004; Teasley, Leventhal, Mynatt, & Rohlman, 1994). Therefore, consistent with a range of debiasing studies using the ‘consider the alternative’ technique, the following hypothesis was derived:

Hypothesis 1: Confirmation bias in weather-related decision making will be reduced after the debiasing technique ‘consider the alternative’. Specifically, pilots will have a reduced tendency to use the positive testing strategy of seeking confirmatory evidence but will seek for disconfirmatory evidence instead.

8.2 Method

8.2.1 Participants

The participants for this study were the same participants who completed Study 4; 101 pilots who were all enrolled in a flight training programme in which they learn to fly and complete an academic programme of study.

8.2.2 Material and Design

8.2.2.1 Overview

The materials and design for Study 5 were similar to those for Study 4. Participants were assigned to one of two experimental groups. One group was exposed to the debiasing lecture, whereas the other had no debiasing intervention (the control group). Both groups then received the same questionnaire, which evaluated the evidence participants used when assessing weather conditions. The benefits and shortcomings of using a questionnaire-based methodology have already been discussed in detail in Chapter 3. As outlined in Study 5, the data for all three debiasing studies (Studies 4–6) were collected at the same time.

8.2.2.2 Development of the Debiasing Intervention

This study used the same cognitive debiasing model as Study 4 and was delivered at the same time (Croskerry et al., 2013). The short lecture contained the following components:

- i) *Awareness*: What is confirmation bias and why does it happen?
- ii) *Motivation*: Examples of how confirmation bias may influence weather assessment and therefore may lead to VFR flight into IMC accidents.
- iii) *Awareness of the direction and magnitude of the bias*: Examples of how to recognise confirmation bias and the magnitude of the bias.
- iv) *Debiasing technique*: An introduction to the ‘consider the alternative’ debiasing technique, with examples of how to use this technique to reduce confirmation bias.

After the short lecture, participants in the debiasing group received the same questionnaire material as the control group. The lecture slides used in Study 5 can be found in Appendix I.

8.2.2.3 Development of Scenarios

The questionnaire material was based on Study 2. The participants received a total of three scenarios rather than five. The number of scenarios was reduced to improve participation rate by reducing survey fatigue. The data from tasks for Study 4 and 6 tasks were also collected at the same time, which resulted in participants receiving a total of 10 scenarios in the combined data collection. The three scenarios were retained for Study 5 based on the internal consistency (Cronbach's alpha) of Study 2. Cronbach's alpha for the original five scenarios was $\alpha = 0.61$, which was reduced to $\alpha = 0.57$ after the removal of two scenarios. Scenario 2 and Scenario 5 had the lowest internal consistency scores ($\alpha = 0.31$ and $\alpha = 0.35$ respectively) and were therefore removed.

Each of the scenarios used the same static images as in Study 2. The criteria for the images have been outlined with Study 2; the key difference was the method by which images were presented. The problem of preserving image quality in paper-based surveys compared to online surveys have already been discussed Study 4. Once again, to ensure that the high quality of the images was retained, each image was printed on high-quality photo print paper before being attached to the paper-based questionnaire.

The remainder of the study was consistent with Study 2. For full details of the scenario design refer to Study 2. The following is a brief summary. Participants were asked to imagine that they were the PIC flying a Cessna 172 on a VFR flight approaching their destination. Participants were provided with an in-flight static image, looking through the cockpit window. Participants were instructed to choose one of three pieces of evidence that they could see to determine if VFR flight could be maintained in their direction of travel. Two pieces of evidence supported the hypothesis that safe flight was possible. For example, the cloud base ahead was high enough. These were designated as the confirmatory choices, which suggest that a positive strategy was being used to test the hypothesis that VFR flight was possible. A third piece of evidence did not support VFR flight. For example, the visibility was too low. This piece of evidence was designated as the disconfirmatory choice or negative testing of the hypothesis that VFR flight was possible.

As in Study 2, in each of the three scenarios, the weather conditions were such that continuing the flight could result in the VFR meteorological minima no longer being met. As a result, inappropriate use of the positive testing strategy may put the pilot at risk of VFR flight into IMC and was interpreted as evidence of confirmation bias.

One of the scenarios follows as an illustration:

Imagine you are flying a Cessna 172 on a VFR cross-country. Halfway through the flight, the aircraft engine begins to make a strange noise. You decide it would be wise to divert to a nearby aerodrome to get it checked out. The aerodrome is just beyond the red box.



Listed below are three statements about the weather conditions in front of the aircraft. Pick one that you feel is most useful in deciding if you can continue towards this aerodrome?

1. The cloud base is too close to the ground.
2. There is a clear patch towards the aerodrome I am heading towards.
3. The heavy showers are remaining to the right of my track.

In this particular scenario, although Items 2 and 3 (confirmatory choices) are evidence which could lead a pilot to decide that he or she can continue towards the

aerodrome, these items could also occur simultaneously in situations when VFR flight is not possible. The forward visibility might be fine (‘There is a clear patch towards the aerodrome I am heading towards.’) but the cloud base might be too low. The heavy showers may currently be to the right of the track, but this could easily change, depending on the wind speed and direction, as well as other weather-related factors. Item 1 (the negative test), if correct, would be enough by itself to disprove the assessment that it is safe to continue the flight. Table 16 presents the evidence that participants received for each scenario. The other two scenarios (three in total) had the same format as this scenario. The questionnaire may be inspected in Appendix G.

Table 16: *Evidence used in each scenario in Study 5*

Scenario No.	Confirmatory Evidence	Disconfirmatory evidence
1	<ul style="list-style-type: none"> • I can see through the rain shower ahead. • Another aircraft ahead safely flew in my direction 10 min earlier. 	<ul style="list-style-type: none"> • I will almost certainly end up having to fly through cloud.
2	<ul style="list-style-type: none"> • Another aircraft ahead has reported that the cloud height is above circuit altitude. • I can see the horizon through the rain shower. 	<ul style="list-style-type: none"> • The cloud base inside the rain shower may be lower than my current flight level.
3	<ul style="list-style-type: none"> • There is a clear patch towards the aerodrome I am heading towards. • The heavy showers are remaining to the right of my track. 	<ul style="list-style-type: none"> • The cloud base is too close to the ground.

8.2.3 Procedure

The procedure for this current study was the same as Study 4 and was completed at the same time. That is, the task was conducted in groups, with approximately half the groups receiving the debiasing intervention then completing the questionnaire immediately afterwards. The other half of participants did not receive any debiasing intervention (the control group) and only received the questionnaire material.

The questionnaire distribution was paper-based; with the design of the questionnaire being similar to that of Study 2. First participants were provided with a brief introduction to the questionnaire and informed that the study had undergone an ethical review. A copy of the ethical approval can be found in Appendix H. Next, the participants evaluated the three scenarios. The order of the scenarios was randomly assigned using a randomised program on Microsoft Excel for each participant. Furthermore, as participants completed the tasks for all three studies during the same session, it was important to also randomise the order in which they completed the tasks for each study.

A priori power analysis using the software G*Power (Erdfelder et al., 1996) was used to determine that with $\alpha = .05$, a total sample size of $n = 64$ would be sufficient for an experimental power of .80 (two-tailed), using an effect size of $d = .5$ for an independent sample *t*-test. It was determined that at least medium effect size would provide a meaningful starting point in the design of any intervention.

8.3 Results

The data were initially screened to identify any outliers and to establish that the data met the requirements for the application of parametric statistical analysis. No outliers were identified, and the skewness and kurtosis of the dependent variable were within the limits recommended by Cameron (2004). The level of statistical significance was set at $\alpha = .05$ for all statistical tests and all tests were conducted as two-tailed.

An independent-sample *t*-test was conducted to compare the mean number of disconfirmatory choices between the two groups: debiased ($n = 47$) and control pilots ($n = 54$). Levene's test for equality of variances was not significant, confirming that the data met the assumptions underlying the independent sample *t*-test. No significant difference between the two groups was found, $t(99) = 0.58$, $p = .56$, $d = .12$. Over half of participants in the debiased group (64%) and the control group (61%) selected the disconfirming choice less than half the time. Only seven (15%) of the debiased and five (9%) of the control group made the disconfirming choice in all three scenarios. The mean number of disconfirming choices across all three scenarios was 1.06 ($SD = 1.13$) for the debiased group and 1.19 ($SD = 0.97$) for the control group.

A single-sample *t*-test (two-tailed) using the total number of disconfirming choices for each participant as data and a test value of 1 (the mean number of disconfirming choices that would be expected if participants selected their responses by chance alone) was conducted for each group. There was no evidence of a difference for the debiased group, $t(46) = 0.39$, $p = .70$ or the control group, $t(53) = 1.40$, $p = .17$. This

indicates that participants in both groups performed no differently from what would be expected by chance. That is, there was no evidence that participants in either group were seeking the disconfirming choice when testing their hypothesis. Participants had an overwhelming preference for using confirmatory evidence when assessing the weather conditions in both the debiased and control group.

Comparison with Study 2

A direct comparison with Study 2 is not possible, as the studies used a different number of scenarios (this study used three scenarios; Study 2 used five). However, an indirect comparison can be made, as the primary analysis for both studies was to compare the number of disconfirmatory choices made to what would be expected by chance. In both studies, the single-sample *t*-test was used to determine if there was any difference from what would be expected by chance. For the number of scenarios presented in both studies, participants selected the disconfirming choice no differently from what would be expected by chance. These results suggest that the results in Study 2 were replicated within the control (non-intervention) group of this study.

8.4 Discussion

The findings presented here suggest that the debiasing intervention used in this study was unsuccessful in reducing the phenomenon of confirmation bias in weather-related decision making, thus not supporting Hypothesis 1. In other words, encouraging pilots to ‘consider the alternative’ when assessing weather conditions had little impact on reducing the rate at which pilots sought confirmatory evidence.

In both groups, participants had an overwhelming preference to seek confirmatory evidence when assessing the weather conditions. In each scenario, this strategy could have provided a false positive. The findings were consistent with Study 4 and with other confirmation debiasing studies (e.g., Gilbey & Hill, 2012; Larrick, 2004). These findings highlighted the major challenge faced in helping pilots avoid being lured into a false perception of the weather conditions by confirmation bias. The debiasing technique used in this study was consistent with the method recommended to reduce confirmation bias, ‘consider the alternative’, which, in the case of confirmation bias, required participants to consider why their hypothesis might be wrong. One reason why the debiasing intervention was unsuccessful in this study may be because the participants did not have enough time to practice the debiasing technique. The debiasing method used was brief and passive, and participants did not have the opportunity to practice or receive feedback on their performance. Therefore, although the debiasing was largely ineffective, it would be unwise to rule out the possibility that more extensive training could improve performance. Extensive training has been shown to be effective in reducing some biases (Cheng, Holyoak, Nisbett, & Oliver, 1986; Larrick et al., 1990). This training could potentially be incorporated into human factors training of *ab initio* pilots.

This study confirms the strong tendency to seek confirmatory information when testing pre-held ideas about a hypothesis. Under normal circumstances, in benign everyday situations, seeking out hypothesis-confirming evidence is quite reasonable, especially when the consequences of a wrong decision are minimal (Croskerry, 2014). Considering that this study was largely ineffective in reducing participants’ tendency to seek confirmatory evidence, and how Study 2 found that experience

levels had little impact on reducing confirmation bias, it would suggest that a wider approach to reduce the potential impact of confirmation bias needs to be explored. As outlined in the discussion in the previous chapter regarding debiasing the anchoring effect, development of displays and clearer information may be one area that could be explored. A richer flow of information could potentially support more effective decision making (Wickens et al., 2013).

One example of how improved information may reduce confirmation bias was highlighted by the NTSB Accident Report (NTSB, 2008) into the fatal Comair Flight 5191 crash in 2005. On 27 August 2005, a CRJ-100 (a 50-seater regional jet) inadvertently lined up on the wrong (shorter) runway at Lexington, Kentucky. Among other factors, confirmation bias was cited as one factor why the crew did not notice they had lined up and attempted to take off from the shorter runway. The aircraft failed to lift off, crashing at the end of the runway, killing 49 of the 50 people on board. The NTSB suggested that clearer information could have helped alert the crew that they had lined up on the wrong runway. The aircraft was located on a different part of the airport than were the crew thought they were located; however, a check of the aircraft heading may have provided a potential cue (but was not an item of the flight crew's check-list). The NTSB suggested a moving map display of the airport may have helped provide the crew with a strong enough cue to realise they had lined up on the wrong runway. In the case of VFR flight into IMC, improved information flow may help weather-related decision making. This will be discussed in more detail in the General Discussion (Chapter 10), which also highlights a number of challenges to this approach.

The limitations of this study were similar to those reported in Study 4. This study used static images, focusing on a similar weather scenario, so therefore care should be taken in generalising these results to a range of differing weather conditions. Caution also needs to be applied when generalising these results to the GA pilot population. This study surveyed pilots with limited flight experience at a New Zealand-based international flight training organisation.

In conclusion, the findings of the current study suggest that using the ‘consider the alternative’ debiasing technique had little effect in reducing confirmation bias in weather-related decision making.

CHAPTER NINE

STUDY 6

DEBIASING OUTCOME BIAS IN WEATHER-RELATED DECISION MAKING

9.1 Introduction

The aim of Study 6 was to explore debiasing of outcome bias in weather-related decision making. If the debiasing was successful, pilots would have a reduced tendency to judge a decision based on its eventual outcome, focusing instead on the quality of the decision process.

Study 3 found that outcome bias influences how pilots judge another pilot's decision to conduct a VFR flight and may also influence a pilot's own decision to conduct a VFR flight. Outcome bias refers to the tendency to judge the quality of a decision by its eventual outcome instead of judging it based on the quality of the decision at the time it was made (Baron & Hershey, 1988). It has consistently been found that people tend to overlook bad decisions where the outcomes have been relatively good (e.g., near misses or situations which went awry but turned out well in the end) but focus on decisions where the outcomes have been bad (e.g., in fatal accidents) (Tinsley et al., 2011).

This bias may mean that lessons may be lost to all except those directly responsible (Dillman et al., 2007). What seems to be overlooked is that it may be just luck or good fortune that decides whether the outcome of a decision is good or bad and a slight change in the conditions that made a pilot ‘lucky’ once may result in a disaster if that luck changes. (Tinsley et al., 2011, p. 90). Exploring methods to reduce the impact of outcome bias may help pilots to accurately assess whether it is safe to conduct a VFR flight. The ‘consider the alternative’ debiasing technique was chosen, as this technique has been shown to mitigate the effects of a range of biases, including outcome bias (Larrick, 2004).

Clarkson et al. (2002) observed some success in debiasing outcome bias using a similar method to that used in this study. On the other hand, a range of studies have also found debiasing to be less successful (Fischhoff, 2002). Therefore, consistent with a range of debiasing studies using the ‘consider the alternative’ technique, the next hypotheses are proposed:

Hypothesis 1: Debiasing will reduce the rate at which outcome information influences how participants rate the decision maker’s judgment.

Hypothesis 2: Debiasing will reduce the rate at which outcome information influences how participants rate the risk undertaken.

Hypothesis 3: Debiasing will reduce the rate at which outcome information influences a participant’s decision to conduct a VFR flight.

9.2 Method

9.2.1 Participants

The participants for this study were the same participants that completed Study 4 and Study 5: 101 pilots who were all enrolled in a flight training programme in which they learn to fly and complete an academic programme of study.

9.2.2 Materials and Design

9.2.2.1 Overview

The materials and design for Study 6 were similar to those of Study 4 and 5. Participants were assigned to one of two experimental groups. One group was exposed to the debiasing lecture while the other had no debiasing intervention (the control group). Both groups then received the same questionnaire, where participants were asked to rate a series of scenarios on a number of relevant dimensions.. The benefits and shortcomings of using a questionnaire-based methodology have already been discussed in detail in Chapter 3.

9.2.2.2 Development of Debiasing Intervention

This study used the same cognitive debiasing model as Studies 4 and 5, and was delivered at the same time (Croskerry et al., 2013). The short lecture contained the following components:

- i) *Awareness*: What is outcome bias and why does it happen?

- ii) *Motivation*: Examples of how outcome bias may influence weather assessments and may therefore lead to VFR flight into IMC.
- iii) *Awareness of the direction and magnitude of the bias*: Examples of recognising outcome bias and the magnitude of the bias.
- iv) *Debiasing technique*: An introduction to the ‘consider the alternative’ debiasing technique, providing an example of how to use this technique to reduce outcome bias.

After the short lecture, participants in the debiasing group received the same questionnaire material as the control group. The lecture slides used in Study 6 can be found in Appendix I.

9.2.2.3 Development of Scenarios

The questionnaire material was based on Study 3 but the participants received a total of three scenarios rather than five. The number of scenarios was reduced to improve participation rate by reducing survey fatigue. The data from tasks for Study 4 and 5 were also collected at the same time, which resulted in participants receiving a total of 10 scenarios in the combined data collection. The three scenarios were retained for Study 6 based on the internal consistency (Cronbach’s alpha) of Study 3. Cronbach’s alpha for the original five scenarios was $\alpha = 0.72$ which was reduced to $\alpha = 0.70$ after the removal of two scenarios. Scenario 2 and Scenario 5 had the lowest internal consistency scores ($\alpha = 0.52$ and $\alpha = 0.30$ respectively) and were therefore removed. These internal consistencies are for the ‘decision quality’ dependent variable; however, the other two main dependent variables (risk assessment and safety assessment) correlated with these scores.

Similar to Study 3, each scenario contained all the information a pilot would require to conduct the flight (weather information, route map and general information about the time of day and the length of the flight). The design criteria used in these scenarios were discussed in Study 3. One of the scenarios reads:

A pilot is conducting a VFR cross-county flight in a Cessna 172. The pilot has flown the same route several times before. The flight should take about an hour, conducted in daylight hours. The weather information available to the pilot for the flight is as follows (a route map was also included in each scenario);

Current weather

Departure (A): wind variable, 3 knots; visibility, 16 km; light rain with a cloud base at 1800 ft.

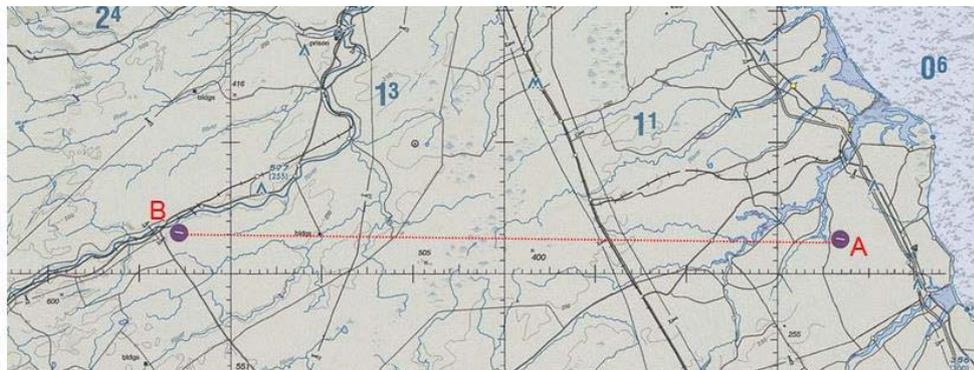
Destination (B): wind variable, 2 knots; visibility, 16 km; light drizzle with a cloud base at 1500 ft.

Area Forecast

Visibility: 16 km with scattered light drizzle.

Cloud broken to overcast at 1,500 ft, with cloud tops at 24,000ft.

Turbulence: nil.



After reviewing the information, participants were randomly provided with information about the eventual outcome of the flight. The type of outcome

information (i.e., negative, positive or no outcome) was the same for all three scenarios once the participants had been assigned to their condition.

Participants in the positive outcome condition group read the following ending to all the scenarios:

‘You later find out that the flight landed safely at its intended destination.’

Participants in the negative outcome condition group read the endings shown in Table 17.

Table 17: *Negative outcome statements for each scenarios in Study 6*

Scenario 1	<i>En route</i> , the weather deteriorated, the pilot decided to conduct a precautionary landing (an off-field landing). On landing in a paddock, the aircraft encountered a soft area, resulting in the nose gear breaking. The pilot escaped injury but the aircraft sustained substantial damage.
Scenario 2	The flight inadvertently flew into cloud, resulting in the pilot losing control of the aircraft and crashing. The pilot received serious injuries.
Scenario 3	The pilot contacted air traffic services and advised them that he/she had inadvertently flown into IMC conditions. The air traffic service was able to identify the aircraft on radar and provide navigational assistance, allowing the pilot to regain visual flight and return to the departure aerodrome.

Participants in the no outcome information group (the control group) read only the common information and no outcome information was provided.

9.2.2.4 Dependent Measures

After reading the scenario, participants were instructed to rate each of the following dimensions using a nine-point Likert-type scale:

- Decision quality ('How do you rate the pilot's decision to conduct the VFR flight?'), where the a scale ranged from 1 = *very good decision* to 9 = *very poor decision*;
- Risk rating ('How much risk do you think the pilot took in attempting this VFR flight?') , where the scale ranged from 1 = *very low risk* and 9 = *very high risk*;
- Whether the respondent would conduct the same flight ('You wish to conduct the same flight. Considering the information available to the pilot before he/she departed, indicate your level of agreement with the following statement: I would be able to safely conduct the flight VFR. '), where the scale ranged from 1 = *strongly agree* to 9 = *strongly disagree*.

The questionnaire used in Study 6 can be found in Appendix G.

9.2.3 Procedure

The procedure for this current study was the same as that for Studies 4 and 5, and was completed at the same time: the task was conducted in groups, with approximately half the groups receiving the debiasing intervention then completing the questionnaire immediately afterwards; the other half of the participants did not receive any debiasing intervention (the control group) and only received the questionnaire material.

The questionnaire distribution was paper-based, with the design of the questionnaire being similar to that of Study 3. First, participants were provided with a brief introduction to the questionnaire and informed that the study had undergone an ethical review. A copy of the ethical approval can be found in Appendix H. Next, the participants evaluated the three scenarios. The order of the scenarios was randomly assigned using a randomised program on Microsoft Excel for each participant, which also randomly assigned participants to one of three outcome conditions: positive, negative and no outcome/control (the outcome condition remained the same for each participant across the three scenarios). Furthermore, because participants completed the tasks for all three studies during the same session, it was important to randomise the order in which they completed each study. This was achieved using a randomised program on Microsoft Excel. Due to ethical requirements no demographic data were collected.

A priori power analysis using the software G*Power (Erdfelder et al., 1996) was used to determine that with $\alpha = .05$, a total sample size of $n = 90$ (split evenly among cells) would be sufficient for experimental power of at least .80, assuming a mean effect size of $f = .4$ for a between-subjects ANOVA.

9.3 Results

9.3.1 General results

The data were initially screened to identify any outliers and to establish that the data met the requirements for the application of parametric statistical analysis. No outliers were identified, and the skewness and kurtosis of the dependent variables were within the limits recommended by Cameron (2004). The level of statistical significance was set at $\alpha = .05$ for all statistical tests and all tests were conducted as two-tailed.

To explore the influence of the debiasing intervention, a 2 (intervention) \times 3 (outcome) ANOVA was conducted. It is important to note the primary interest in these analyses is the results for the *interaction*. Evidence of a significant interaction would suggest a difference exists between the two intervention groups (debiased vs. control).

9.3.2 Decision Judgment

A 2 (debiased vs. control) \times 3 (outcome condition) ANOVA revealed no evidence of a main effect among the three outcome conditions for the perceived quality of the decision, $F(2,95) = 1.82, p = .17, \eta^2 = .035$. There was no evidence of a main effect between intervention groups (debiased vs. control), $F(1,95) = 1.39, p = .24, \eta^2 = .014$, and no evidence of an interaction between the outcome and intervention groups $F(2,95) = 1.63, p = .20, \eta^2 = .032$.

9.3.3 Risk Assessment

A 2 (debiased vs. control) \times 3 (outcome condition) ANOVA revealed no evidence of a main effect among the three outcome conditions for risk assessment of the flight, $F(2,95) = 1.94, p = .15, \eta^2 = .038$. There was no evidence of a main effect between intervention groups (debiased vs. control), $F(1,95) = 1.51, p = .22, \eta^2 = .015$, and no evidence of an interaction between the outcome and intervention groups, $F(2,95) = 0.70, p = .50, \eta^2 = .014$.

9.3.4 Would The Respondent Fly?

A 2 (debiased vs. control) \times 3 (outcome condition) ANOVA revealed no evidence of a main effect among outcome groups regarding the decision to conduct the same VFR flight, $F(2,95) = 1.82, p = .17, \eta^2 = .036$. There was no evidence of a main effect between intervention groups (debiased vs. control), $F(1,95) = 2.61, p = .11, \eta^2 = .026$, and no evidence of an interaction between the outcome and intervention groups, $F(2,95) = 0.51, p = .60, \eta^2 = .010$.

9.3.5 Comparison with Study 3

The findings of the current study were compared to those of Study 3 to explore if the outcome bias found in the original study was replicated in the control group of this study. A 3 (outcome) \times 2 (Study 3 v Study 6) ANOVA was conducted for each of the three dependent variables. There was no significant interaction for any of the three dependent variables. Judgment of the flight, $F(2,270) = 0.12, p = .89, \eta^2 = .001$; risk assessment of the flight, $F(2,270) = 0.073, p = .93, \eta^2 = .001$; safety assessment to

conduct the flight, $F(2,270) = 0.24$, $p = .78$, $\eta^2 = .002$. These findings suggest that the outcome bias observed in Study 3 was also observed in the control group of this study.

It should be noted that at least three differences between the two studies existed: the delivery method (paper vs. online), the sample size between the two studies (Study 6, $n = 54$; Study 3, $n = 300$) and the characteristics of the participants in each study (Study 3 = a wide spectrum of pilots; Study 6 = young and inexperienced pilots).

9.4 Discussion

The findings presented here suggest that the debiasing intervention used in this study was unsuccessful in reducing the phenomenon of outcome bias in weather-related decision making. These findings do not support Hypotheses 1, 2 and 3 for this study. In other words, encouraging pilots to consider why the outcome might have been different when judging another flight, assessing the risk of that flight or assessing whether or not they would have conducted the flight themselves had little impact on reducing the influences of outcome information on these judgments.

A comparison of the participants' assessments of the quality of the decision to conduct a VFR flight suggests that after receiving the debiasing intervention, participants had the same tendency to rate a decision based on the outcome. This was observed in both the positive and negative outcome conditions. The findings for risk assessment and the pilot's own willingness to conduct the flight also showed the same tendency to use outcome information even after receiving the debiasing

intervention. The main effect in each of the three dependent variables suggested no evidence of outcome bias; however, the analysis of the control group confirmed outcome bias was observed, similar to Study 3. One explanation for this finding is that a few participants may have responded to the debiasing intervention, while the majority did not. These findings suggest that there may be some individual differences in the impact of debiasing. Unfortunately due to the lack of demographic data collected in this study this could not be explored further.

Similarly to Study 5, the reason why the debiasing intervention was unsuccessful may be partly due to the debiasing intervention being delivered too passively. Clarkson et al. (2002) successfully debiased outcome bias but this was achieved using a strong debiasing intervention (for example, encouraging participants to reflect more deeply on the evidence). Interventions which only advised participants on potential biases did not have a statistically significant effect. The findings of this present study are surprising, considering that the participants would easily have been able to determine which way to apply the debiasing model. Apart from the control group, participants could clearly see which direction they were required to apply the ‘consider the alternative’ technique.

Out of the three cognitive biases explored in this thesis, the debiasing model was expected to be able to be applied to outcome bias with minimal practice and feedback. However, the findings of this study suggest that stronger debiasing interventions would be necessary, which may produce more encouraging results. A detailed discussion of how to support pilots make accurate weather-related decisions are discussed in the next chapter (General Discussion).

The limitations of this study were similar to those of the previous studies. This study focused on similar weather scenarios, so care should be taken in generalising these results to a range of weather situations. Caution also needs to be applied when generalising these results to the GA pilot population. This study looked at pilots with limited flight experience at a New Zealand-based international flight training organisation.

In conclusion, the findings of the current study suggest that using the ‘consider the alternative’ debiasing technique had little effect in reducing outcome bias in weather-related decision making in this study group.

CHAPTER TEN

GENERAL DISCUSSION

10.1 Chapter Overview

This chapter discusses the implications and the theoretical contribution of the studies that encompass this thesis. The research questions are first addressed, followed by a discussion of the implications of cognitive biases and debiasing on pilot weather-related decision making. Speculations regarding possible solutions to resolve the issues of debiasing are discussed. This chapter concludes with a direction for future research and the limitations of the research.

10.2 Summary of Findings

A review of the literature indicated that a substantial body of evidence has identified cognitive biases as being a significant factor in decision error in a range of fields (e.g., Croskerry, 2002; English et al., 2006). Anecdotal evidence from aircraft accident reports has suggested that cognitive biases may help explain pilots' impaired situational awareness during flight towards IMC (i.e., CAA NZ, 2005a; NTSB, 2008) but have received limited attention in prior studies of pilots' weather-related decision making.

10.2.1 Research Questions

A series of research questions was developed in order to explore the research problem identified in this thesis: can cognitive biases help to explain why pilots make inappropriate or ineffective decisions when approaching IMC? The findings from the first three research questions led to three additional research questions. These focused on whether the effects of cognitive biases can be mitigated by practical debiasing.

Research Question 1: Are pilots influenced by the anchoring effect when making weather-related decisions? Specifically, do participants anchor on outdated weather forecasts and under-adjust when assessing the actual weather conditions, and does the anchoring effect influence a pilot's decision to continue a VFR flight?

The findings of Study 1 provided evidence that pilots may be influenced by the anchoring effect when making weather-related decisions. Specifically, when pilots were exposed to a high anchor (forecasted high cloud and good visibility), their assessments of cloud height and visibility distance were greater than when they were exposed to a low anchor (forecast low cloud and poor visibility). Even though pilots were exposed to the same in-flight weather conditions, their judgment of the actual weather conditions was influenced by the 'initial' weather report, which acted as a high or low anchor.

In Study 1, there was also evidence to suggest that a pilot's perception of whether it was safe to continue a VFR flight was influenced by the anchoring effect. After

having been exposed to the high anchor, pilots felt it was safer to continue the flight compared to when they were exposed to the low anchor.

Pilots with a range of experience and age levels were influenced by the anchoring effect to a similar extent. Specifically, pilots with just a few dozen hours of flying experience were influenced by the anchoring effect to a similar extent to pilots with several thousand hours.

Research Question 2: Does confirmation bias occur in weather-related decision making? Specifically, what type of hypothesis testing strategy do pilots favour when judging the weather conditions?

The findings of Study 2 provided evidence that pilots have a tendency to favour a positive testing strategy by seeking confirmatory evidence when determining if the conditions are suitable to continue a VFR flight. Over-reliance on seeking confirmatory evidence (a positive testing strategy), as demonstrated in Study 2, can lead to serious consequences. Little difference in the choice of hypothesis testing strategy was observed between experience and age groups. However, there was evidence to suggest that pilots who had experienced VFR flight into IMC had a greater tendency to favour confirmatory evidence than pilots who had not experienced VFR flight into IMC.

Research Question 3: Does outcome bias occur in weather-related decision making? Specifically, does outcome information influence how the quality of a decision is rated? In addition, does outcome information influence subsequent decisions?

The findings of Study 3 provided evidence that outcome bias can affect how a weather-related decision is judged. The participants' assessments of the decision to conduct a VFR flight suggest that participants who were exposed to the negative outcome were more likely to rate the decision to conduct the flight as being worse compared to participants who were exposed to the positive outcome. Even though the information available to the pilot was the same for both outcome scenarios, the participant's judgment of the decision to conduct the flight was influenced by an outcome assigned by the researcher. Outcome bias was also observed in the evaluation of the pilot's assessment of risk, with scenarios with positive outcomes assessed as being less risky than those with negative outcomes.

Study 3 also found evidence to suggest that outcome information may influence how pilots make subsequent decisions when conducting a flight. When participants were exposed to a positive outcome, they were more likely to state that they would conduct the same flight themselves than if they had been exposed to a negative outcome. Similar to the first two studies, pilot characteristics were unreliable indicators for predicting the level or impact of outcome bias.

Research Question 4: Can debiasing reduce the rate at which the anchoring effect occurs in weather-related decision making? Specifically, after having been exposed

to the debiasing intervention ‘consider the alternative’ technique, do pilots have a reduced tendency to anchor on outdated weather forecasts?

Study 4 did not find evidence that debiasing could reduce the rate of the anchoring effect in weather-related decision making. After receiving the debiasing intervention, pilots still had a tendency to anchor on an outdated weather forecast to the same degree as pilots who were not exposed to the debiasing intervention.

Research Question 5: Can debiasing reduce the rate at which confirmation bias occurs in weather-related decision making? Specifically, after having been exposed to the debiasing intervention ‘consider the alternative’ technique, do pilots have a reduced tendency to seek confirmatory evidence?

Study 5 did not find evidence that debiasing can reduce confirmation bias in weather-related decision making. Specifically, after receiving the debiasing intervention, pilots still had a tendency to favour the positive testing strategy by seeking confirmatory evidence when selecting what type of evidence was most useful to them in deciding whether or not to continue their flight.

Research Question 6: Can debiasing reduce the rate at which outcome bias occurs in weather-related decision making? Specifically, after having been exposed to the debiasing intervention ‘consider the alternative’ technique, do pilots have a reduced tendency to use outcome information?

Study 6 did not find evidence that debiasing can reduce the rate of outcome bias in weather-related decision making. After receiving the debiasing intervention, outcome information still influenced how participants judged another flight, assessed the risk of that flight and assessed whether they would have conducted the flight themselves.

10.3 Discussion

10.3.1 Cognitive Biases in Weather-Related Decision Making

Flights operating under VFR that fly into IMC continue to be a major hazard for GA pilots (Batt & O'Hare, 2005; NTSB, 1989). Faulty pilot decision making has been shown to be a major factor behind this accident type (Hunter et al., 2003; Madhavan & Lacson, 2006). This research sought to further understand why pilots who were qualified only to fly under VFR flew into IMC, by exploring how the use of cognitive heuristics may lead to systematic biases, whereby VFR pilots make inappropriate or ineffective decisions when faced with IMC.

One of the key contributions this thesis makes to the literature is the strong evidence that cognitive biases influence weather-related decision making. Three cognitive biases that may potentially occur in pilot decisions to fly from VFR into IMC were identified: the anchoring effect, confirmation bias and outcome bias. Three vignette-based studies found evidence that pilots tended to anchor on and under-adjust using initial information ($n = 201$), favour a confirmatory strategy when testing a

hypothesis ($n = 278$) and evaluate judgments by the outcome rather than the decision process ($n = 300$).

In everyday situations, using heuristics can be useful in reducing cognitive load, especially when the consequences of a wrong decision are minimal (Croskerry, 2014; Baron & Hershey, 1988). Even within the aviation environment, it would be unreasonable to suggest heuristics should not be used. Gigerenzer and Gaissmaier (2011) suggested that not all heuristics result in a trade-off between effort and accuracy, therefore heuristics, when used appropriately may be beneficial. However, when dealing with uncertain situations, like those which occur in weather-related decision making, caution should be used, as inappropriate use of heuristics could increase the chances of an adverse outcome. When a VFR pilot is approaching deteriorating weather conditions, all three of these cognitive biases may result in impaired situational awareness. This could lead the pilot to perceive the weather conditions as being better than they really are, thus putting the VFR flight at a greater risk of flight into IMC.

It is, possibly, not surprising that cognitive biases are present in weather-related decision making, not least because they have been observed in a range of fields, including medicine (Croskerry, 2002). Medical decisions often have to be made under time and resource limitations, and the costs of making the wrong decision can involve loss of life or serious injury. Indeed, Gruppen et al. (1994) and Croskerry (2002) highlighted the dangers of cognitive biases in medicine, especially in the field of emergency medicine.

Consistent with previous research (Englich et al., 2006; Northcraft & Neale, 1987), experience was not associated with reducing the rate at which the cognitive biases occurred. These findings indicate that even as a pilot gains experience, he or she does not develop or retain the cognitive tools needed to avoid or reduce the effects of cognitive biases. This is consistent with findings in the area of risk perception; for example, Hunter's (2002) study of pilots' risk perception in adverse weather found that as pilots gained experience, their perception of risk reduced. One potential explanation is that as pilots gain experience flying in and around adverse weather with no significant consequences, their perception of the risk reduces. This may result in a pilot taking greater risks around adverse weather. However, if conditions shift slightly or if luck does not intervene a pilot may find themselves flying into IMC. A similar process may occur with the use of heuristics: each time a pilot uses a heuristic decision making strategy with no adverse consequences, this might reinforce the use of this strategy. The rewards from using this strategy (e.g., time saving), may further reinforce the benefits of heuristics. Despite this, over-reliance on a heuristic decision strategy may eventually result in an adverse outcome.

The findings of this thesis provide some insight into why some pilots continue to fly VFR into IMC, particularly within the context of the impaired situational assessment explanation. If a pilot is able to correctly recognise that VFR flight is no longer possible, the rational decision—and the decision which the average pilot is likely to choose—is to alter course to avoid entering IMC. The part cognitive biases may play in this decision making process is they may influence a pilot's ability to assess the weather environment accurately. To make the decision process even more challenging, cognitive biases do not necessary operate in isolation. Each of the

cognitive biases explored in this thesis may reinforce other cognitive biases. For example, the anchoring effect can feed back into confirmation bias, which may be supported by outcome bias (Nickerson, 1998; Wickens & Hollands, 2000).

In the context of VFR flight into IMC, all three cognitive biases can result in a pilot believing the conditions to be better than they are in reality on a flight near adverse weather. On any given flight, a VFR pilot will assess the weather information to determine if the conditions are suitable for a VFR flight. This initial assessment of the situation may be influenced by an anchor (e.g., a high anchor resulting in conditions being assessed more favourably). As the VFR flight progresses, with this initial assessment anchored within a pilot's mind, he or she is likely to seek cues that are consistent with their initial hypothesis (Vidulich, Wickens, Tsang, & Flach, 2010). This stage of the flight may also be influenced by a number of other factors identified by VFR flight into IMC literature. For example, if the pilot is meeting up with someone or returning to his or her home base, he/she may feel additional pressure to complete the flight. The more hypothesis-consistent cues a pilot is able to find, the more this reinforces their initial judgment that safe flight is possible. One potential cue a pilot may use is the outcome information of other flights. For example, a pilot may overhear another aircraft arrive at the planned destination. Even though this outcome information does not necessarily reflect that the conditions are suitable for VFR flight, it may reinforce the hypothesis that it is safe to continue. Each of these cognitive biases may result in the VFR pilot continuing further towards adverse weather.

If a pilot does find him/herself inadvertently entering IMC, he/she may have luck on their side to find a safe passage back to VFR conditions. Unfortunately for a VFR pilot who has inadvertently entered IMC, luck only needs to fail once for the flight to end in tragedy.

Considering the findings of this thesis, it is possible to speculate how a fatal VFR flight into IMC accident in British Columbia was potentially influenced by cognitive biases (TSB, 1996). The aircraft was on a VFR flight approaching the airport in poor weather. The pilot's decision to continue towards the airport may have been influenced by a potential anchor in the form of a report from another pilot ahead, which stated that the cloud base was 900 feet, although the actual cloud ceiling had been varying between 300 and 500 feet. The accident report suggested that confirmation biases may have influenced the cues the pilot used once the pilot had selected the course of action. A very compelling cue would be required to change one's mind. Indeed, the report suggested the pilot probably had a tendency to use cues to confirm the validity of the intended plan of action. The positive outcome information, in the form of another aircraft having also recently landed, may have reinforced this decision to continue. Unfortunately, the aircraft flew into IMC and crashed a short time later (TSB, 1996).

10.3.2 Debiasing Weather-Related Decision Making

Considering the serious consequences that can arise from decision error, it is important to explore methods to help pilots accurately assess the weather conditions by reducing the impact of cognitive biases. Research into correcting or preventing

cognitive biases (debiasing) has largely been overshadowed by cognitive bias research, but is a critical area for improving decision making (Larrick, 2004). The strong presence of cognitive biases in Studies 1–3 led to an expansion of the original research questions: can the effects of cognitive biases be corrected or prevented through practical debiasing? Aiding pilots to avoid cognitive bias will enable them to make better informed decisions by forming an accurate assessment of the dynamic weather environment.

Using the debiasing technique ‘consider the alternative’, 101 pilots participated in the debiasing studies, which required them to consider a wider range of information to counteract the effect of cognitive biases. Evidence suggested that the ‘consider the alternative’ debiasing technique held the most promise and was the most practical to implement. This technique has been shown to mitigate the effects of a range of biases, including the biases investigated in this thesis (Larrick, 2004). The findings of the three debiasing studies indicated that the debiasing intervention was largely unsuccessful in reducing all three cognitive biases and therefore not supporting the three debiasing hypotheses. The findings indicated that even after receiving the debiasing intervention, pilots still had a tendency to anchor on initial pieces of information, a tendency to favour confirmatory evidence when testing a hypothesis and a tendency to evaluate judgments by their outcome.

There has been some success in debiasing, although, a range of studies have found debiasing to be less successful (Fischhoff, 2002). Studies that have successfully reduced cognitive bias have generally done so by using intrusive methods, such as requiring participants to write down long lists of alternative or contrary information

(e.g., Mussweiler et al., 2000; Roese, 2004). However, this method of debiasing does not necessarily reflect how pilots' decision making can be improved in a real-world environment. This is particularly relevant in the situation which a VFR pilot faces during flight towards IMC, when the pilot is under high workload and time pressure.

10.3.3 Implications of Cognitive Biases in Weather-Related Decision Making

The primary aim of this thesis was to investigate whether cognitive biases occur in weather-related decision making and if they can be reduced through debiasing. Previous research has found that if pilots have an accurate assessment of the weather conditions, they are more likely to make rational decisions. Putting this another way, if pilots have accurately assessed the visibility or cloud height to be too low, they are more likely to recognise that VFR flight may no longer be possible and will take the necessary action to avoid flying into IMC (Goh & Wiegmann, 2001; Wiegmann et al., 2002). The research presented in this thesis found evidence that when faced with deteriorating weather conditions, cognitive biases may influence a pilot's ability to form an accurate situational assessment of the dynamic weather environment.

Similarly to the challenge faced in combating VFR flight into IMC, when addressing the problem of cognitive biases in weather-related decision making, a number of measures may be required. The findings of the debiasing studies highlight the major challenge faced with debiasing, not just in regards to the three cognitive biases explored in this thesis but cognitive biases in general. Therefore, it is important not to rely on debiasing as the sole tool to improve pilots' decision making in

deteriorating weather conditions. The findings of the debiasing studies suggest that for debiasing to be successful across a range of cognitive biases, more extensive training may provide more encouraging results. Extensive training would need to provide the necessary practice and feedback to ensure that pilots could consolidate the debiasing technique. This type of training could easily be incorporated into the current human factors training of *ab initio* pilots.

Weather-related decision making is often considered a skill that a pilot will gradually develop as he or she is exposed to different weather-related environments (Wiggins & O'Hare, 1995). However, the findings that experience had little impact on reducing the rate at which the cognitive biases occurred suggest that mere practice in making decisions is not necessarily an effective way to avoid cognitive biases and improve decision quality (Wickens et al., 2013). As a result, further research is necessary to explore practical decision support tools for pilots.

10.4 Direction for Future Research

Ultimately, the goal is to improve safety, so a realistic and practical framework is important to help pilots manage the threat of cognitive biases. Wickens et al. (2013) suggested four areas that could be adopted to address the problem of cognitive biases: training in debiasing, proceduralization, displays and information, and automation and decision support tools. Individually, each of these measures is unlikely to combat all cognitive biases in weather-related decision making, so a combined approach is likely to be required. Training in debiasing was discussed in

the previous section; the other three areas suggested by Wickens et al. (2013) are discussed as potential areas that could be explored.

10.4.1 Proceduralization

Proceduralization has been shown to improve decision making in some situations by encouraging decision makers to follow a predefined set of rules, ensuring that all the necessary information is processed (Rasmussen, 1983; Ricchiute, 1998). Adopting a fully proceduralized approach to VFR flight is impracticable; however, exploring aspects of VFR GA flying could have value. One example that already exists is the concept of encouraging pilots to set and follow personal meteorological minima. Personal meteorological minima were first developed by Kirkbride, Jensen, Chubb, and Hunter (1996), which require pilots to write down their own personalised weather conditions that they are comfortable flying in. For most pilots, these are likely to be well above the legal VFR meteorological weather conditions.

This type of decision making tool is likely to help reduce some, but not all, cognitive biases. For example, the impact of confirmation bias is likely to benefit from personal meteorological minima, as it requires pilots to check key cues (e.g., cloud height and visibility) against pre-defined criteria. The main disadvantage is this that type of procedure is it is still susceptible to other 'front end' cognitive biases such as the anchoring effect. Personal meteorological minima rely on pilots having an accurate perception of the weather conditions. However, the findings of this thesis have demonstrated that pilots' assessments of the conditions can be influenced by cognitive biases. This is particularly the case in areas with limited weather information, such as during the cruise phase of flight, which also happens to be the

phase VFR flight into IMC is more likely to occur (Detwiler et al., 2008; Goh & Wiegmann, 2001). In the cruise phase of flight, the pilot is more likely to rely on broad regional weather forecasts (rather than aerodrome specific forecasts) and therefore is more likely to be at risk from the anchoring effect and outcome bias.

10.4.2 Displays and Information

Displays and information may be one area that could help pilots to avoid some cognitive biases, especially cognitive biases that occur at the ‘front end’ of the decision process, such as the anchoring effect (Mosier & Fischer, 2010). A number of cognitive biases, including confirmation bias, anchoring and outcome bias, are partly caused by a person focusing on a narrow or outdated set of evidence (Larrick, 2004). Improving the flow of information to the pilot, particularly in flight, may support the pilot in making more informed decisions and possibly counter the effects of some cognitive biases.

Introducing advanced displays into the VFR aircraft cockpit requires some caution. First of all, a number of cognitive biases stem from how our brain makes complex decisions (Tversky & Kahneman, 1974). Adding even more information may make matters worse by saturating the pilot with information, therefore encouraging pilots to use heuristics to process the vast amount of information the VFR pilot faces when flying in deteriorating weather conditions.

The introduction of advanced technology into the general aviation cockpit is receiving more attention. Johnson et al. (2006) investigated the effect of advanced cockpit displays (synthetic displays and moving maps) on general aviation pilot decision making. Wiggins (2007) have explored the impact GPS on decision making strategies, with a focus on moving map displays in the general aviation cockpit. Advanced displays enable complex tasks, like navigation to be conducted with reduced pilot workload, which in theory enables more cognitive resources to be allocated towards other tasks, such as weather-related decision making (Casner, 2005).

Even though development of advanced displays and technologies is an appealing response to VFR flight into IMC accidents, there are considerable challenges to overcome (Wiggins, 2007). Johnson et al. (2006) found that pilots who were operating an aircraft with synthetic vision and moving map displays were more likely to continue flight into deteriorating weather conditions compared to pilots without these advanced technologies. This is in part because pilots spend more time looking inside the cockpit at various displays rather than looking outside (Johnson et al., 2006; Williams, Yost, Holland, & Tyler, 2002). Furthermore, it has been suggested advanced technology may result in some pilots taking greater risks near adverse weather. Beringer and Ball (2004) found that pilots using high-resolution weather radar displays were more likely to continue a flight towards adverse weather with the expectation that they could fly around the adverse weather with the aid of the displays.

Research into the impact of how basic meteorological information is displayed to the pilot has resulted in promising findings. O'Hare and Stenhouse (2009) have explored

graphical displays, which transform the traditional text based weather information into graphical form. Graphic displays have shown to improve recall of meteorological information, and resulted in less information omitted. However, Ahlstrom and Suss (2015), found pilots have difficulty detecting changes in graphical weather displays. The findings of this study suggest it is not only important to consider how weather information is displayed but also consider the likelihood of pilots detecting a change has occurred.

Research in the area of advanced technologies has significant potential, not just in reducing cognitive biases but also generally in reducing VFR flight into IMC. One positive side effect of advanced technologies has been shown in how a pilot handles the aircraft once they enter IMC. Johnson et al. (2006) found that pilots with advanced technology all successfully flew the aircraft out of hazardous conditions and landed safely.

10.4.3 Automation and Decision Support Tools.

Automation and decision support tools aim to help pilots bridge the gap between the vast amount of information presented to them and putting the information into a meaningful state. Automated decision support systems have had some success in the medicine field, which could potentially be adapted for weather-related decision making. For example, automated reminders to encourage pilots to check for the latest weather information, or automated guidance through the weather assessment information and decision process could be beneficial (Karsh, 2010; Vashitz et al., 2009). The principal benefit of carefully designed automation systems is the reduced

user workload that can result (Wickens et al., 2013). Workload reductions could potentially mitigate the fundamental reason why heuristics are used in the first place by providing the pilot with sufficient mental capacity to carefully consider the available information.

10.4.4 Direction of Aviation Cognitive Bias Research

Numerous cognitive heuristics and biases have been identified. For example, Dobelli (2013) reviewed over 100 biases in his recent book and at least 40 biases have been found to affect clinical reasoning (Mamede et al., 2010). Future research could explore the use of other heuristics and whether they lead to cognitive biases in pilots' decisions. The cognitive biases discussed in this thesis concentrated on a fairly specific situation: VFR pilots approaching IMC. Pilots make a range of decisions during any given flight which can have serious implications for the safety of that flight. The impact of cognitive bias in other areas of aviation may further our understanding of other problem areas in aviation, such as why pilots on commercial airlines frequently continue to fly unstabilised approaches, thus putting the flight at risk (Flight Safety Foundation, 2009).

Finally, in this thesis, the author explored the debiasing technique, 'consider the alternative' to help improve decision making. Previous literature has indicated that this technique is likely to hold the most promise in regards to debiasing the three cognitive biases within this thesis. This particular technique could therefore be explored further. As discussed previously, this thesis delivered the 'consider the alternative' technique in a relatively passive and brief session, and possibly more

extensive training may provide more encouraging results. There is also scope to explore other debiasing techniques. As outlined in Chapter Six, a range of techniques have been investigated, from motivational strategies to extensive training. The key aspect to keep in mind when exploring other debiasing techniques is to consider the debiasing technique's ability to be implemented into the weather-related decision making environment in real world.

10.5 Limitations

It is important to note a number of potential limitations in the studies that comprise this thesis. Each study in this thesis outlined specific limitations and only the limitations that apply to the thesis in general are discussed here. All six studies involved the use of a scenario-based methodology. A flight simulator-based platform would have been preferable, but was beyond the resources available for this thesis. The gap between the controlled experiments presented in this thesis and the real-world environment are likely to be considerable. Whether similar results would be obtained in actual flight conditions cannot be determined. However, in real life situations, the time and workload stress associated with flying the aircraft may lead to increased stress levels, which have been shown to deteriorate decision making ability by inducing a person to place greater reliance upon heuristics (Wickens et al., 1993). It has been suggest that this type of methodology, at the very least, acts as an indicator to real life behaviour (see Exum et al., 2012 for a more complete review); therefore, the findings of this thesis should at least provide some indication of the impact of cognitive biases in the real-world environment.

CHAPTER ELEVEN

CONCLUSION

This thesis began with the identification of a problem that has been a concern in GA for some time: the high fatality rate that continues to occur when a VFR flight enters IMC. This research explored how the use of cognitive heuristics can lead to systematic biases, whereby VFR pilots make inappropriate or ineffective decisions when faced with adverse weather. A substantial body of evidence has identified cognitive biases as being a significant factor in decision error in a range of fields (e.g., Croskerry, 2002; English et al., 2006) but this topic has received limited attention in the pilot weather-related decision making literature.

In this thesis, the author found strong evidence to suggest that when a VFR pilot is making a weather-related decision, he or she could be influenced by three cognitive biases: anchoring effect, confirmation bias and outcome bias. Pilots had a tendency to anchor and under-adjust on initial pieces of information, have a tendency to favour confirmatory evidence when testing hypotheses and to evaluate decisions based upon their outcome rather than their quality at the time they were made. Of particular concern in VFR flight is how each of these cognitive biases could potentially lead a pilot to perceive the weather conditions as being better than they are in reality. This could result in a pilot continuing towards deteriorating weather conditions, when the rational decision would be to alter course.

Considering the serious consequences that can arise from cognitive biases, it is appropriate to consider techniques to counter them. This research found strong evidence that pilots with a range experience levels were influenced by each of the cognitive biases to a similar extent. This would suggest that pure practice is not an appropriate counter-measure against cognitive biases.

In this thesis, the author attempted to combat cognitive biases using a promising debiasing technique, 'consider the alternatives'. The debiasing studies were unsuccessful in reducing all three cognitive biases in this thesis. Even after receiving the debiasing intervention, pilots still had a tendency to be influenced by the three cognitive biases. The debiasing studies confirmed the major difficulty faced in providing pilots with practical real-life solutions to avoid cognitive biases. It would be unreasonable to suggest cognitive heuristics should not be used at all; cognitive heuristics are a valuable decision making tool which can ease a pilot's cognitive load during decision making. The ideal solution would be to aid pilots find the right balance between using heuristics when appropriate and switching to an analytical decision making strategy when heuristics should be avoided. Such procedures will enable them to make more informed decisions by forming an accurate assessment of the dynamic weather environment.

Similar to the challenge faced in combating VFR flight into IMC, when addressing the problem of cognitive biases in weather-related decision making, a number of measures may be required. Although the 'consider the alternative' method was unsuccessful, training in debiasing may provide some benefit to avoiding cognitive biases; however, more extensive training is likely to be necessary to produce more

encouraging results. Along with exploring other debiasing techniques, there are a number of potential future research areas that can be investigated. For example, the role of how and when information is displayed or presented to pilots plays in helping reduce some cognitive biases could be explored.

Considering the findings of this thesis, it is important to develop practical applications to aid pilots manage the threat of cognitive biases. Because of the nature of pilot training, especially in the early stages, novice pilots are not often exposed to the type of weather conditions that a VFR pilot approaching IMC would experience. As a result, during pilot training, a pilot may not be in a position to practice in-flight weather related decision making (Johnson & Wiegmann, 2015). One challenge is safely simulating VFR flight towards adverse weather conditions to enable novice pilots to practice weather related decision making. Unlike obtaining important aircraft handling skills, such as handling an engine failure in flight (where a safe reduction in power allows the emergency to be safely simulated), it is difficult to simulate deteriorating weather conditions without putting the aircraft into danger.

One potential way forward is for pilots to use computer based training (CBT). CBT has the potential to provide a safe environment for pilots to practice the skills necessary to identify the threat of cognitive biases during weather related decision making. Similar training support tools have had some success. Wiggins and O'Hare (2003b) developed a CBT program called 'weatherwise', which aided pilots to recognise key cues during deteriorating weather conditions. A similar product could be used as a practical training tool for training in cognitive biases. A CBT product could incorporate similar scenario based situations as presented in this thesis. As well

as highlighting the potential impact each of these cognitive biases has on weather-related decision making, the CBT could also provide some guidance on mitigating their impact. For example, the anchoring effect scenario could be used to illustrate the importance of obtaining the most current weather information.

In conclusion, the three common cognitive biases (the anchoring effect, conformation bias and outcome bias) were each found to affect the decision making processes of pilots when they were considering whether to continue a VFR flight towards adverse weather. Importantly, in each case, the effect was likely to place pilots in greater danger. Considering the serious consequences that can result from cognitive biases, it is important that pilots at least understand the mechanisms leading to these biases and that continued efforts are made to reduce the effects of cognitive bias.

REFERENCES

- Ahlstrom, U., & Suss, J. (2015). Change Blindness in Pilot Perception of METAR Symbology. *International Journal of Industrial Ergonomics*, 46, 44-58.
- Air Accident Investigation Branch [AAIB]. (2014). *Aircraft Accident Report EW/G2014/06/18* (Report No. EW/G2014/06/18). Farnborough, UK: AAIB.
- Aircraft Owner and Pilots Association [AOPA]. (2014). *23rd Joseph T. Nall Report: General Aviation Accident in 2011*. Frederick, MD: AOPA Air Safety Foundation.
- Alicke, M. D., & Davis, T. L. (1989). The role of a posteriori victim information in judgments of blame and sanction. *Journal of Experimental Social Psychology*, 25(4), 362-377.
- Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin*, 110(2), 486-498.
- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1), 124-140.
- Arkes, H. R., Faust, D., Guilmette, T. J., & Hart, K. (1988). Eliminating the hindsight bias. *Journal of Applied Psychology*, 73(2), 305-307.

REFERENCES

- Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*(4), 471-499.
- Ask, K., & Granhag, P. A. (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling, 2*(1), 43-63.
- Australian Transport Safety Bureau [ATSB]. (2013). *Aviation Occurrence Investigation AO-2012-076*. Canberra, Australia: Australian Transport Safety Bureau.
- Baron, J. (2008). *Thinking and deciding* (4th ed.). New York, NY: Cambridge University Press.
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54*(4), 569-579.
- Batt, R., & O'Hare, D. (2005). Pilot behaviour in the face of adverse weather: A new look at an old problem. *Aviation, Space and Environmental Medicine, 76*(6), 552-559.
- Beringer, D. B., & Ball, J. D. (2004). *The Effects of NEXRAD Graphical Data Resolution and Direct Weather Viewing on Pilots' Judgments of Weather Severity and Their Willingness to Continue a Flight*. (Report No. DOT/FAA/AM-04/5). Washington, DC: Office of Aerospace Medicine.

REFERENCES

- Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19, 7–42.
- Cameron, A. C. (2004). *Kurtosis*. SAGA Research Methods. Retrieved from <http://srmo.sagepub.com>
- Casner, S. M. (2005). The effect of GPS and moving map displays on navigational awareness while flying under VFR. *International Journal of Applied Aviation Studies*, 5, 153-165.
- Cervone, D., & Peake, P. K. (1986). Anchoring, efficacy and action: The influence of judgment heuristics on self-efficacy judgment and behaviour. *Journal of Personality and Social Psychology*, 50(3), 492-501.
- Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioural Decision Making*, 7(4), 223-242.
- Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, 18(3), 293-328.
- Civil Aviation Authority of New Zealand [CAA NZ]. (2005a). *CAA Occurrence Number 05/2471*. Wellington, New Zealand: Civil Aviation Authority of New Zealand.
- Civil Aviation Authority of New Zealand [CAA NZ]. (2005b). *CAA Occurrence Number 05/3469*. Wellington, New Zealand: Civil Aviation Authority of New Zealand.

REFERENCES

- Civil Aviation Authority of New Zealand [CAA NZ]. (2011). *Rule Part 61: Pilot Licences and Ratings*. Wellington, New Zealand: Civil Aviation Authority of New Zealand.
- Civil Aviation Authority of New Zealand [CAA NZ]. (2012). *CAA Occurrence Number 11/1504*. Wellington, New Zealand: Civil Aviation Authority of New Zealand.
- Civil Aviation Authority of New Zealand [CAA NZ]. (2015a). *Civil Aviation Authority of New Zealand Safety Information*. Retrieved from https://www.caa.govt.nz/safety_info/posters.htm
- Civil Aviation Authority of New Zealand [CAA NZ]. (2015b). *Rule Part 91: General Operating and Flight Rules*. Wellington, New Zealand: Civil Aviation Authority of New Zealand.
- Civil Aviation Safety Authority of Australia [CASA]. (2013). *Annual Report 2012-2013, Operating Statistic, Flight operations and personnel*. Canberra, Australia: Civil Aviation Safety Authority.
- Clarkson, P. M., Emby, C., & Watt, V. W. (2002). Debiasing the outcome effect: The role of instruction in an audit litigation setting. *Auditing: A Journal of Practice & Theory*, 21, 7-20.
- Coyne, J., Baldwin, C., & Latorella, K. (2008). Pilot weather assessment: Implications for visual flight rules flights into instrument meteorological conditions. *The International Journal of Aviation Psychology*, 18(2), 153-166.

REFERENCES

- Croskerry, P. (2002). Achieving quality in clinical decision making: Cognitive strategies and detection of bias. *Academic Emergency Medicine, 9*(11), 1184-1204.
- Croskerry, P. (2014). Bias: A normal operating characteristic of the diagnosing brain. *Diagnosis, 1*(1), 23-27.
- Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: Origins of bias and theory of debiasing. *BMJ Quality & Safety, 22*, 58-64.
- Detwiler, C., Holcomb, K., Hackworth, C., & Shappell, S. (2008). *Understanding the human factors associated with visual flight rules flight into instrument meteorological conditions* (Report No. DOT/FAA/AM-08/12). Washington, DC: Office of Aerospace Medicine.
- Dillman, B. G., Voges, J., & Robertson, M. (2007). Safety occurrences student perceptions regarding failures to report. *Journal of Aviation Management and Education, 1*, 1-14.
- Dobelli, R. (2013). *The art of thinking clearly*. New York, NY: HarperCollins.
- Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57*(6), 1082-1090.
- Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influences of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin, 32*(2), 188-200.

REFERENCES

- Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter provided anchors. *Psychological Science, 12*(5), 391-396.
- Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17*(4), 311-318.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behaviour Research Methods, Instruments, & Computers, 28*(1), 1-11.
- Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13*(3), 378-395.
- Exum, M. L., Turner, M. G., & Hartman, J. L. (2012). Self-reported intentions to offend: All talk and no action?. *American Journal of Criminal Justice, 37*(4), 523-543.
- Federal Aviation Administration [FAA]. (2004). *Airplane Flying Handbook* (Report No. FAA-H-8083-3A). Washington, DC: U.S. Department of Transportation.
- Federal Aviation Administration [FAA]. (2009). *U.S. Civil Airmen Statistics, 2009*. Retrieved from www.faa.gov/data_statistic/aviation_data_statistic/civil_airmen_statistic
- Fischhoff, B. (1975). Hindsight≠foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 288-299.

REFERENCES

- Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases* (pp. 422–444). Cambridge, UK: Cambridge University Press.
- Fischhoff, B. (2002). Heuristics & Biases in application. In T. Gilovich, D. Griffin, & D. Kahneman, D.(Eds.), *Heuristics and Biases: The psychology of intuitive judgment* (pp. 730-748). New York, NY: Cambridge University Press.
- Flight Safety Foundation [FSF]. (2009). *Reducing the Risk of Runway Excursion*. Retrieved from <http://flightsafety.org/files/RERR/fsf-runway-excursions-report.pdf>
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, *18*, 253–92.
- Giffin, W. C., & Rockwell, T. H. (1984). Computer-aided testing of pilot response to critical in-flight events. *Human factors*, *26*(5), 573-581.
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451-482.
- Gilbey, A., & Hill, S. (2012). Confirmation bias in general aviation lost procedure. *Applied Cognitive Psychology*, *26*(5), 785-795.
- Gilbey, A., Tani, K., & Tsui, W. H. K. (2015). Outcome bias and under-reporting in aviation. *Applied Cognitive Psychology*, *30*, 141–151.

REFERENCES

- Gino, F., Moore, D. A., & Bazerman, M. H. (2009). No harm, no foul: The outcome bias in ethical judgments. *Harvard Business School NOM Working Paper*, (08-080).
- Gino, F., Shu, L. L., & Bazerman, M. H. (2010). Nameless+ harmless= blameless: When seemingly irrelevant factors influence judgment of (un)ethical behavior. *Organizational Behavior and Human Decision Processes*, *111*, 93-101.
- Goh, J., & Wiegmann, D. D. (2001). Visual flight rules flight into instrument meteorological conditions: An empirical investigation of the possible causes. *International Journal of Aviation Psychology*, *11*(4), 359-379.
- Green, R. G., James, H. M. M., Gradwell, D., & Green, R. L. (1996). *Human Factors for Pilots* (2nd Ed.). England: Ashgate Publishing Limited.
- Gruppen, L. D., Margolin, J., Wisdom, K., & Grum, C. M. (1994). Outcome bias and cognitive dissonance in evaluating treatment decisions. *Academic Medicine*, *69*(10).
- Halberstadt, J. B. & Levine, G. M. (2006). Effects of Reason Analysis on the Accuracy of Predicting Basketball Games. *Journal of Applied Social Psychology*, *29*(3), 517-530.
- Hastie, R., & Dawes, R. M. (2010). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Oak, CA: SAGE Publication.

REFERENCES

- Hunter, D. R. (2002). *Risk perception and risk tolerance in aircraft pilots* (Report No. DOT/FAA/AM-02/17). Washington, DC: Office of Aerospace Medicine.
- Hunter, D. R. (2006). Risk perception among general aviation pilots. *The International Journal of Aviation Psychology*, *16*(2), 135-144.
- Hunters, D. R., Martinussen, M., & Wiggins, M. (2003). Understanding how pilots make weather-related decisions. *The International Journal of Aviation Psychology*, *13*(1), 73-87.
- Hunter, D. R., Martinussen, M., Wiggins, M., & O'Hare, D. (2011). Situational and personal characteristics associated with adverse weather encounters by pilots. *Accident Analysis and Prevention*, *43*(1), 176-186.
- Hutton, R. J. B., & Klein, G. (1999). Expert decision making. *Systems Engineering*, *2*(1), 32-45.
- Hyams, D. S., Fan, T. P., & Kuchar, F. J. (1998). Survey of In-Flight Replanning Performed on the Flight Deck. *17th Digital Avionics Systems Conference*, Seattle, WA.
- Iglewicz, B., & Hoaglin, D. (1993). *How to detect and handle outliers*. Milwaukee, WI: American Society for Quality Control.
- International Council of Aircraft Owner and Pilot Associations [ICAOPA]. (2014). *27th IAOPA World Assembly Statistical Report*. International Council of Aircraft Owner and Pilot Associations.

REFERENCES

Jensen, R. S. (1995). *Pilot judgment and crew resource management*. Brookfield, VT: Ashgate Publishing Limited.

Jensen, R. S., Adrion, J., & Maresh, J. (1986). *A preliminary investigation of the application of the DECIDE model to aeronautical decision making*. Columbus, OH: The Ohio State University Aviation Psychology Laboratory.

Jensen, R. S., & Benel, R. A. (1977). *Judgment evaluation and instruction in civil pilot training*. (FAA-RD-78-24). Washington, DC: Federal Aviation Administration.

Johnson, M., & Wiegmann, C. M. (2015). VFR Into IMC: Using Simulation to Improve Weather-Related Decision-Making. *The International Journal of Aviation Psychology*, 25(2), 63-76.

Johnson, N., Wiegmann, D., & Wickens, C. (2006). Effects of advanced cockpit displays on general aviation pilots' decision to continue visual flight rules into instrument meteorological conditions. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*.

Kahneman, D. (2011). *Thinking fast and slow*. New York, NY: Farrar, Stratus & Giroux.

Karsh, B. T. (2010). *Clinical practice improvement and redesign: How change in workflow can be supported by clinical decision support*. (AHRQ Publication No. 09-0054-EF). Rockville, MD: Agency for Healthcare Research and Quality.

REFERENCES

- Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioural confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behaviour, 27*(2), 187-203.
- Kennedy, J. (1995). Debiasing the curse of knowledge in audit judgment. *The Accounting Review, 70*(2), 249-273.
- Kirkbride, L. A., Jensen, R. S., Chubb, G. P., & Hunter, D. R. (1996). *Developing the Personal Minimums Tool for Managing Risk During Preflight Go/No-Go Decisions*. (Report No. DOT/FAA/AM-96/19). Washington, DC: Office of Aviation Medicine.
- Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*(2), 211-228.
- Kochan, J. A., Jensen, R. S., Chubb, G. P., & Hunter, D. R. (1997). *A New Approach to Aeronautical Decision-Making: The Expertise Method*. (Report No. DOT/FAA/AM-97/6). Washington, DC: Office of Aerospace Medicine.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 107-118.
- Koslowski, B., & Maqueda, M. (1993). What is confirmation bias and when do people actually have it?. *Merrill-Palmer Quarterly, 39*(1), 104-129.

REFERENCES

- Larrick, R. P. (2004). Debiasing. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316-357). Oxford, UK: Blackwell.
- Larrick, R. P., Morgan, J. N., & Nisbett, R. E. (1990). Teaching the use of cost benefit reasoning in everyday life. *Psychological Science, 1*(6), 362-370.
- Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare?. *Perspectives on Psychological Science, 4*(4), 390-398.
- Lopes, L. L. (1987). Between hope and fear: The psychology of risk. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, (pp. 255–295). San Diego, CA: Academic Press.
- Madhavan, P., & Lacson, F. (2006). Psychological factors affecting pilots decisions to navigate in deteriorating weather. *North American Journal of Psychology, 8*(1), 47-62.
- Mamede, S., Gog, T., Berge, K., Rikers, R. M. J. P., Saase, J. L. C., Guldener, C., & Schmidt, H. G. (2010). Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine resident. *The Journal of the American Medical Association, 304*(11).
- Mazzocco, P. J., Alicke, M. D., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Psychology, 26*, 131-146.

REFERENCES

- Midkiff, A., & Hansman, R. J. (1992). *Identification of important party line informational elements and the implication for situational awareness in the datalink environment* (Report: ASL-92-2), Cambridge, MA: Massachusetts Institute of Technology.
- Mosier, K. L., & Fisher, U. (2010). Judgment and decision making by individuals and teams: Issues, models, and applications. *Review of Human Factors and Ergonomics*, 6(1), 198-255.
- Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Society for Personality and Social Psychology*, 26(9), 1142-1150.
- Muthard, E. K., & Wickens, C. D. (2001). *Change detection in a flight planning task environment: An examination of the confirmation bias and its relation to decision making error* (Technical Report ARL-01-18/NASA-01-9). Savoy, IL: University of Illinois at Urbana-Champaign, Aviation Research Lab.
- Nagel, D. C. (1988). Human error in aviation operations. In E.L. Wiener (Ed.), *Human Factors in Aviation*. San Diego, CA: Academic Press.
- National Transport Safety Board [NTSB]. (1968). *An analysis of U.S. general aviation accidents involving weather as a cause/related factor 1966* (Notion 155). Washington, DC: NTSB.

REFERENCES

National Transport Safety Board [NTSB]. (1989). *General aviation accidents involving VFR flight into IMC* (Report no. NTSB/SR-89-01). Washington, DC: NTSB.

National Transport Safety Board [NTSB]. (2005). *Safety study: Risk factors associated with weather related GA accidents*. (Report No. NTB/SS-05/01). Washington, DC: NTSB.

National Transport Safety Board [NTSB]. (2008). *Attempted takeoff from wrong runway Comair Flight 5191, August 27, 2006*. (Accident Report NTSB/AAR-07/05). Washington, DC: NTSB.

National Transport Safety Board [NTSB]. (2012). *Aircraft Accident Report ERA11GA066* (Report No. NTB/ERA11GA066). Washington, DC: NTSB.

National Transport Safety Board [NTSB]. (2013). *Aircraft Accident Report ERA12FA205*. (Report No. NTB/ERA12FA206). Washington, DC: NTSB.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175-220.

Nisbett, R. E., Krantz, D. H., Jepson, D., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90*(4), 339-63.

REFERENCES

- Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behaviour & Human Decision Processes*, 39(1), 84-97.
- O'Hare, D. (1990). Pilots' perception of risks and hazards in general aviation. *Aviation, Space & Environmental Medicine*, 61(7), 599-603.
- O'Hare, D., Hunter, D, R., Martinussen, M., & Wiggins, M. (2011). Classification and predication of pilot weather encounters: A discriminant function analysis. *Aviation, Space, and Environmental Medicine*, 82(5), 543-549.
- O'Hare, D., Mullen, N., & Arnold, A. (2010). Enhancing Aeronautical Decision Making through Case-Based Reflection. *The International Journal of Aviation Psychology*, 20(1), 48-58.
- O'Hare, D., & Owen, D. (2001). The 'Where' and 'Why' of cross-country VFR crashes: Database and simulation analyses. *Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors & Ergonomics Society.
- O'Hare, D., & Roscoe, S, N. (1990). *Flight deck performance: The human factor*. Ames: Iowa State University
- O'Hare, D., & Smitheram, T. (1995). Pressing on into deteriorating weather conditions: An application of behavioural decision theory to pilot decision making. *The International Journal of Aviation Psychology*, 5(4), 189-197.

REFERENCES

- O'Hare, D., & Stenhouse, N. (2009). Under the weather: An evaluation of different modes of presenting meteorological information for pilots. *Applied Ergonomics*, *40*(4), 688-693.
- O'Hare, D., & Waite, A. (2012). Effect of pilot experience on recall of information from graphical weather displays. *The International Journal of Aviation Psychology*, *22*(1), 1-17.
- O'Hare, D., Wiggins, M., Batt, R., & Morrison, D. (1994). Cognitive failure analysis for aircraft accident investigation. *Ergonomic*, *37*(11), 1855-1869.
- Orasanu, J., Martin, L., & Davison, J. (2009). Cognitive and contextual factors in aviation accidents: Decision error. In E. Salas & G. Klein (Eds.). *Linking Expertise and Naturalistic Decision Making* (pp. 209-225). New York, NY: Lawrence Erlbaum Associates Inc.
- Pauley, K., O'Hare, D., & Wiggins, M. (2008). Risk tolerance and pilot involvement in hazardous events and flight into adverse weather. *Journal of Safety Research*, *39*, 403-411.
- Pines, J. M. (2006). Profiles in patient safety: Confirmation bias in emergency medicine. *Academic Emergency Medicine*, *13*(1), 90-94.
- Plous, S. (1989). Thinking the unthinkable: The effects of anchoring on likelihood estimates of nuclear war. *Journal of Applied Social Psychology*, *19*(1), 67-91.

REFERENCES

- Pogarsky, G. (2004). Projected offending and contemporaneous rule-violation: Implications for heterotypic continuity. *Criminology*, 42(1), 111-136.
- Pronin, E. (2007). Perception and misperception of bias in human judgment. *Trends in Cognitive Sciences*, 11, 37-43.
- Rasmussen, J. (1983). Skills, rules, and knowledge; signal, signs and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(2), 257-266.
- Rhoda, D. A., & Pawlak, M. L. (1999a). The thunderstorm penetration / deviation decision in the terminal area. *Eighth Conference on Aviation, Range and Aerospace Meteorology*, American Meteorological Society, Dallas, TX.
- Rhoda, D. A., & Pawlak, M. L. (1999b). *An assessment of thunderstorm penetration and deviation by commercial aircraft in the terminal area*. (Project Report NASA - A/2), MIT Lincoln Laboratory.
- Ricchiute, D. N. (1998). Evidence, memory, and casual order in a complex audit decision task. *Journal of Experimental Psychology: Applied*, 4(1), 3-15.
- Roese, N. J. (2004). Twisted pair: Counterfactual thinking and the hindsight bias. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316-357). Oxford, UK: Blackwell.
- Sarter, N. B., & Schroeder, B. (2001). Supporting decision making and action selection under time pressure and uncertainty: The case of in-flight icing. *Human Factors*, 43(4), 573-583.

REFERENCES

- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2), 129–138.
- Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3), 437-446.
- Teasley, B. E., Leventhal, L M., Mynatt, C. R., & Rohlman, D. S. (1994). Why software testing is sometimes ineffective: Two applied studies of positive test strategy. *Journal of Applied Psychology*, 79, 142-155.
- Thom, T. (2003). *The air pilot's manual: Human factors and pilot performance* (2nd ed.). England: Airlife Publishing.
- Thorpe, J. A., Martin, E. L., Edwards, B. J., & Eddows, E. E. (1976). *Situational emergency training: F-15 Emergency Procedures Training Program*. (AFHRL-TR-76-47(1).). Williams Air Force Base, AZ: Air Force Human Resources Laboratory.
- Tinsley, C. H., Dillon, R. L., & Madsen, P. M. (2011). How to avoid catastrophe. *Harvard Business Review*, 89(4), 90-97.
- Transport Safety Board of Canada [TSB]. (1990). *Report of a safety study on VFR flight into adverse weather*. (TSB Report No. 90-SP002). Gatineau, Canada: Transportation Safety Board of Canada.

REFERENCES

- Transport Safety Board of Canada [TSB]. (1996). *VFR INTO IMC – Controlled Flight into Terrain. Western Straits Air De Havilland DHC-3 (Turbine) Otter C-FEBX, Campbell River, British Columbia, 27th September 1995*. (Aviation Investigation A95H0012). Gatineau, Canada: Minister of Public Works and Government Services Canada.
- Transport Safety Board of Canada [TSB]. (2007). *Runway overrun and fire Air France Airbus A340-313 F-GLZQ, Toronto/Lester B. Pearson International Airport, Ontario, 02 August 2005* (Aviation Investigation A05H0002). Gatineau, Canada: Minister of Public Works and Government Services Canada.
- Transport Safety Board of Canada [TSB]. (2010a). *Controlled Flight into Terrain (CFIT), Pacific Coastal Airlines, Grumman G-21A C-FPCK, South Thormanby Island, British Columbia, 16th November 2008*. (Aviation Investigation A08P0353). Gatineau, Canada: Minister of Public Works and Government Services Canada.
- Transport Safety Board of Canada [TSB]. (2010b). *Loss of Control and Collision with Terrain, Cessna 150J, C-GPXN, Courtland, Ontario, 12th August 2009*. (Aviation Investigation A09O0171). Gatineau, Canada: Minister of Public Works and Government Services Canada.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453-458.

REFERENCES

- Vashitz, G., Meyer, J., Parmet, Y., Peleg, R., Goldfarb, D., Porath, A., & Gilutz, H. (2009). Defining and measuring physicians response to clinical reminders. *Journal of Biomedical Informatics*, 42(2), 317-326.
- Vidulich, M. A., Wickens, C. D., Tsang, P. D., & Flach, J. M. (2010). Information Processing in Aviation. In E. Salas, & D. Maurino (Eds.). *Human Factors in Aviation* (pp175-213), Elsevier.
- Wagtendonk, W. J. (2011). *Meteorology for professional pilots* (4th Ed). Aviation Theory Centre (NZ) Ltd trading as Pilot Books.
- Wason, P. C. (1960). On the failure to eliminate hypothesis in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12(3), 129-140.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed), *New horizons in psychology* (pp. 135-151). Harmondsworth: Penguin.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39(5), 806-820.
- Weinstein, N. D., & Klein, W. M. (1995). Resistance of personal risk perception to debiasing interventions. *Health Psychology*, 14(2), 132-140.
- Wickens, C., & Flach, J. (1988). Information processing. In E. Wiener & D. Nagel (Eds.), *Human factors in aviation*. New York, NY: Academic.

REFERENCES

- Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*. (3rd ed.), Prentice Hall.
- Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2013). *Engineering Psychology and Human Performance*. (4th ed.), Pearson Education.
- Wickens, C. D., Stokes, A., Barnett, B., & Hyman, F. (1993). The effects of stress on pilot judgment in a MIDIS simulator. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision-making* (pp. 271-292). New York: Plenum Press.
- Wiegmann, D. A., Goh, J., & O'Hare, D. (2002). The role of situation assessment and flight experience in pilots' decision to continue visual rules flight into adverse weather. *Human Factors*, *44*(2), 189-197.
- Wiggins, M. (2007). Perceived GPS use and weather-related decision strategies among pilots. *The International Journal of Aviation Psychology*, *17*(3), 249-264.
- Wiggins, M., Azar, D., Hawken, J., Loveday., & Newman, D. (2014). *Cue-utilisation typologies and pilots' pre-flight and in-flight weather decision-making*. *Safety Science*, *65*, 118-124.
- Wiggins, M. W., Hunter, D. R., O'Hare, D., & Martinussen, M. (2012). Characteristics of pilots who report deliberate versus inadvertent visual flight into Instrument Meteorological Conditions. *Safety Science*, *50*(3), 472-477.

REFERENCES

- Wiggins, M., Martinussen, M., & Hunter, D. R. (1999). Three continents but a single mind when it comes to weather. *Proceedings of the 10th International Symposium on Aviation Psychology*. Columbus: Ohio State University.
- Wiggins, M. W., & O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: A cross sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied*, 1(14), 304-319.
- Wiggins, M. W., & O'Hare, D. (2003a). Expert and novice pilot perceptions of static in-flight images of weather. *International Journal of Aviation Psychology*, 13(2), 173-187.
- Williams, K. W., Yost, A., Holland, J., & Tyler, R. R. (2002). *Assessment of advanced cockpit displays for general aviation aircraft: The capstone program*. (Tech. Rep. No. NTIS DOT/FAA/AM-02/21). Washington, DC: Federal Aviation Administration.
- Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, 116(1), 117-142.
- Wilson, D. R., & Fallshore, M. (2001). Optimistic and ability biases in pilots' decisions and perceptions of risk regarding VFR flight into IMC. *Proceedings of the 11th International Symposium on Aviation Psychology*, Columbus, OH.
- Wright, W. F., & Anderson, U. (1989). Effects of situation familiarity and financial incentives on use of the anchoring and adjustment heuristic for probability assessment. *Organizational Behaviour and Human Decision Processes*, 44(1), 68-82.

APPENDICES

Appendix A – Sample of Study 1 Questionnaire

Section ONE: INTRODUCTION

I would like to invite you to participate in an exploratory study regarding weather-related decision making. This study is being carried out by Stephen Walmsley (Post-graduate student at Massey University).

Please try to answer as genuinely and as honestly as possible.

It should take you approximately 10 minutes to complete the questionnaire. The data you supply will be anonymous.

Completion of the questionnaire implies consent to participate.

This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher is responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher, please contact Professor John O'Neill, Director (Research Ethics), telephone 06 350 5249, email humanethics@massey.ac.nz.

Next

0%

Section TWO (1 of 5)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind at 270°T at 5knot, Visibility is 27km, Cloud broken at 4000 feet, Temp. 9°C, Dew Point 6°C, QNH 1012

1. The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



The image may take a few seconds to load

2. What is your best guess of the Cloud Height, in feet above ground level?

3. What is your best guess of the Horizontal Visibility (in km)?

4. How confident are you of these estimates?

	1	2	3	4	5	6	7	8	9	
Not Confident	<input type="radio"/>	Very Confident								

Indicate your level of agreement in regards to flying towards the destination;

5. It is safe to continue towards your destination;

	1	2	3	4	5	6	7	8	9	
Strongly Disagree	<input type="radio"/>	Strongly Agree								

Back		Next
14%		

Section TWO (2 of 5)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind at 030°T at 3knot, Visibility is 8kmm, Cloud broken at 1300 feet, Temp. 22°C, Dew Point 16°C, QNH 1002

6. The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



7. What is your best guess of the Cloud Height, in feet above ground level?

8. What is your best guess of the Horizontal Visibility (in km)?

9. How confident are you of these estimates?

	1	2	3	4	5	6	7	8	9	
Not Confident	<input checked="" type="radio"/>	<input type="radio"/>	Very Confident							

Indicate your level of agreement in regards to flying towards your destination;

10. It is safe to continue towards your destination;

	1	2	3	4	5	6	7	8	9	
Strongly Disagree	<input type="radio"/>	Strongly Agree								

Back		Next	
29%			

Section TWO (3 of 5)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind at 080°T at 8knot, Visibility is 26km, Cloud broken at 3900 feet, Temp. 19°C, Dew Point 6°C, QNH 1004

11. The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



12. What is your best guess of the Cloud Height, in feet above ground level?

13. What is your best guess of the Horizontal Visibility (in km)?

14. How confident are you of these estimates?

	1	2	3	4	5	6	7	8	9	
Not Confident	<input type="radio"/>	Very Confident								

Indicate your level of agreement in regards to flying towards your destination;

15. It is safe to continue towards your destination;

	1	2	3	4	5	6	7	8	9	
Strongly Disagree	<input type="radio"/>	Strongly Agree								

Back		Next	
43%			

Section TWO (4 of 5)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind at 180°T at 6knots, Visibility 7km, Cloud broken at 1200 feet, Temp. 12°C, Dew Point 6°C, QNH 1001

16. The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



17. What is your best guess of the Cloud Height, in feet above ground level?

18. What is your best guess of the Horizontal Visibility (in km)?

19. How confident are you of these estimates?

	1	2	3	4	5	6	7	8	9	
Not Confident	<input type="radio"/>	Very Confident								

Indicate your level of agreement in regards to flying towards your destination;

20. It is safe to continue towards your destination;

	1	2	3	4	5	6	7	8	9	
Strongly Disagree	<input type="radio"/>	Strongly Agree								

Back		Next
57%		

Section TWO (5 of 5)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind at 210°T at 9knot, Visibility is 23km, Cloud broken at 3600 feet, Temp. 18°C, Dew Point 15°C, QNH 1005

21. The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



22. What is your best guess of the Cloud Height, in feet above ground level?

23. What is your best guess of the Horizontal Visibility (in km)?

24. How confident are you of these estimates?

	1	2	3	4	5	6	7	8	9	
Not Confident	<input type="radio"/>	Very Confident								

Indicate your level of agreement in regards to flying towards your destination;

25. It is safe to continue towards your destination;

	1	2	3	4	5	6	7	8	9	
Strongly Disagree	<input type="radio"/>	Strongly Agree								

Back		Next
71%		

Section THREE: DEMOGRAPHIC

26. Indicate on the scale below how realistic you found the material in the three scenario's in this survey;

	1	2	3	4	5	6	7	8	9	
Not at all realistic	<input type="radio"/>	Absolutely Realistic								

Total Flying Hours (approx):

What is your age:

Type of Licence you Hold;

- Student Pilot
- Private Pilot Licence (PPL)
- Commercial Pilot Licence (CPL)
- Air Transport Pilot Licence (ATPL)

		Back	Submit		
86%					

Appendix B – Low Risk Ethics Notification for Study 1



MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA

7 February 2013

Stephen Walmsley
96 King Street
CAMBRIDGE 3434

Dear Stephen

Re: Anchoring and Adjustment in VFR-INC Accidents

Thank you for your Low Risk Notification which was received on 13 December 2012.

Your project has been recorded on the Low Risk Database which is reported in the Annual Report of the Massey University Human Ethics Committees.

The low risk notification for this project is valid for a maximum of three years.

Please notify me if situations subsequently occur which cause you to reconsider your initial ethical analysis that it is safe to proceed without approval by one of the University's Human Ethics Committees.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

A reminder to include the following statement on all public documents:

"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director (Research Ethics), telephone 06 350 5249, e-mail humanethics@massey.ac.nz".

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to provide a full application to one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely

John G O'Neill (Professor)
**Chair, Human Ethics Chairs' Committee and
Director (Research Ethics)**

cc Dr Andrew Gilbey
School of Aviation
PN833

Mr Ashok Poduval, CEO
School of Aviation
PN833

Massey University Human Ethics Committee
Accredited by the Health Research Council

Research Ethics Office

Massey University, Private Bag 11222, Palmerston North 4442, New Zealand T +64 6 350 5573 +64 6 350 5575 F +64 6 350 5822
E humanethics@massey.ac.nz animalethics@massey.ac.nz gti@massey.ac.nz www.massey.ac.nz

Appendix C – Sample of Study 2 Questionnaire

Section ONE: Introduction

I would like to invite you to participate in an exploratory study looking at how pilots make weather-related decisions. This study is being carried out by Stephen Walmsley (Post-graduate student at Massey University).

The study consist of answering a question about **5 in-flight images**. Please try to answer as genuinely and as honestly as possible.

It should take you approximately **5 to 10 minutes** to complete the questionnaire. The data you supply will be anonymous.

Completion of the questionnaire implies consent to participate.

This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher is responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher, please contact either Dr Andrew Gilbey (Study Supervisor), telephone 06 356 9099 Ext 84212, email A.P.Gilbey@massey.ac.nz or Professor John O'Neill, Director (Research Ethics), telephone 06 350 5249, email humanethics@massey.ac.nz.

Next

0%

Section TWO (1 of 5)

Imagine you are flying a Cessna 172 on a VFR cross country, your destination aerodrome is just beyond the red box. As you approach the aerodrome your passenger begins to feel air sick and wishes to land as soon as possible. To avoid too much discomfort for your passenger it is important that you proceed towards the aerodrome if the conditions permit, the nearest suitable aerodrome is an hour away.

Listed below are three statements about the weather conditions in front of the aircraft. Pick *one* that you feel is most useful in deciding if you can continue towards this aerodrome?

- Another aircraft ahead safely flew in my direction 10min earlier
- Almost certainly I will end up having to fly through cloud
- I can see through the rain shower ahead

Back

Next

14%

Section TWO (2 of 5)

Imagine you are flying a Cessna 172 on a VFR cross country flight, flying in the direction of the red box. You are running late for a meeting which is being held near your destination aerodrome, which is just beyond the red box. There are no other aerodromes in the immediate area. The meeting is very important therefore it would be ideal if you could proceed towards your destination.

Listed below are three statements about the weather conditions in front of the aircraft. Pick *one* that you feel is most useful in deciding if you can continue towards this aerodrome?

- You have just received the Met report (ATIS) for your aerodrome which indicates light winds
- The cloud base is remaining above my current level
- The dark cloud above my flight path may produce a heavy rain shower at any time

Back Next

29%

Section TWO (3 of 5)

Imagine you are flying a Cessna 172 on a VFR cross country. Half way through the flight the aircraft engine begins to make a strange noise, you decide it would be wise to divert to a nearby aerodrome to get it checked out. The aerodrome is just beyond the red box.

Listed below are three statements about the weather conditions in front of the aircraft. Pick *one* that you feel is most useful in deciding if you can continue towards this aerodrome?

- There is a clear patch towards the aerodrome I am heading towards
- The cloud base is too close to the ground
- The heavy showers are remaining to the right of track

Back Next

43%

Section TWO (4 of 5)

Imagine you are flying a Cessna 172 on a local VFR flight. You have hired the aircraft from the local aeroclub who have just radioed you to say they need the aircraft back ASAP. The aerodrome which you need to return the aircraft to is located just beyond the red box. It is important that you return ASAP if the weather conditions permit to ensure you can hire an aircraft in the future.

Listed below are three statements about the weather conditions in front of the aircraft. Pick *one* that you feel is most useful in deciding if you can continue towards this aerodrome?

- I can see the horizon through the rain shower
- The cloud base inside the rain shower may be lower than my current flight level
- Another aircraft ahead has reported the cloud height is above circuit altitude

Back

Next

57%

Section TWO (5 of 5)

Imagine you are flying a Cessna 172 on a VFR cross country flight, your destination is just beyond the hills towards the red box. The aerodrome is due to close shortly for the rest of the day to allow for maintenance work to be carried out on the runway, however you have time to complete a normal approach. It is important that you proceed towards the aerodrome if the weather permits as you would have to fly another 2 hours to a suitable alternative.

Listed below are three statements about the weather conditions in front of the aircraft. Pick *one* that you feel is most useful in deciding if you can continue towards this aerodrome?

- An aircraft ahead has reported light turbulence flying over the hills
- The visibility is very good towards my destination
- The cloud is lowering towards the hill tops to the left of track

[Back](#)[Next](#)

71%

Section THREE: Demographic

1. In comparison to other pilots with similar flight background and experience as yourself, what do you feel are your chances of experiencing an accident due to inadvertent flight into instrument meteorological conditions (i.e., cloud or fog)?

	1	2	3	4	5	6	7	8	9	
Less Than Average	<input type="radio"/>	More Than Average								

2. Have you flown into areas of instrument meteorological conditions (IMC), when you were not on an instrument flight plan (VFR flight into IMC)?

- Yes
- No

3. What is your Total Flying Hours (Approx) ?

4. What type of Licence do you hold?

- Student
- PPL
- CPL
- ATPL

5. Which of the following regarding an Instrument Rating applies to you?

- Instrument Rated & Current
- Instrument Rated but not Current
- Never Held an instrument Rating

6. What is your AGE?

7. Indicate on the scale below how realistic you found the material in the five scenarios;

	1	2	3	4	5	6	7	8	9	
NOT at all Realistic	<input type="radio"/>	VERY Realistic								

Back

Submit

Appendix D – Low-Risk Ethics Notification for Study 2



MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA

3 September 2013

Stephen Walmsley
96 King Street
CAMBRIDGE 3434

Dear Stephen

Re: Confirmation Bias in Weather-Related Decision-Making

Thank you for your Low Risk Notification which was received on 9 August 2013.

Your project has been recorded on the Low Risk Database which is reported in the Annual Report of the Massey University Human Ethics Committees.

The low risk notification for this project is valid for a maximum of three years.

Please notify me if situations subsequently occur which cause you to reconsider your initial ethical analysis that it is safe to proceed without approval by one of the University's Human Ethics Committees.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

A reminder to include the following statement on all public documents:

"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director (Research Ethics), telephone 06 350 5249, e-mail humanethics@massey.ac.nz".

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to provide a full application to one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely

A handwritten signature in black ink that reads "J. O'Neill".

John G O'Neill (Professor)
**Chair, Human Ethics Chairs' Committee and
Director (Research Ethics)**

cc Dr Andrew Gilbey
School of Aviation
PN833

Mr Ashok Poduval, CEO
School of Aviation
PN833

**Massey University Human Ethics Committee
Accredited by the Health Research Council**

Research Ethics Office, Research and Enterprise

Massey University, Private Bag 11222, Palmerston North 4442, New Zealand T 06 3505573; 06 3505575 F 06 350 5622
E humanethics@massey.ac.nz; animalethics@massey.ac.nz; gtc@massey.ac.nz www.massey.ac.nz

Appendix E – Sample of Study 3 Questionnaire

Section ONE: Introduction

I would like to invite you to participate in an exploratory study looking at how pilots make weather-related decisions. This study is being carried out by Stephen Walmsley (who is a Post-graduate student at Massey University and also a flight instructor).

The study consists of evaluating **5 VFR Flights** in marginal weather. Please try to answer as genuinely and as honestly as possible.

It should take you approximately **5 to 10 minutes** to complete the questionnaire. The data you supply will be anonymous.

Completion of the questionnaire implies consent to participate.

This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher is responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher, please contact either Dr Andrew Gilbey (Study Supervisor), telephone 06 356 9099 Ext 84212, email A.P.Gilbey@massey.ac.nz or Professor John O'Neill, Director (Research Ethics), telephone 06 350 5249, email humanethics@massey.ac.nz.

Next

0%

Section TWO: (1 of 5)

Please try to imagine the following situation. A pilot is conducting a VFR cross country flight in a Cessna 172. The pilot has flown the same route several times before, which is over flat terrain. The flight should take about an hour, conducted in daylight hours. The weather information and route map available to the pilot is as follows;

Current Weather

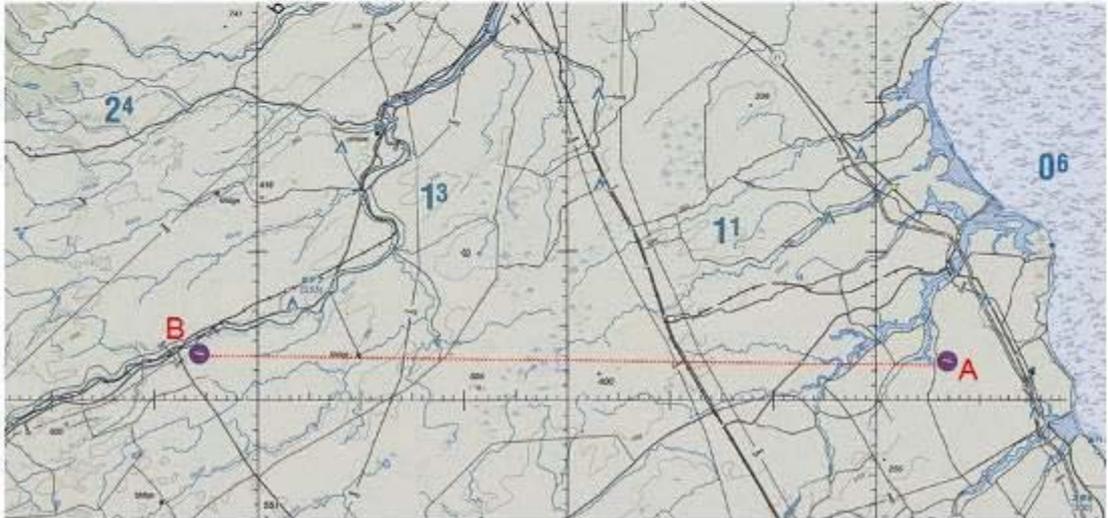
Departure (A) : Wind Variable 3knots, Visibility 16km, Light Rain with Cloud base at 1800ft.

Destination (B) : Wind Variable 2knots, Visibility 16km, Light Drizzle, Cloud base 1500ft.

Area Forecast

Visibility 16km with scattered light drizzle.
 Cloud broken to overcast at 1,500 feet, with cloud tops 24,000ft.
 Turbulence NIL.

All cloud heights in AGL (above ground level). Pilot traveling from Aerodrome A to Aerodrome B



The pilot decides to **go ahead** with the VFR flight.
 You later find out the flight safely landed at its intended destination.

Evaluate the pilots decision to go ahead with the flight;

1. Considering the information available to the pilot, how do you rate the pilots decision to conduct the VFR flight;

	1	2	3	4	5	6	7	8	9	
Very GOOD Decision	<input type="radio"/>	Very POOR Decision								

2. How much risk do you think the pilot took in attempting this VFR flight?

	1	2	3	4	5	6	7	8	9	
Very LOW Risk	<input type="radio"/>	Very HIGH Risk								

3. You wish to conduct the same flight. Considering the information available to the pilot before they departed, indicate your level of agreement with the following statement;

I would be able to **safely** conduct the flight VFR;

	1	2	3	4	5	6	7	8	9	
Strongly Agree	<input type="radio"/>	Strongly Disagree								

				Back	Next					
14%										

Section TWO: (2 of 5)

Please try to imagine the following situation. A pilot is conducting a local flight in a Cessna 172. The pilot is familiar with the area, having completed most of his basic flight training in the area. The local terrain is mostly flat, and the flight is to be conducted in daylight hours. The pilot intends to fly for about 1 hour. The weather information and local area map available to the pilot is as follows;

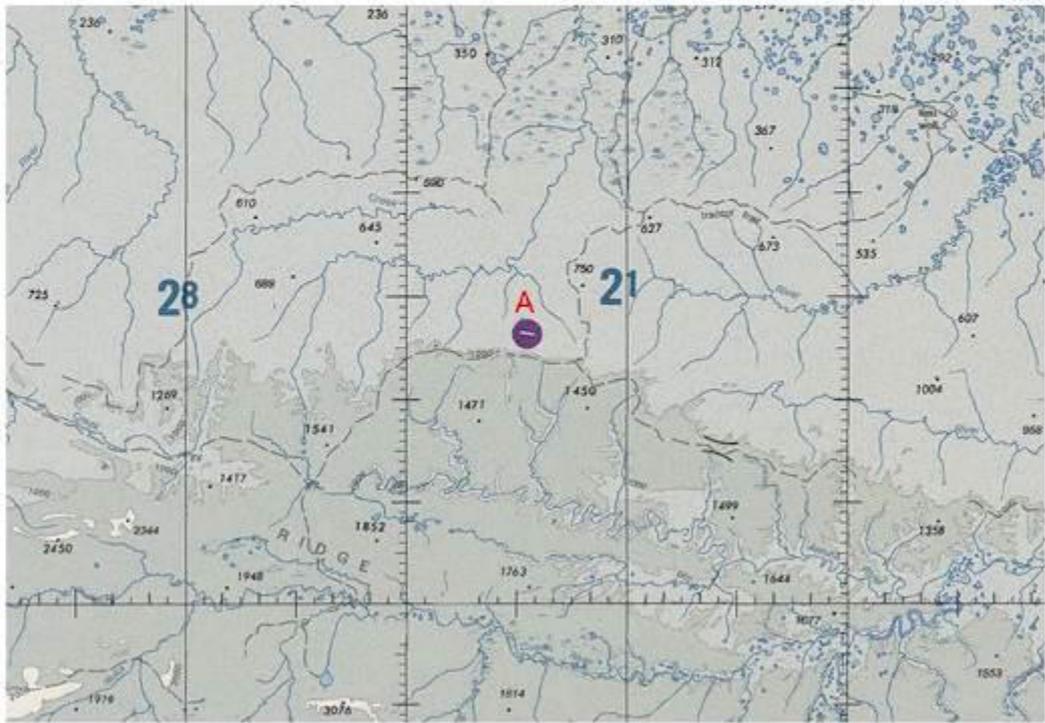
Current Weather

Departure (A): Wind 280 at 3knots, Visibility 10km, Cloud Scattered at 1000ft, Broken at 1600ft, Overcast 3300ft.

Forecast

No Significant Changes to current weather expected in next two hours.

All cloud heights in AGL (above ground level). Pilot flying from Aerodrome A



The pilot decides to **go ahead** with the VFR flight.

You later find out the pilot was able to safely complete the local flight.

Evaluate the pilots decision to go ahead with the flight;

4. Considering the information available to the pilot, how do you rate the pilots decision to conduct the VFR flight;

	1	2	3	4	5	6	7	8	9	
Very GOOD Decision	<input type="radio"/>	Very POOR Decision								

5. How much risk do you think the pilot took in attempting this VFR flight?

	1	2	3	4	5	6	7	8	9	
Very LOW Risk	<input type="radio"/>	Very HIGH Risk								

6. You wish to conduct the same flight. Considering the information available to the pilot before they departed, indicate your level of agreement with the following statement;

I would be able to **safely** conduct the flight VFR;

	1	2	3	4	5	6	7	8	9	
Strongly Agree	<input type="radio"/>	Strongly Disagree								

				Back	Next					
29%										

Section TWO: (3 of 5)

Please try to imagine the following situation. A pilot is conducting a local flight in a Cessna 172 in which he intends to take some aerial photos of a friend's house. The friend's house is about 30min flight time from the departure aerodrome. The pilot has plenty of experience of flying in this area. The weather information and local area map available to the pilot is as follows;

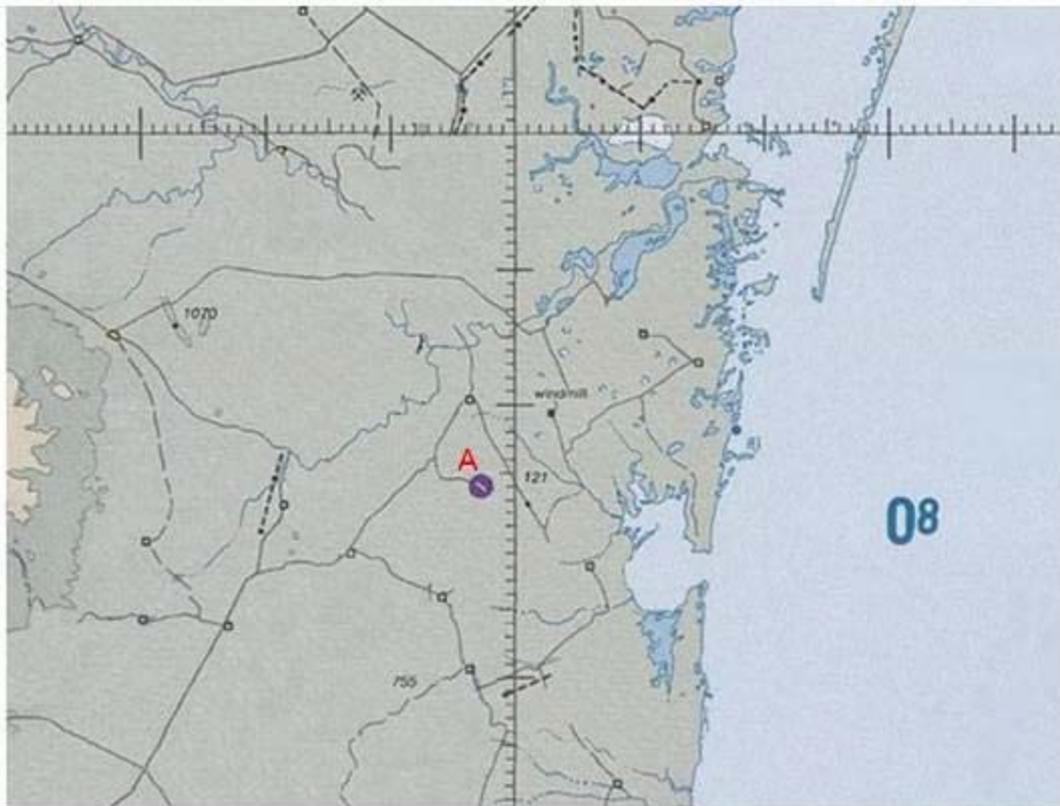
Current Weather

Departure (A) : Wind 150 at 19knots, Visibility 5km in light Rain, Cloud Scatters at 2200ft, Overcast 6500ft.

Forecast

No Significant Changes to current weather expected in next two hours.

All cloud heights in AGL (above ground level). Pilot flying from Aerodrome A



The pilot decides to **go ahead** with the VFR flight. You later find out the pilot was able to safely complete the local flight.

Evaluate the pilots decision to go ahead with the flight;

7. Considering the information available to the pilot, how do you rate the pilots decision to conduct the VFR flight;

	1	2	3	4	5	6	7	8	9	
Very GOOD Decision	<input type="radio"/>	Very POOR Decision								

8. How much risk do you think the pilot took in attempting this VFR flight?

	1	2	3	4	5	6	7	8	9	
Very LOW Risk	<input type="radio"/>	Very HIGH Risk								

9. You wish to conduct the same flight. Considering the information available to the pilot before they departed, indicate your level of agreement with the following statement;

I would be able to **safely** conduct the flight VFR;

	1	2	3	4	5	6	7	8	9	
Strongly Agree	<input type="radio"/>	Strongly Disagree								

				Back	Next					
43%										

Section TWO: (4 of 5)

Please try to imagine the following situation. A pilot is conducting a VFR cross country flight in a Cessna 172. The pilot has been at a business meeting all morning and is now returning home. The pilot flew in yesterday, and the return leg should take approximately 60min. The weather information and route map available to the pilot is as follows;

Current Weather

Departure (A) : Wind 190 at 8knots, Visibility 30km, Sky Clear.

Destination (B) : Wind 240 at 8knots, Visibility 5km in Patch's of Mist, Cloud base 700ft.

Area Forecast

Visibility reducing to 5km in area of Mist and Fog.

Areas of Cloud 1000ft with tops 15,000ft.

Turbulence NIL.

All cloud heights in AGL (above ground level). Pilot traveling from Aerodrome A to Aerodrome B



The pilot decides to **go ahead** with the VFR flight.

You later find out the flight safely landed at its intended destination.

Evaluate the pilots decision to go ahead with the flight;

10. Considering the information available to the pilot, how do you rate the pilots decision to conduct the VFR flight;

	1	2	3	4	5	6	7	8	9	
Very GOOD Decision	<input type="radio"/>	Very POOR Decision								

11. How much risk do you think the pilot took in attempting this VFR flight?

	1	2	3	4	5	6	7	8	9	
Very LOW Risk	<input type="radio"/>	Very HIGH Risk								

12. You wish to conduct the same flight. Considering the information available to the pilot before they departed, indicate your level of agreement with the following statement;

I would be able to **safely** conduct the flight VFR;

	1	2	3	4	5	6	7	8	9	
Strongly Agree	<input type="radio"/>	Strongly Disagree								

					Back	Next				
57%										

Section TWO: (5 of 5)

Please try to imagine the following situation. A non-instrument rated pilot is conducting a VFR cross country flight in a Cessna 172. The pilot is returning to his home base which he flew from yesterday. The return flight should take just over an hour. The weather information and route map available to the pilot is as follows;

Current Weather

Departure (A) : Wind 340 at 4knots, Visibility 16km, Sky Clear.

Destination (B) : Wind 310 at 9knots, Visibility 16km in Light Rain, Cloud base 800ft.

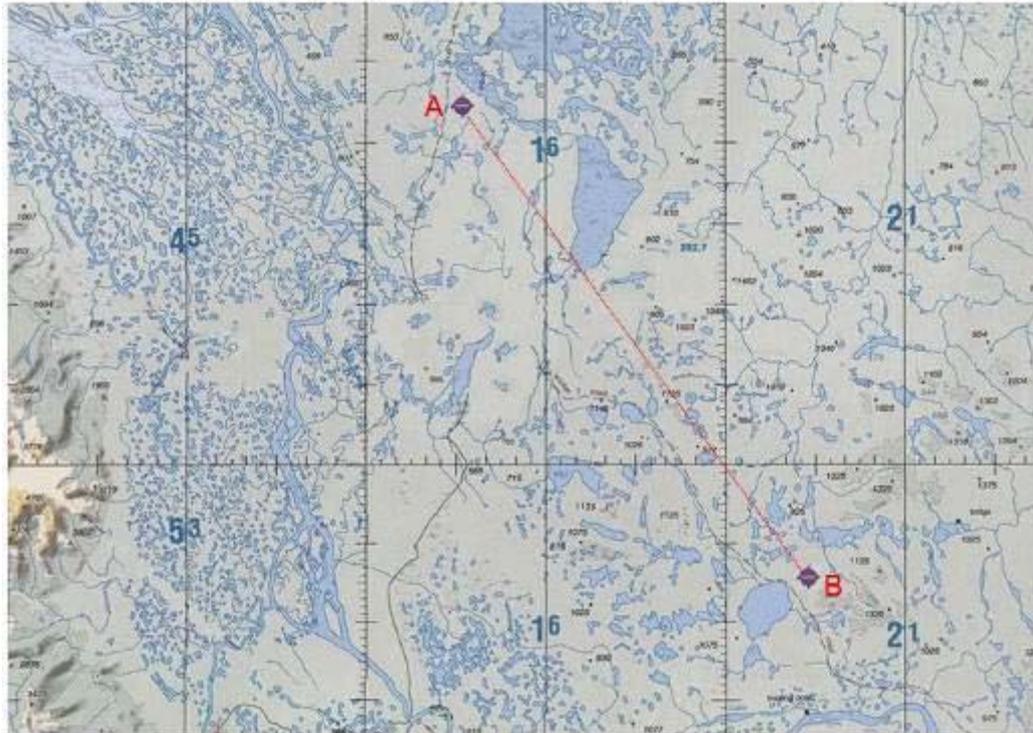
Area Forecast

Visibility reducing 10km in areas of light rain and Mist.

Cloud cover at 4000ft, with top of 20,000ft.

Turbulence NIL.

All cloud heights in AGL (above ground level). Pilot traveling from Aerodrome A to Aerodrome B



The pilot decides to **go ahead** with the VFR flight. You later find out the flight safely landed at its intended destination. Evaluate the pilots decision to go ahead with the flight;

13. Considering the information available to the pilot, how do you rate the pilots decision to conduct the VFR flight;

	1	2	3	4	5	6	7	8	9	
Very GOOD Decision	<input type="radio"/>	Very POOR Decision								

14. How much risk do you think the pilot took in attempting this VFR flight?

	1	2	3	4	5	6	7	8	9	
Very LOW Risk	<input type="radio"/>	Very HIGH Risk								

15. You wish to conduct the same flight. Considering the information available to the pilot before they departed, indicate your level of agreement with the following statement;

I would be able to **safely** conduct the flight VFR;

	1	2	3	4	5	6	7	8	9	
Strongly Agree	<input type="radio"/>	Strongly Disagree								

				Back	Next					
71%										

Section THREE: Demographic

16. What is your Total Flying Hours (Approx) ?

17. Which of the following regarding an Instrument Rating applies to you?

- Instrument Rated & Current
- Instrument Rated but not Current
- Never Held an instrument Rating

18. What type of Licence do you hold?

- Student
- PPL
- CPL
- ATPL

19. What is your AGE?

Back

Submit

86%

Appendix F – Low-Risk Ethics Notification for Study 3



MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA

28 March 2014

Stephen Walmsley
96 King Street
CAMBRIDGE 3434

Dear Stephen

Re: Outcome Bias in Weather-Related Decision-Making

Thank you for your Low Risk Notification which was received on 4 March 2014.

Your project has been recorded on the Low Risk Database which is reported in the Annual Report of the Massey University Human Ethics Committees.

You are reminded that staff researchers and supervisors are fully responsible for ensuring that the information in the low risk notification has met the requirements and guidelines for submission of a low risk notification.

The low risk notification for this project is valid for a maximum of three years.

Please notify me if situations subsequently occur which cause you to reconsider your initial ethical analysis that it is safe to proceed without approval by one of the University's Human Ethics Committees.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

A reminder to include the following statement on all public documents:

"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director (Research Ethics), telephone 06 350 5249, e-mail humanethics@massey.ac.nz".

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to provide a full application to one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely

A handwritten signature in black ink that reads "J. O'Neill".

John G O'Neill (Professor)
Chair, Human Ethics Chairs' Committee and
Director (Research Ethics)

cc Dr Andrew Gilbey
School of Aviation
PN833

Mr Ashok Poduval, CEO
School of Aviation
PN833

Massey University Human Ethics Committee
Accredited by the Health Research Council

Research Ethics Office

Massey University, Private Bag 11222, Palmerston North 4442, New Zealand T +64 6 350 5573 +64 6 350 5575 F +64 6 350 5622
E humanethics@massey.ac.nz animaethics@massey.ac.nz gtc@massey.ac.nz www.massey.ac.nz

Appendix G – Sample of Study 4–6 Questionnaire

Weather-Related Decision Making Study

***INFORMATION SHEET* [#1 – debiasing group]**

I would like to invite you to participate in an exploratory study looking at how pilots make weather-related decisions.

This study is being carried out by Stephen Walmsley (who is a Post-graduate student at Massey University and also a flight instructor), exploring weather-related decision making.

You have been invited to participate in this study as you are in the target audience that this project is exploring; student pilots currently conducting their ground school phase of their flight training. No discomfort or risk will result from participation.

Project Procedures

You will be provided with a **short lecture** (5-10min) which intends to educate you in an area which can cause decision error when evaluating the weather (cognitive bias) along with a simple method to avoid this type of decision error.

You will then evaluate **10 VFR Flights** in marginal weather using a paper based questionnaire. It should take about 20-25 minutes to complete the questionnaire.

As you are also a student of the researcher, it is important that if you feel pressured to participate that you do not conduct this study. Non-participation will have no influence on your studies.

Data Management

The data you supply will be anonymous. To protect your identity no demographic information will be collected. The data will be securely held by the researcher.

Completion and return of the questionnaire implies consent. You have the right to decline to answer any particular question.

If you would like to receive a summary of the studying findings once they have been completed, please provide your email: _____
(please note your email will only be used to send summary of findings).

Project Contacts

Please feel free to contact the researcher (Stephen Walmsley). Telephone; 07 843 3304 ext 839 email; walmsleystephen@gmail.com

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher, please contact Dr Andrew Gilbey (Study Supervisor), Telephone; 06 356 9099 Ext: 84212. Email A.P.Gilbey@massey.ac.nz

This project has been reviewed and approved by the Massey University Human Ethics Committee: Southern B, Application 14/43. If you have any concerns about the conduct of this research, please contact Prof John O'Neill, Acting Chair, Massey University Human Ethics Committee: Southern B, telephone 06 350 5799 x 81090, email humanethicsouthb@massey.ac.nz.

Weather-Related Decision Making Study
***INFORMATION SHEET* [#2 – control group]**

I would like to invite you to participate in an exploratory study looking at how pilots make weather-related decisions.

This study is being carried out by Stephen Walmsley (who is a Post-graduate student at Massey University and also a flight instructor), exploring weather-related decision making.

You have been invited to participate in this study as you are in the target audience that this project is exploring; student pilots currently conducting their ground school phase of their flight training. No discomfort or risk will result from participation.

Project Procedures

You will be evaluating **10 VFR Flights** in marginal weather using a paper based questionnaire. It should take about 20-25 minutes to complete the questionnaire.

As you are also a student of the researcher, it is important that if you feel pressured to participate that you do not conduct this study. Non-participation will have no influence on your studies.

Data Management

The data you supply will be anonymous. To protect your identity no demographic information will be collected.

The data will be securely held by the researcher.

Completion and return of the questionnaire implies consent. You have the right to decline to answer any particular question.

If you would like to receive a summary of the studying findings once they have been completed, please provide your email: _____
(please note your email will only be used to send summary of findings).

Project Contacts

Please feel free to contact the researcher (Stephen Walmsley), telephone, 07 843 3304 ext 839 email; walmsleystephen@gmail.com

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher, please contact either Dr Andrew Gilbey (Study Supervisor), telephone 06 356 9099 Ext 84212, email A.P.Gilbey@massey.ac.nz

This project has been reviewed and approved by the Massey University Human Ethics Committee: Southern B, Application 14/43. If you have any concerns about the conduct of this research, please contact Prof John O'Neill, Acting Chair, Massey University Human Ethics Committee: Southern B, telephone 06 350 5799 x 81090, email humanethicsouthb@massey.ac.nz.

Section ONE (1 of 4)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind 270°T at 5kt, Visibility 25km, Cloud broken at 3800', Temp. 9°C, DP 6°C, QNH 1012.

The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



What is your best guess of the Cloud Height, in ft above ground level?

What is your best guess of the Horizontal Visibility (in km)?

Section ONE (2 of 4)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind at 030°T at 3kt, Visibility 8km, Cloud broken at 1300', Temp 22°C, DP 16°C, QNH 1002

The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



What is your best guess of the Cloud Height, in ft above ground level?

What is your best guess of the Horizontal Visibility (in km)?

Section ONE (3 of 4)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind at 080°T at 8kt, Visibility is 26km, Cloud broken at 3900', Temp 19°C, DP 6°C, QNH 1004

The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



What is your best guess of the Cloud Height, in ft above ground level?

What is your best guess of the Horizontal Visibility (in km)?

Section ONE (4 of 4)

Imagine that you are flying a Cessna 172 on a VFR cross country flight. You are approaching your destination with an elevation close to sea level. The aerodrome's current weather is not available.

You have obtained the weather forecast prior to the flight which indicates the weather at your destination will be;

Wind at 180°T at 6kt, Visibility 7km, Cloud broken at 1200', Temp 12°C, DP 6°C, QNH 1001

The in-flight image below is the cockpit view in the direction of your destination. Is the weather forecast better or worse than the actual weather conditions?

- Weather forecast BETTER than actual conditions
- SAME
- Weather forecast WORSE than actual conditions



What is your best guess of the Cloud Height, in ft above ground level?

What is your best guess of the Horizontal Visibility (in km)?

Section TWO (1 of 3)

Imagine you are flying a Cessna 172 on a VFR cross country, your destination aerodrome is just beyond the red box. As you approach the aerodrome your passenger begins to feel air sick and wishes to land as soon as possible. To avoid too much discomfort for your passenger it is important that you proceed towards the aerodrome if the conditions permit, the nearest suitable aerodrome is an hour away.

Listed below are three statements about the weather conditions in front of the aircraft. Pick *one* that you feel is most useful in deciding if you can continue towards this aerodrome?

- Another aircraft ahead safely flew in my direction 10min earlier
- Almost certainly I will end up having to fly through cloud
- I can see through the rain shower ahead

Section THREE (2 of 3)

Imagine you are flying a Cessna 172 on a local VFR flight. You have hired the aircraft from the local aeroclub who have just radioed you to say they need the aircraft back ASAP. The aerodrome which you need to return the aircraft to is located just beyond the red box. It is important that you return ASAP if the weather conditions permit to ensure you can hire an aircraft in the future.

Listed below are three statements about the weather conditions in front of the aircraft. Pick *one* that you feel is most useful in deciding if you can continue towards this aerodrome?

- I can see the horizon through the rain shower
- The cloud base inside the rain shower may be lower than my current flight level
- Another aircraft ahead has reported the cloud height is above circuit altitude

Section TWO (3 of 3)



Imagine you are flying a Cessna 172 on a VFR cross country. Half way through the flight the aircraft engine begins to make a strange noise, you decide it would be wise to divert to a nearby aerodrome to get it checked out. The aerodrome is just beyond the red box.

Listed below are three statements about the weather conditions in front of the aircraft. Pick *one* that you feel is most useful in deciding if you can continue towards this aerodrome?

- There is a clear patch towards the aerodrome I am heading towards
- The cloud base is too close to the ground
- The heavy showers are remaining to the right of track

Section THREE: (1 of 3)

Please try to imagine the following situation. A pilot is conducting a VFR cross country flight in a Cessna 172. The pilot has flown the same route several times before, which is over flat terrain. The flight should take about an hour, conducted in daylight hours. The weather information and route map available to the pilot is as follows;

Current Weather

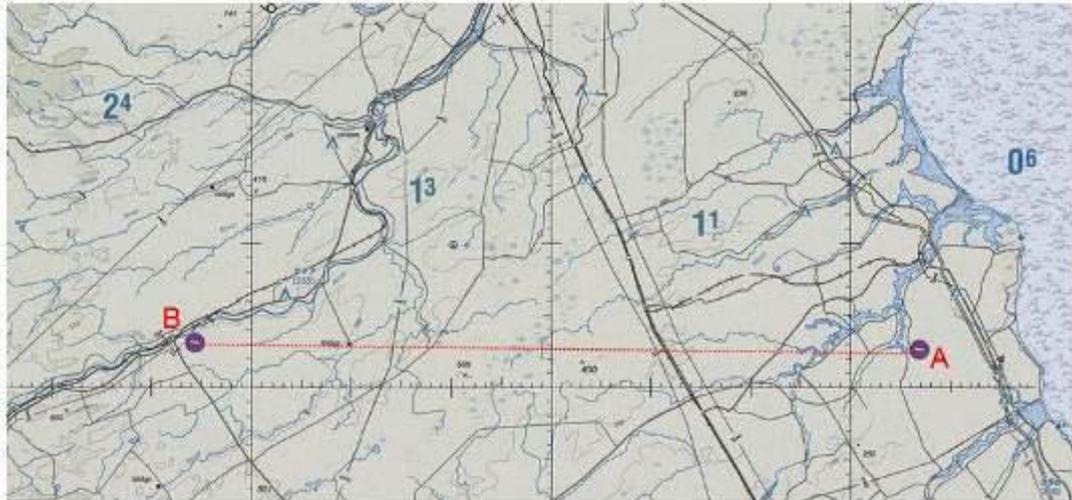
Departure (A): Wind Variable 3knots, Visibility 16km, Light Rain with Cloud base at 1800ft.

Destination(B): Wind Variable 2knots, Visibility 16km, Light Drizzle, Cloud base 1500ft.

Area Forecast

Visibility 16km with scattered light drizzle.
 Cloud broken to overcast at 1,500 feet, with cloud tops 24,000ft.
 Turbulence NIL.

All cloud heights in AGL (above ground level). Pilot traveling from Aerodrome A to Aerodrome B



The pilot decides to **go ahead** with the VFR flight.
 You later find out the flight safely landed at its intended destination.

Evaluate the pilots decision to go ahead with the flight;

Considering the information available to the pilot, how do you rate the pilots decision to conduct the VFR flight;

	1	2	3	4	5	6	7	8	9	
Very GOOD Decision	<input type="radio"/>	Very POOR Decision								

How much risk do you think the pilot took in attempting this VFR flight?

	1	2	3	4	5	6	7	8	9	
Very LOW Risk	<input type="radio"/>	Very HIGH Risk								

You wish to conduct the same flight. Considering the information available to the pilot before they departed, indicate your level of agreement with the following statement;

I would be able to **safely** conduct the flight VFR;

	1	2	3	4	5	6	7	8	9	
Strongly Agree	<input checked="" type="radio"/>	<input type="radio"/>	Strongly Disagree							

How much risk do you think the pilot took in attempting this VFR flight?

	1	2	3	4	5	6	7	8	9	
Very LOW Risk	<input type="radio"/>	Very HIGH Risk								

You wish to conduct the same flight. Considering the information available to the pilot before they departed, indicate your level of agreement with the following statement;

I would be able to **safely** conduct the flight VFR;

	1	2	3	4	5	6	7	8	9	
Strongly Agree	<input type="radio"/>	Strongly Disagree								

Section THREE (3 of 3)

Please try to imagine the following situation. A non-instrument rated pilot is conducting a VFR cross country flight in a Cessna 172. The pilot is returning to his home base which he flew from yesterday. The return flight should take just over an hour. The weather information and route map available to the pilot is as follows;

Current Weather

Departure (A): Wind 340 at 4knots, Visibility 16km, Sky Clear.

Destination (B): Wind 310 at 9knots, Visibility 16km in Light Rain, Cloud base 800ft.

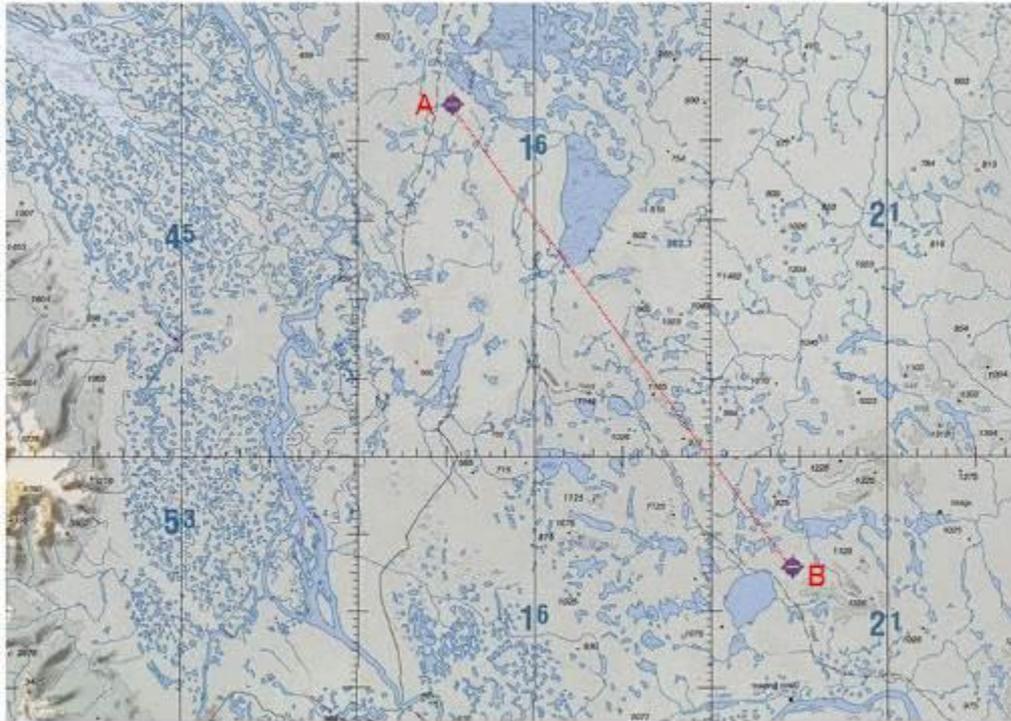
Area Forecast

Visibility reducing 10km in areas of light rain and Mist.

Cloud cover at 4000ft, with top of 20,000ft.

Turbulence NIL.

All cloud heights in AGL (above ground level). Pilot traveling from Aerodrome A to Aerodrome B



The pilot decides to **go ahead** with the VFR flight. You later find out the flight safely landed at its intended destination. Evaluate the pilots decision to go ahead with the flight;

Considering the information available to the pilot, how do you rate the pilots decision to conduct the VFR flight;

	1	2	3	4	5	6	7	8	9	
Very GOOD Decision	<input type="radio"/>	Very POOR Decision								

How much risk do you think the pilot took in attempting this VFR flight?

	1	2	3	4	5	6	7	8	9	
Very LOW Risk	<input type="radio"/>	Very HIGH Risk								

You wish to conduct the same flight. Considering the information available to the pilot before they departed, indicate your level of agreement with the following statement;

I would be able to **safely** conduct the flight VFR;

	1	2	3	4	5	6	7	8	9	
Strongly Agree	<input type="radio"/>	Strongly Disagree								

Appendix H –Human Ethics Approval for Studies 4–6



MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA

4 September 2014

Stephen Walmsley
96 King Street
CAMBRIDGE 3434

Dear Stephen

**Re: HEC: Southern B Application – 14/43
Debiasing weather related decision making**

Thank you for your letter dated 26 August 2014.

On behalf of the Massey University Human Ethics Committee: Southern B I am pleased to advise you that the ethics of your application are now approved. Approval is for three years. If this project has not been completed within three years from the date of this letter, reapproval must be requested.

If the nature, content, location, procedures or personnel of your approved application change, please advise the Secretary of the Committee.

Yours sincerely

Prof John O'Neill, Acting Chair
Massey University Human Ethics Committee: Southern B

cc Dr Andrew Gilbey
School of Aviation
PN833

Mr Ashok Poduwal, CEO
School of Aviation
PN833

Massey University Human Ethics Committee
Accredited by the Health Research Council

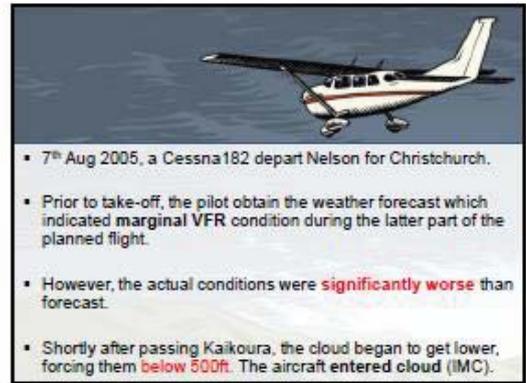
Research Ethics Office, Research and Enterprise

Massey University, Private Bag 11222, Palmerston North 4442, New Zealand T 06 3505573; 06 3505575 F 06 350 5622
E humanethics@massey.ac.nz; animalethics@massey.ac.nz; gtc@massey.ac.nz www.massey.ac.nz

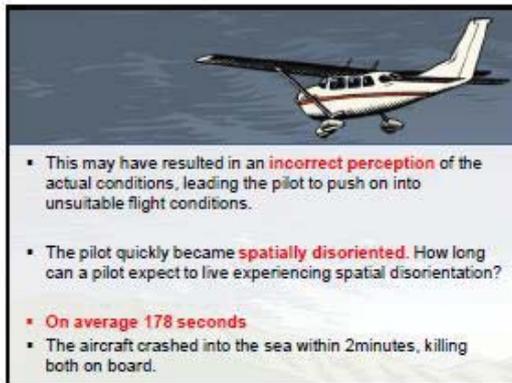
Appendix I – Debiasing Lecture Slides



Weather Related Decision Making: Cognitive Biases



- 7th Aug 2005, a Cessna 182 depart Nelson for Christchurch.
- Prior to take-off, the pilot obtain the weather forecast which indicated **marginal VFR** condition during the latter part of the planned flight.
- However, the actual conditions were **significantly worse** than forecast.
- Shortly after passing Kaikoura, the cloud began to get lower, forcing them **below 500ft**. The aircraft entered cloud (IMC).

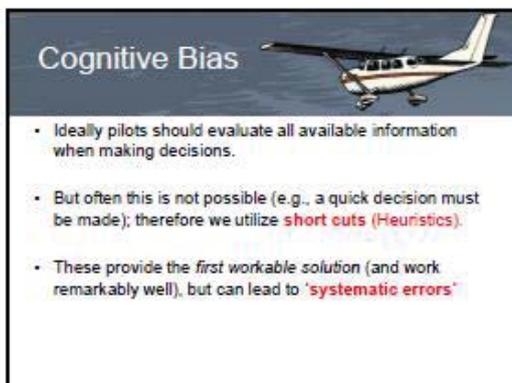


- This may have resulted in an **incorrect perception** of the actual conditions, leading the pilot to push on into unsuitable flight conditions.
- The pilot quickly became **spatially disoriented**. How long can a pilot expect to live experiencing spatial disorientation?
 - On average 178 seconds**
- The aircraft crashed into the sea within 2minutes, killing both on board.



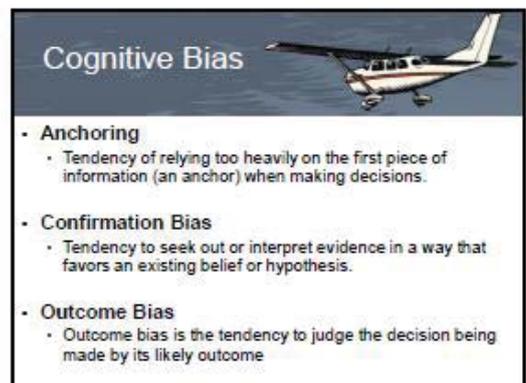
Weather Related Decision Making

- Significant danger** with VFR pilots flying into Instrument Meteorological Conditions (IMC).
- When a VFR pilot enters IMC conditions it is often down to **luck** if there is a safe outcome.
- Faulty pilot **decision making** has been shown to be considerable factor behind this type of accident.



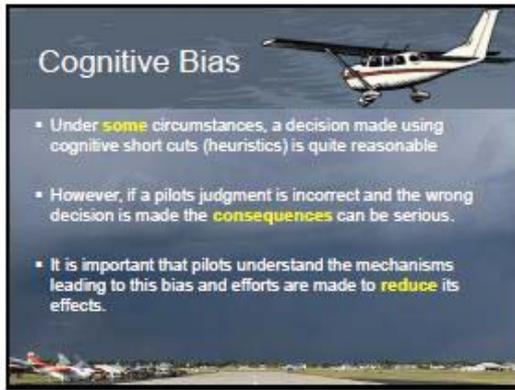
Cognitive Bias

- Ideally pilots should evaluate all available information when making decisions.
- But often this is not possible (e.g., a quick decision must be made); therefore we utilize **short cuts (Heuristics)**.
- These provide the **first workable solution** (and work remarkably well), but can lead to **'systematic errors'**



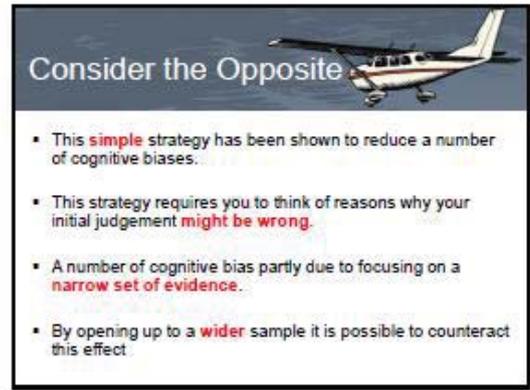
Cognitive Bias

- Anchoring**
 - Tendency of relying too heavily on the first piece of information (an anchor) when making decisions.
- Confirmation Bias**
 - Tendency to seek out or interpret evidence in a way that favors an existing belief or hypothesis.
- Outcome Bias**
 - Outcome bias is the tendency to judge the decision being made by its likely outcome



Cognitive Bias

- Under **some** circumstances, a decision made using cognitive short cuts (heuristics) is quite reasonable
- However, if a pilots judgment is incorrect and the wrong decision is made the **consequences** can be serious.
- It is important that pilots understand the mechanisms leading to this bias and efforts are made to **reduce** its effects.



Consider the Opposite

- This **simple** strategy has been shown to reduce a number of cognitive biases.
- This strategy requires you to think of reasons why your initial judgement **might be wrong**.
- A number of cognitive bias partly due to focusing on a **narrow set of evidence**.
- By opening up to a **wider** sample it is possible to counteract this effect.



Anchoring

Consider anchor **inconsistent** arguments.

- Reasons why first piece of information might be inadequate
- Reasons why forecast conditions are not what they say they are

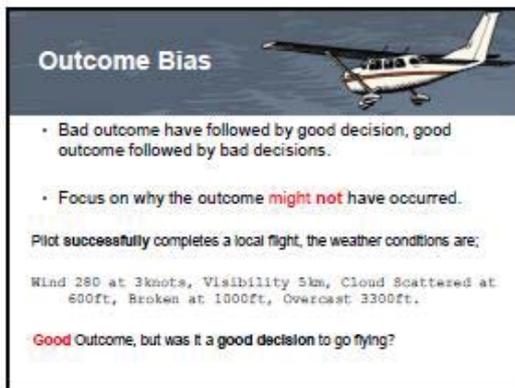
Forecast Cloud Height 2000ft, Visibility 30km.

- Why might the forecast conditions be wrong?
 - Why might the cloud be higher or lower than 2000ft
 - Why might the Visibility be higher or lower than 30km



Confirmation Bias

- Seek for evidence that would prove your initial judgment might be **wrong**.
- Just one piece of evidence require to **disprove** a hypothesis.
- Safe** to continue the flight?
 - Seek **disconfirming** evidence (reasons why it is **NOT** safe)
 - The cloud too low?
 - The cloud base is too low to the ground?
 - The visibility too low?
 - I will fly through the cloud at this height?



Outcome Bias

- Bad outcome have followed by good decision, good outcome followed by bad decisions.
- Focus on why the outcome **might not** have occurred.

Pilot **successfully** completes a local flight, the weather conditions are:

Wind 280 at 3knots, Visibility 5km, Cloud Scattered at 600ft, Broken at 1000ft, Overcast 3300ft.

Good Outcome, but was it a good decision to go flying?



QUESTIONS